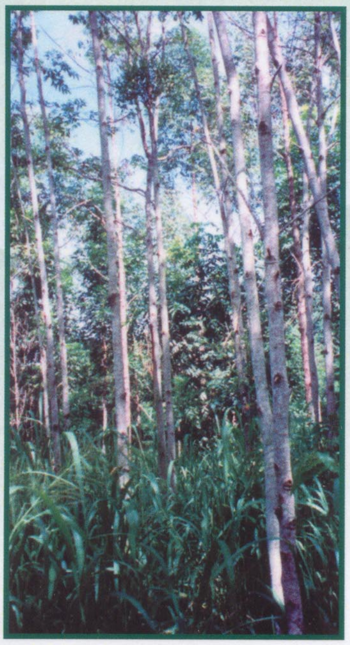


Socio-economic Research Methods in Forestry: A Training Manual

Edited by S. Harrison, J. Herbohn
E. Mangaoang and J. Vanclay

Rainforest CRC

Cooperative Research Centre for Tropical Rainforest Ecology and Management

SOCIO-ECONOMIC RESEARCH METHODS IN FORESTRY: A TRAINING MANUAL

Proceedings of an International Training Workshop held in the College of Forestry at Leyte State University, Visca, Baybay, The Philippines, over the period 4-10 February, 2002

Edited by
Steve Harrison, John Herbohn,
Ed Mangaoang and Jerry Vanclay

Rainforest CRC

ISBN 0 86443 691 2

Published by the Cooperative Research Centre for Tropical Rainforest Ecology and Management. Further copies may be requested from the Cooperative Research Centre for Tropical Rainforest Ecology and Management, James Cook University, PO Box 6811 Cairns, QLD, Australia 4870.

This publication should be cited as: Harrison, S., Herbohn, J., Mangaoang, E. and Vanclay, J (2002) *Socio-economic Research Methods in Forestry: A Training Manual.* Cooperative Research Centre for Tropical Rainforest Ecology and Management. Rainforest CRC, Cairns.

December 2002

ACKNOWLEDGEMENTS

We would like to thank all those people in the College of Forestry at Leyte State University who contributed to the arrangements and smooth running of the training workshop. Our thanks also go to the other contributors in this Training Manual, and to the workshop participants for their insightful feedback on many issues.

We would also like to thank the residents of Conalum barangay, Inopocan, for their hospitality and patience on the visit associated with this workshop, and on earlier visits, in providing us with a practical introduction to smallholder forestry.

We wish to acknowledge the financial support from Australian Centre for International Agricultural Research (ACIAR) which allowed us to hold the workshop and to sponsor the attendance of Australian and Filipino participants. The assistance of the Rainforest Cooperative Research Centre in the production of this manual is gratefully acknowledged.

Steve Harrison, John Herbohn, Ed Mangaoang and Jerry Vanclay

THE UNIVERSITY OF QUEENSLAND

CONTRIBUTORS

Mr Nicholas Emtage, PhD candidate, School of Natural and Rural Systems Management, The University of Queensland, Qld. 4072, Australia.

Dr Steve Harrison, Associate Professor of Economics, School of Economics, The University of Queeensland, Qld. 4072, Australia.

Dr John Herbohn, Senior Lecturer in Natural Systems Management, School of Natural and Rural Systems Management, The University of Queeensland, Gatton 4343, Australia.

Dr Eduardo O. Mangaoang, Associate Professor and Head, College of Forestry, Leyte State University, Visca, Baybay, Leyte, The Philippines.

Dr Jungho Suh, Postdoctoral Research Fellow, School of Economics and School of Natural and Rural Systems Management, The University of Queeensland, Qld. 4072, Australia.

Professor Jerry Vanclay, Professor of Sustainable Forestry, School of Environmental Science and Management, Southern Cross University, Lismore, NSW 2480, Australia.

TABLE OF CONTENTS

PARTICIPANTS AT THE LEYTE STATE UNIVERSITY WORKSHOP ON SOCIO-ECONOMIC RESEARCH METHODS IN FORESTRY

NAME	POSITION/DESIGNATION	AGENCY
1. Dr Paciencia P. Milan	President	LSU
2. Dr Jose L. Bacusmo	Director of Research	ODRD, LSU
3. Dr Belita A. Vega	Director of Center for Social Research	CSR, LSU
4. Dr Salome B. Bulay-og	Assoc. Professor, Ag. Econ. Dept.	DAEAB, LSU
5. Dr Remberto A. Patindol	Assoc. Professor, Stat. Dept.	DAEAM, LSU
6. Dr Pastor Garcia	Assoc. Professor	FARMI, LSU
7. Dr Eduardo O. Mangaoang	Assoc. Professor	DOF, LSU
8. For. Nestor O. Gregorio	Instuctor	DOF, LSU
9. For. Anatolio N. Polinar	Instructor	DOF, LSU
10. For. Dennis P. Peque	Instructor	DOF, LSU
11. For. Mylene D. Napiza	Instructor	DOF, LSU
12. Prof. Arturo E. Pasa	Instructor	DOF, LSU
13. Dr Ernesto C. Bumatay	Professor	DOF, LSU
DENR Region 8		
14. For. Emma M. Germano	Researcher	DENR-8
15. For Edilberto E. Nasayao	Researcher	DENR-8
PCARRD		
16. For. Reynaldo Dimla	Researcher	PCARRD
Local Government Unit/Provincial Government		
17. Rev. Manuel E. Constatino Jr.	CWMP In-charge	LGU, Isabel
18. Rodolph Hernandez	Environment Officer	ENRO, Leyte (PASAR)
NGO/Industry		
19. Louie O. Pable	Executive Director	LGM, Isabel
20. Engr Amos S Calda	Community Relations Officer	LIDE, Isabel
Research & Development Organization		
21. ICRAF Staff (3)	Research Officers	
22. LIP-GTZ Staff (1)	Research Officer	UQ, Australia
23. Dr John Herbohn	Senior Lecturer	UQ, Australia
24. Dr Steve Harrison	Associate Professor	UQ, Australia
25. Prof. Jerry Vanclay	Professor	Southern Cross University
26. Nicholas Emtage	PhD Student	UQ, Australia
27. Dr Jungho Suh	Post-doctoral Research Fellow	UQ, Australia
28. Darwin Posas	MSF Student	CF, LSU
29. Jusifer Mendoza	MSF Student	CF, LSU
30. Edwin Cedamon	Sc. Res. Asst.	ACIAR-CF

PREFACE

The training workshop on socio-economic research methods in forestry on which these papers are based was a consequence of a long-held ambition to offer such a training package, arising from research activities in the socio-economics program of the Rainforest Cooperative Research Centre. This ambition came to fruition at Leyte State University (LSU), Visca, Baybay, The Philippines, over the period 4-10 February, 2002.

The workshop – a hectic but enjoyable experience – was the first run of such a program by our research group within the ACIAR Project ASEM/2000/088 – Redevelopment of a Timber Industry Following Extensive Land Clearing. It represents part of the capacity building activity planned for this project. Other objectives of the workshop were: to allow an exchange of views on the forestry research mindset of the Australian and Filipino project members; to advance the planning of and participation in the ACIAR project; and to document the various research methods we have employed in this and other forestry research projects.

It has been a rewarding experience to be a part of an enthusiastic and active research group, and fortunately several of the members of this group were able to contribute at the training workshop. Hence a wide range of skills and interests are represented in the papers contained here. It is our hope that this training manual will be of assistance to others in planning training workshops, and to researchers in various disciplines who wish to gain skills in the particular research areas addressed here.

Steve Harrison

1. A Forester's Perspective of the Socio-economic Information Requirements for Forestry in Leyte

Eduardo O. Mangaoang

In this module, a number of observations are made about the need for socio-economic research in forestry, on the basis of experience in the College of Forestry at Leyte State University (LSU). These comments draw particularly on teaching and outreach activities of the College. They are also influenced by collaboration with international research agencies, and with other research units within LSU. The module is designed to provide a context for the subsequent presentations on research techniques.

1. BACKGROUND AND RATIONALE

In the recent years, the focus on forest management and conservation has significantly shifted from the highly technical commercial forestry to a more people-oriented social forestry orientation. Gone are the days when forestry is looked upon as solely management and utilization of trees by large-scale timber-product-oriented logging corporations to meet demands for wood and wood-based products. The more recent scenario is a paradigm shift in the forestry sector to small-scale, multiple-product-based, people-oriented and community-based sustainable forest management. This shift recognizes that sustainable forest management and conservation can only be achieved through the commitment and active participation of the local people in communities.

Experiences in Leyte Island, Philippines, reveal that local people in communities – commonly called *smallholders* – have important knowledge and skills to offer for sustainable forest management and conservation. Studies conducted by the ViSCA-Darwin Project of the College of Forestry (CF) at LSU (Lawrence *et al*. 1999), and the collaborative research undertaking of CF, LSU and the International Centre for Research in Agroforestry (ICRAF), Visayas, (CF and ICRAF 2002) have provided clear evidence of the high level of knowledge and skills and active involvement of smallholder farmers in forest management and conservation. Various development activities on people-oriented community-based forestry had been conducted by government and non-government organizations, as well as people's organizations and private institutions. Various Local Government Units (LGUs), for example, have initiated activities for the rehabilitation and protection of forest areas in their respective localities. The rise of effective people's organizations in communities with concerns for forestry and the environment is evident.

Despite the apparent shift in focus and support for people-oriented community-based forestry, until quite recently not much had been achieved in view of the expectations for sustainable forest management and conservation. This lack of progress stimulated researchers in forestry and related fields to identify the major forestry impediments and measures by which they may be overcome. At LSU, much concern had been given to the socio-economic research in forestry. The initiative undertaken by the ViSCA-Darwin Project on local knowledge and biodiversity conservation and the collaborative undertaking of CF, LSU and ICRAF on local management practices of indigenous trees, are testimonies to the importance of socio-economic research for this present-day forestry situation. Careful examination of the socio-economic factors is a major requisite in arriving at appropriate answers to the failures and shortcomings of people-oriented community-based forestry in Leyte.

2. SUGGESTED SOCIO-ECONOMIC INFORMATION REQUIREMENTS

Lessons learned from research conducted on smallholder community forestry and experiences acquired from various

community-based forestry projects indicate the need to examine carefully the key socio-economic factors that may have encouraged or discouraged local people in communities to engage in tree farming and plantation development. For example, observations in selected upland communities in Leyte reveal that small farmers would engage in tree farming if seedlings are conveniently available, and there is a clear market demand and reasonable price for the expected wood product. This section identifies the kinds of socio-economic information that are necessary in determining and analyzing factors that may make a major contribution to the promotion of farm and community forestry in Leyte.

Adoption and suitability of trees for smallholder farmers' needs

Farmers grow trees on their land and in the process are selective in the kind of trees to cultivate or maintain. Some are enthusiastic to raise fast-growing exotic trees for immediate cash income and home consumption such as for house construction purposes. Others prefer indigenous premium timber species because of durability and finishing quality. Designing measures to promote sustainable tree cultivation among smallholders therefore requires a knowledge of the kinds of forest trees they want to adopt in view of their respective needs and objectives.

People's attitude and perceptions about forest and trees

Adoption of tree farming and community forestry in a given locality depends largely on how people perceive the value and usefulness of the undertaking in their day-to-day living. Farmers who perceive that cultivation of trees is beneficial, most especially in economic sense, are more likely to engage in tree farming.

Local knowledge and skills on trees and tree management

Smallholders have important knowledge and skills about trees and their management that scientists need to know. This arises from passed-down knowledge, years of experience, informal experimentation, and discussions with other farmers. Promotion of tree cultivation on farms can be easily attained if forest researchers and extension agents start doing things with what the farmers know and starting from what they initially have. It is vital to recognize and draw upon the store of knowledge they have to share on sustainable forest management and conservation.

Smallholders' awareness and understanding of forest policies, rules and regulations

Smallholders' desires and eagerness to cultivate trees on farms are often hindered by their perceptions about inconsistencies in government policies, rules and regulations. Many believe mistakenly that, for example, harvesting and marketing of planted indigenous trees on farms is prohibited by law. This policy, in their thoughts, is not supportive of the government's intention to encourage small farmers to engage or participate in the nation's reforestation program. Comprehending better the smallholders' level of awareness and understanding about forest rules and regulations can pave the way for development agencies to develop simple and effective mechanisms that will help farmers to access and understand better the relevant policies governing tree farming and community forestry.

Processing, marketing and pricing of tree farm products

The economic aspect of tree farming and community forestry is an important farmers' consideration as to why they would engage in the venture. Experienced farmers in Inopacan, Leyte, for example, are interested to grow trees on farms because of the perceived substantial amount of cash income that they can realize upon harvest, which may be enough to support their daily needs and to send their children to school. The amount of cash income that farmers can realize in due time, however, is dictated by the kind and quality of wood products that they can produce, together with how these are valued in the market. Often, the opportunity to improve quality and value of wood products sold in the market is not

accessible to the smallholders.

Local policies and arrangements regarding tree growing, harvesting and marketing

In upland communities, a considerable amount of interesting information is available about policies and arrangements formulated and agreed upon by the local people in relation to the growing, harvesting and marketing of tree products. Access to this information may be beneficial not only to researchers, but also to development workers and policy makers. These local policies and arrangements may have played a key role in areas where the forest environment has been sustainably managed and protected over time.

Financial requirement of tree-growing versus smallholders' financial capability

One of the major constraints against smallholders engaging in tree farming is the financial requirement of establishing and maintaining the stand during the first about five years. Detailed information about the cost of establishing a hectare of tree farm and how much cost an average smallholder can shoulder in doing so are critical information. These are relevant to designing programs that can support smallholders' desires for tree cultivation. For example, in a given community forestry project where people's participation is indispensable, one can have a clear idea about appropriate sharing of resource responsibility among partners (development organization and local people) if they have a sound knowledge about financial requirements of tree farming and what people can afford to offer.

Alternative source of livelihood

Tree farming is an attractive business venture but the long wait before the tree crop can be harvested makes smallholders hesitant to commit their scarce land, labour and cash resources. Alternative sources of funds to support day-to-day living needs is usually the common demand of people in communities. Since the kind of alternative livelihood demanded varies from one locality to another, information is needed on people's desires and preferences for alternative livelihood activities while engaging in tree farming.

3. DISCUSSION

The Philippines, including Leyte Island, has undergone large-scale deforestation, such that timber harvesting from native forests has disappeared, and only a fledgling farm and community-based timber industry now exists. A priority area for forestry research is therefore the promotion of reforestation, for timber production and environmental purposes. The focus of forest management in recent years has switched to smallholder forestry. This module has identified the kinds of socio-economic information that will assist in the promotion of farm and community forestry in Leyte. Research into other areas of socio-economic research, timber supply chain analysis, and wood processing and marketing, will no doubt be important. However, the current priority is to increase the level of planting, within forestry systems that integrate with smallholder farming and meet socio-economic and environmental sustainability requirements.

REFERENCES

CF and ICRAF (College of Forestry, Leyte State University and the International Centre for Research in Agroforestry, Visayas) (2002), Local knowledge on indigenous trees in Central Philippines, Summary report submitted to ICRAF Tree Domestication Program, Southeast Asian Regional Office, Bogor, Indonesia.

Lawrence, A.L., Mangaoang, E.O. and Barrow, S. (1999), *Foresters, Farmers and Biodiversity: New Issues for the Forestry Curriculum*, Proceedings of the National Workshop on Local Knowledge and Biodiversity Conservation in Forestry Practice and Education, held on October 19-23, 1998 at ViSCA, Baybay, Leyte.

2. Socio-economic Research Techniques in Tropical Forestry

Steve Harrison

A wide variety of techniques from the areas of economics, management, sociology and planning are relevant for carrying out research into non-industrial forestry. This module introduces some of the more widely used of these techniques, which are discussed in more detail in later modules. On the basis of experience in a forestry research group, comments are made about relevant applications of these research techniques.

1. INTRODUCTION

To provide input to forest policy, it is necessary understand the priority that different groups within the community place on forestry, and the costs and benefits of various kinds arising from forestry. There is a need to generate information which will assist decision makers, in government and the private sector, in regard to forestry investment and management. Government is concerned with strategies to promote forestry for timber production and wider community benefits, and also to regulate the use of native forests and plantation land for the public good. Investors and landholders seek to manage their resources in forestry in such a way as to make a satisfactory return on their investment and effort and promote their livelihood.

Forestry training has traditionally concentrated on large-scale production and silvicultural technology, relevant to the maximization of forest biomass production, from monoculture plantations of a few well-known conifer species. A policy interest in small-scale forestry – particularly farm and community forestry – is relatively new. In these areas, financial aspects usually remain important, but there is also considerable emphasis on social and environmental objectives. Research in this situation requires a different mindset to that of traditional silviculture.

Another point of departure in this module is that it is concerned with forestry in the tropics. Some broad differences can be recognized between tropical and temperate forestry. In particular, reforestation is in most cases of relatively recent origin in the tropics, and the challenge is to induce more landholders to plant trees. In contrast, in some European countries, forests have been in the hands of the same families for hundreds of years, and approach the concept of a *normal* or even-aged forest. In such cases, the research interest may then be in monitoring economic performance of forestry and its relationship with other farm activities, harvest scheduling and marketing strategies, and intergenerational transfer of forest resources.

A wide variety of economic and other social science research techniques are available for tropical forestry research. These approaches are the central theme of this module, although it is necessary to place them in a decision-support perspective. Material in this module on socio-economic research techniques in small-scale forestry arises in particular from activities carried out in the Rainforest Cooperative Research Centre in Queensland, and draws on Harrison (2001).

This module first outlines the nature and role of socio-economic analysis, with particular emphasis on natural resource and environmental economics. The role of government in forestry in then examined, including the instruments which governments use to support farm and community forestry. Next, various research techniques are outlined briefly, and comments are made about the relevance of these techniques to address particular planning and management issues. Finally, some speculation is made about research issues and methods which are likely to figure prominently in the future. A brief discussion section follows.

2. THE NATURE OF SOCIO-ECONOMIC RESEARCH IN RELATION TO SMALL-SCALE FORESTRY IN THE TROPICS

Plantation forestry as a long-term investment, usually with the predominant objective of producing lumber and other wood products for sale, and providing uncertain financial returns to growers. It also generating positive externalities for the community. In the case of small-scale forestry in particular, plantations are often grown for multiple objectives. The management of native forests, on the other hand, is increasingly dominated by environmental objectives. For both plantation and native forests, government has an important role to play, though largely as a facilitator in the former case and as a regulator in the latter case.

Formulation of appropriate government policy has high information needs. Economics is one the disciplines which can provide policy-relevant information. The orientation of economics is essentially *anthropocentric*, i.e. management policies are examined in terms of their contribution to satisfying the needs and wants of humankind.

A distinction is made between *positive* and *normative* economics. Positive (or 'what is' economics) attempts to explain how economic systems work, by hypothesising and testing theories on the basis of empirical evidence. Normative (or 'what aught to be') economics attempts to analyses consequences of alternative courses of action in terms of specific objectives, and hence derive prescriptions for management. Sometimes the term *conditional normative* is used, in that optimal management policies are prescribed, conditional on specified objectives, constraints and environments.

A distinction is sometimes drawn between *financial* and *economic* analysis. Financial analysis takes account of costs and returns which can be identified through market transactions. Economic analysis takes a broader perspective, and includes for example the opportunity cost the unpaid inputs, e.g. of the labour and financial inputs of managers of small firms including farms. The term 'social economic' is also used sometimes, e.g. in social cost-benefit analysis. This does not refer to combination of economics and sociology, but rather to making allowance for impacts beyond the boundaries of the firm, i.e. typically called spillover effects or externalities. This would include impacts of environment and communities, viewed in terms of their (unpriced) costs and benefits.

Within the discipline of economics, there are many areas of specialisation. In broad terms, these involve macroeconomics (dealing with national economies) and microeconomics (dealing with firms and industries). At a more specific level, there are specialisations in international economics, monetary economics, development economics, transport economics, health economics, econometrics, and so on. During the last 20 years, there has been rapid development of the fields of natural resource and environmental economics. These are essentially areas of applied microeconomics, although they also draw on macroeconomics (e.g. in relation to green national accounts) and development economics. Within natural resource economics, there is further specialisation, e.g. into land economics, forestry economics, fisheries economics, water resource economics, energy economics, and the economics of outdoor recreation. The areas dealing with agriculture, forestry, fisheries and wildlife (a sub-group of the renewable resources) are sometimes referred to as bioeconomics. In the last decade, ecological economics has also emerged as a recognizable economics specialisation.

The discipline of natural resource management has also undergone rapid development in the last 20 years. This is a field concerned with economics, geography, conservation biology and other areas, but with a focus on improving the performance of bioeconomic systems.

Economists and natural resource management professionals working in the fields of agriculture and forestry require a sound understanding of biology. Often they will work in *multidisciplinary* teams, which also comprise specialists in other fields, including the natural sciences. No one discipline has a mortgage on all the

knowledge and analysis skills necessary to guide efficient, equitable and sustainable management of natural resources. If research is carried out from a multidisciplinary (or even interdisciplinary) perspective, then managers can have more confidence that the advice provided to them will be sound. There is a lower risk that decisions will be criticised for overlooking important perspectives. Sometimes it is argued that what is needed is a *transdisciplinary* team, in which members are multi-skilled. While this has obvious benefits, it is also necessary that disciplinary rigour be applied in research.

While early forestry research tended to take a *reductionist* approach of investigating specific relationships, there is now strong emphasis on a systems philosophy, which takes an *holistic* approach to defining and examining forestry systems.

Any structured research activity can be viewed in the context of the scientific method of inquiry, in that a set of implicit steps can be recognized. In general terms, these include identifying a decision problem and an associated system (e.g. a collection of resources and their interactions), posing questions or making hypotheses about this system, collecting data, and analysing these data to draw inferences.

In economics, there has been a tendency to rely heavily on *quantitative* research techniques, and mathematical and statistical models, particularly in North America. However, in recent years there has been increased interest in *qualitative* techniques, the usefulness of which has been clearly demonstrated in other social sciences (e.g. see Patton 1990).

When collecting and analysing social and economic data (as indeed in most technical research), it must be recognized that while *objectivity* is sought, nevertheless a good deal of *subjectivity* will be involved. That is, to some extent researchers subjectively chose their research topics, sites and methods, and the data they will collect, through to subjectively choosing the significance level at which they will declare significant differences.

3. THE ROLE OF SOCIO-ECONOMIC RESEARCH

Governments play a major role in forest industries, and particularly non-industrial forestry. This includes support programs to enhance the profitability of forestry and the attitudes of potential tree growers, and to overcome various impediments to tree planting. It also includes the regulation of forestry, and designed to achieve environmental and social sustainability.

In general, the objective of economists and natural resource management professionals is to provide information to agencies and individuals who have the task of managing natural resources. That is, they produce quantitative or qualitative information, within a normative framework, which fulfills a decision-support role. This information augments (but does not replace) the information, judgments, hunches and tentative decisions of managers.

Socio-economic research is in general designed to support government policy-making, i.e. a social rather than private-producer perspective is taken. The research output augments the information that decision-makers already have, confirming or challenging tentative decisions. The aim is to generate qualitative and quantitative information with will assist in the making of good decisions from the viewpoint of the wider community. The spatial focus may be on a community, district, watershed or wider region.

Socio-economic research provides one of the inputs to government policy. It moves beyond technical considerations and examines the impact of forestry policies on people and on the environment. In the case of environmental impacts, these may be assessed by scientists, but again the view of the community towards environmental issues is critical to policy. For example, if deforestation is leading to flooding of coastal cities, sedimentation of fisheries and depletion of groundwater supplies, then these issues very much affect human settlements and place high priority on reforestation.

Desirable policies are not static, but rather evolve over time, for example as

technology, environmental conditions and community attitudes change. Governments do not automatically support environmental and social causes, but need to be convinced or pressured into recognizing their worth. Strong environmental lobbies are now well established, and do not appear to be losing their influence. In some cases, 'green parties' have gained considerable power with parliaments. In low-income developing countries, environmental pressure often comes for external loan agencies and donors.

The emphasis of this module is on research into small-scale forestry in the tropics and sub-tropics, and this has some differences from small-scale forestry in developed countries in temperate zones. Typically, there is a history of relatively recent deforestation, and recognition of the critical need for reforestation. Forestry activities typically involve reforestation of agricultural land, rather than management of a relatively 'normal' forest. There is greater emphasis on impediments to planting, benefits of reforestation (including carbon sequestration and watershed protection benefits) and government assistance programs, and less on financial monitoring and harvesting systems.

4. SOCIO-ECONOMIC RESEARCH METHODS IN SMALL-SCALE FORESTRY

A wide variety or research techniques, both qualitative and quantitative, have been applied to examine socio-economic issues in small-scale forestry. Some of the more important of these are listed in Table 1.

When a problem in relation to natural resource management has been defined, it is important to determine who are the affected parties of *stakeholders* who a vested interest in policy outcomes (Harrison and Qureshi 2000). It is then necessary to collect data which will shed light on resource management. If appropriate experts can be identified, it may be possible to collect data by discussion with these experts and from informal surveys not using a probability sampling framework. Considerable use was made of informed sources in the study of potential for adoption of Australian tree species in the Philippines reported by Harrison and Herbohn (2000).

More formal methods can be applied to elicit expert opinion. *Delphi* surveys seek to obtain group consensus views while minimizing the interactions between experts so as to prevent domination on the basis of personality or rank. This was used to predict harvest ages and stand yields to be used in financial modeling of non-traditional tree species in farm forestry in North Queensland, Australia (Herbohn *et al*. 1999, Herbohn and Harrison 2000). *SWOT* analysis (group identification of strengths, weaknesses, opportunities and threats) is sometimes used when evaluating a specific program or enterprise and exploring improvement measures, e.g. see Hobbs *et al*. (2001). *Focus group* meetings are a means of generating and testing ideas as an aid to further analysis, such as setting up scenarios for non-market valuation.

Case studies of particular systems (e.g. nurseries, Community Based Forest Management (CBFM) sites, timber processing plants, lumber dealers, furniture manufacturers) may provide important insights into the role and management of these systems. There has been increasing recognition of these qualitative research techniques in recent years, particularly when a relevant population of substantial size is not available from which to draw data (e.g. see Patton 1990).

The various methods of participatory rural appraisal (PRA) also fit within this qualitative research sphere. PRA has been found useful in designing research programs in relation to small-scale forestry in the Philippines (e.g. Singzon *et al*. 1993). PRA is '*a systematic, semi-structured approach and method of assessing and understanding . . . village situations with the participation of the people and through the eyes of the people*. It comprises a rich menu of visualisation, interviewing, and group work methods that have been proven valuable for understanding the local functional values of resources, for revealing the complexities of social structures and for mobilizing and organizing local people. It is therefore a family of methods and approaches to enable local people to present, share, and analyze their

knowledge of life and conditions, to plan, to act, monitor, and evaluate' (PROCESS Foundation 1996, p. 2)[1]. Pratt (2001) provided a recent evaluation of PRA applications in Nepal.

An alternative to these qualitative methods is to identify a target or reference population and develop a sampling frame (containing as comprehensive representation of this population as possible), and conduct a sample survey. For example, landholder surveys are used to obtain information about attitudes and impediments to small-scale forestry.

While these 'stated' opinions may not match exactly the real decisions of landholders, attitude surveys are much easier to carry out than observing actual behaviour, and in practice a combination of both may be the optimal approach. For sample surveys, it is necessary to choose a *sampling design* (often a form of stratified or multistage sampling), and to develop and test a questionnaire.[2] Sometimes a semi-structured approach will be useful, where some questions are of closed format (e.g. recall a specific fact, respond yes or no, state a degree of belief or preference on a Likert scale), some are open ended (e.g. list reasons for a belief) and some allow free-form discussion (which may be voice recorded for later qualitative analysis).

Survey results are typically entered onto a computer package. Electronic spreadsheets have become extremely popular for data entry and storage, and for deriving relatively simple *descriptive statistics* (frequency distributions, means and variances) and making presentation (line and bar graphs, pie diagrams). Spreadsheets also have some capability for *statistical inferences* to be made concerning the underlying population. The Statistical Package for the Social Sciences (SPSS) has proved useful for more complex analysis, e.g. cluster analysis for identifying distinctive groups of landholders in terms of their attitudes to tree planting (e.g. Emtage *et al*. 2001). A variety of statistical time series analysis techniques have been developed, which are powerful methods of explaining observed data and making forecasts of future values of variables. Scenario development is designed to describe and understand what future types of situations might arise.

Over about the last 20 years, non-market valuation techniques have become widely applied in forestry research. The (zonal) travel cost method (TCM) allows demand to be estimated for recreation sites. The hedonic price method involves a multivariate analysis of past transaction records to determine the relationship between the market value of an asset and its characteristics, including environmental charactersitics such as freedom from pollution and good views. In this way, the value placed by the market on environmental characteristics can be estimated statistically. The contingent valuation method (CVM) and choice modeling (environmental choice modelling, choice experiments) are used to estimate total economic value (TEV) of natural assets such as forests, including use and non-use values. CVM has been controversial due to the large number of potential biases and apparently unrealistically high values obtained in some applications. Benefit transfer – inference of values from a source site to target site – provides a time-saving alternative to making new estimates for each specific new site. In practice, benefit transfer methodology is the most widely used approach to non-market valuation, and is being supported by development of databases of environmental values (e.g. see Morrison 2001).

Agencies concerned with forest management have a reporting responsibility to government in regard to the achievements from spending public funds.

---

[1] PROCESS Foundation (1996) observed that PRA draws on five traditions – activist participatory research, agroecosystem analysis, applied anthropology, field research in farming systems, and rapid rural appraisal – in an attempt to achieve 'people empowerment' and avoid the mistakes of 'rural development tourists'.

[2] Some frequently misused terms are 'survey', 'questionnaire', 'sample' and 'observation'. Typically, a single survey is carried out, in which there are a number of sample members, to each of which is administered a single questionnaire. The sample consists of a subset of members from the population; a number of observations of each variable under investigator make up a single sample to be used in a single survey.

Table 1. Socio-economic research methods in small-scale forestry

Data collection	Stakeholder analysis Elicitation of expert opinion (including consultations with experts, SWOT analysis, the Delphi method, focus groups) Case studies Participatory rural appraisal Sample surveys using probability sampling
Data analysis	Analysis of survey data – descriptive statistics Multivariate analysis (including cluster analysis and factor analysis) Price forecasting (time series models) Scenario analysis
Non-market valuation	Valueing non-wood forest products and services Evaluation of forest recreation benefits using the travel cost method Estimation of total economic value – the contingent valuation method Choice modelling or choice experiments The hedonic price method Benefit tranfer
Reporting	Reporting systems for forest enterprises and agencies
Physical and financial modelling	Stand yield modelling (including under sparse data) The optimal economic rotation (the Faustmann formula) Discounted cash flow analysis and sensitivity analysis Development of financial models of forestry investments and overall enterprises Modelling carbon sequestration Cost-benefit analysis (CBA) and cost-effectiveness analysis (CEA) Risk or venture analysis
Watershed and regional modelling	Geographical information systems (farm, watershed and regional level) Interindustry input-output analysis Transhipment modelling (locational efficiency and logistical analysis) Multicriteria analysis (and the analytic hierarchy process) Resource allocation models – linear programming Resource allocation models – goal programming Regional development models
Policy analysis	Synthesis of policy directions (transferring research to policy)

While financial outcomes are normally reported, it is only in recent times that serious attempts are being made in environmental reporting (Herbohn 2000), and these reporting systems are still very much at a research stage.

A special type of forestry reporting – that of carbon accounting, reporting and monitoring – is now under development and will become critical if carbon sequestration credits from plantation forestry become available and subject to trade (e.g. see Lamb 2000).

A variety of modelling approaches are used by researchers in relation to small-scale forestry. On the physical side, it is necessary to generate estimates of the yield of woodlots and small border plantings. This can involve particular difficulties in the case of small-scale forestry, where yield observations upon which to base modelling are scarce and performance is generally considerably

below that of research trial and commercial plantation yields. This problem is compounded when non-traditional species are grown. Stand yield prediction in such situations using the Chapman-Richards model has been employed by Venn *et al*. (2000).

The Faustmann optimal economic rotation model (Pearse 1990) provides an appropriate economic framework for estimating the returns from forestry and comparing the economics of alternative forestry systems. Financial modelling – which requires yield and price estimates, and is usually carried out using a spreadsheet package and applying discounted cash flow functions – allows the payoff from forestry to be predicted; this may be restricted to modelling of a forestry enterprise or involve modelling the overall business undertaking (e.g. farm business or community development program). This analysis could be extended to determining the financial performance and rent sharing at various stages of the timber production pipeline, including milling and timber merchandising.

Economic modelling may be viewed as an extension to financial modelling, in which an effort is made to include shadow prices rather than simply market prices, and in which non-market values are included, typically in an extended cost-benefit framework. Where a specific objective is to be achieved (e.g. the planting of 200 ha of community forestry) or performance can be measured in specific units (costs per hectare planted), cost-effective analysis (CEA) rather than CBA may be employed, thus averting the need for estimation of economic benefits.

Since forestry is a long-term enterprise, with uncertain stand yield and future timber price, some form of risk analysis is normally attached to the financial analysis, as a *sensitivity analysis* or a *risk simulation*, say using the @RISK simulation add-on to Excel or Lotus 1-2-3 (Harrison *et al*. 2001).

A variety of approaches have been developed for forestry planning at a regional level. Methods of spatial modeling using geographical information systems (GIS) are being used increasingly in socio-economic studies, particularly with regard to conservation objectives. The application of GIS in riparian revegetation studies is reviewed with a case study by Harrison *et al*. (1998).

Inter-industry input-output analysis is designed to estimate the impacts of a change of expenditure (e.g. large on-off investment) in an enterprise, and yields various types of 'multipliers' (income, output and employment) which are indicators of community benefit from the investment. Multiplier values for a 'normal forest' are reasonably well established, but when it comes to reforestation the evidence is sparse. In practice, the upstream investment in tree planting (from expenditure on inputs and the consequent economic activity driven by this expenditure) is usually modest, with a long delay to harvesting, so that the multipliers tend to be quite small (e.g. Todd *et al*. 1999), raising questions about the usefulness of estimating multipliers.

When dealing with a number of stakeholder groups with often conflicting objectives, it has become apparent that the method of analysis should take account of multiple goals. In recent years, there has been much interest in multicriteria analysis (MCA) or multi-objective decision-support systems (MODSS) as an approach to planning landuse at a catchment level, including for reforestation planning (e.g. RAC 1992; Robinson 2000; Qureshi and Harrison 2001). The analytic hierarchy process (Saaty 1995) is sometimes used to elicit stakeholder preference weights in relation to various goals in MCA (e.g. see Harrison and Herbohn 2002). MCA and MODSS approaches are able to take into account the preferences of the various stakeholder groups, to utilize both quantitative and qualitative information, and are reasonably rapid to apply. A criticism can be the high level of subjectivity involved.

One form of MCA which can be carried out using mathematical programming software is *goal programming*. Here a number of aspiration levels for specific goals are set and the objective is to minimize the sum of appropriately weighted shortfalls (weighted goal programming) or to select activities such that the goals are satisfied in priority

order (lexicographic or preemptive goal programming). Goal programming is being applied by Venn to compare alternative forest utilization policies by the indigenous community in Cape York in Australia.

Mathematical programming has further applications in forestry, such as harvest scheduling and determining the least-cost transport allocation between supply origins and demand destinations and the optimal location of processing facilities.

Devising strategies for small-scale forest industry development in any particular region is a challenging task. Theoretical foundations for this kind of analysis are provided by Tykkyäinen *et al*. (1997). The FLORES model of Vanclay *et al.* (2000) is an attempt to develop structured methodology for examining the requirements for more rapid adoption of small-scale forestry, and is to be trialed in the Philippines. In contrast, the collection of papers of Hyttinen *et al.* (1996) examine the role of forestry in regional development.

The output of research has to be communicated to policy makers, and taken up by them, if it is to have practical outcomes. This requires assembling the information as an integrated package which can be comprehended by officers of the relevant administrative staff, and is viewed by them as sensible and politically acceptable.

5. DISCUSSION

Socio-economic analysis of forestry systems has been a neglected research area. It is unlike silvicultural research, and draws on the techniques of the social scientist, recognizing the community setting and multi-goal nature of small-scale forestry. Particular issues arise in the tropics and sub-tropics, where reforestation is urgently needed following extensive forest logging and clearing. There is a severe lack of information about the performance of non-traditional species and mixed-species plantations. Small-scale forestry is found to face a large number of constraints and to present a wide variety of policy issues. Critical amongst these issues is how to promote new forest industries, the role of government (if any) in supporting farm and community forestry investments, and the potential contribution of venture capital, bank loans, NGO funds, and ethical or green investments. A number of innovative methods appear to offer promise for enhancing the economic attractiveness of small-scale forestry.

A wide variety of research approaches are available in relation to small-scale forestry, with techniques of the economist, sociologist, environmentalist and planner found relevant to analysis of forestry issues, to generate information for policy-makers. Economics provides both a conceptual framework for analysis (e.g. the cost-benefit analysis framework) and specific tools (e.g. discounted cash flow analysis, goal programming, transshipment models, interindustry input-output analysis). Some approaches consist of a diverse collection of related techniques, e.g. MODSS. Some are tools which allow models to be developed, e.g. the simulation modeling computer languages such as Simile (used in development of FLORES).

Further improvement in – and greater acceptability of – socio-economic research methods can be expected in the future. Attention may be paid to refinement of research methodology. It is likely that there will be greater integration of methods of analysis, e.g. of MCA, financial and discounted cash flow (DCF) analysis and GIS, to development a more powerful research approach.

The need to have sound information to support policy for small-scale forestry can be expected to intensify, particularly as governments appear to be reducing the level of support for tree planting. While it is critical to appreciate the various research techniques, any research program needs to be viewed in a broader context, including research objectives, project design, team building and project management, and the policy context in which results will be viewed.

REFERENCES

Emtage, N.F., Harrison, S.R. and Herbohn, J.L. (2001), 'Landholder Attitudes to and Participation in Farm Forestry Activities in Sub-Tropical and Tropical Eastern

Australia', in S.R. Harrison and J.L. Herbohn, eds., *Sustainable Farm Forestry in the Tropics: Social and Economic Analysis and Policy*, Edward Elgar, Cheltenham, pp. 195-210.

Harrison, S. (2001), 'Research Approaches to Environmental-economic Issues in Small-scale forestry', in A. Niskaanen and J. Vayrynen, eds., *Economic Sustainability of Small-Scale Forestry*, EFI Proceedings No. 36, European Forest Institute, Joensuu, pp. 201-214.

Harrison, S.R., Herbohn J.L. eds. (2001), *Socio-Economic Evaluation of the Potential for Australian Tree Species in the Philippines*, Australian Centre for International Agricultural Research, Canberra.

Harrison, S.R. and Herbohn, J.L. eds. (2002), *A Whole-Farm and regional Agroforestry Decision-Support System*. Rural Industries Research and Development Corporation Research Report Series, Canberra.

Harrison, S.R., Herbohn, J.L. and Emtage, N.F. (2001), 'Estimating investment risk in small-scale plantations of rainforest cabinet species and eucalypts', in S.R. Harrison and J.L. Herbohn, eds., *Sustainable Farm Forestry in the Tropics: Social and Economic Analysis and Policy*, Edward Elgar, Cheltenham, pp. 47-59.

Harrison, S.R. and Qureshi, M.E. (2000), 'Choice of Stakeholder Groups and Members in Multicriteria Decision Models', *Natural Resources Forum*, 24(1): 11-19.

Harrison, S.R., Qureshi, M.E., Sharma P.C. and Tidey, M.E. (1998), 'Using GIS in Multicriteria Analysis of Riparian Revegetation Options', in P. Lai, Y Leung and W. Shi, eds., *Proceedings of the International Conference on Modelling Geographical and Environmental Systems with Geographical Information Systems*, Chinese University of Hong Kong, Hong Kong, Vol. 1, pp. 216-222.

Herbohn, J.L. and Harrison, S.R. (2000), 'Assessing Financial Performance of Small-Scale Forestry', in S.R. Harrison, J.L. Herbohn and K.F. Herbohn, eds., *Sustainable Small-scale Forestry: Socio-economic Analysis and Policy*, Edward Elgar: Cheltenham, pp.37-49.

Herbohn, J.L., Harrison, S.R. and Emtage, N. (1999), 'Potential Performance of Rainforest and Eucalypt Cabinet Timber Species in Plantations in North Queensland', *Australian Forestry*, 62(1): 79-87.

Herbohn, K.F. (2000), 'Accounting and Reporting of Forest Enterprises', in S.R. Harrison, J.L. Herbohn and K.F. Herbohn, eds., *Sustainable Small-scale Forestry: Socio-economic Analysis and Policy*, Edward Elgar, Cheltenham, pp. 104-119.

Hobbs, M, Hytönen, L. and Kangas, J. (2001), *Factors Affecting the Economic Sustainability of the Non-industrial Private Forest Enterprise: A Comparison of Stakeholder Perceptions*, paper presented at the International Symposium on Economic Sustainability of Small-scale Forestry, IUFRO Group 3.08.00, Joensuu, Finland.

Hyttinen , P., Mononen, A. and Pelli, P. eds. (1996), *Regional Development Based on Forest Resources – Theories and Practices*, EFI Proceedings No. 9, European Forest Institute, Joensuu, Finland.

Lamb, K. (2000), 'Carbon-based Marketing Opportunities for Small-scale Farm Forestry', in S.R. Harrison, J.L. Herbohn and K.F. Herbohn, eds., *Sustainable Small-scale Forestry: Socio-economic Analysis and Policy*, Edward Elgar, Cheltenham, pp. 89-103.

Morrison, D. (1991), 'Non-market Valuation Databases: How Useful Are They?', *Economic Analysis and Policy*, 31(1): 33-56.

Patton, M.G. ed. (1990), *Qualitative Evaluation and Research Methods*, 2nd edn., Sage Publications, Newbury Park.

Pearse, P.H. (1990), *Introduction to Forestry Economics*, University of British Columbia Press, Vancouver.

Pratt, G. (2001), *Practitioners Critical Reflections on PRA and Participation in Nepal*, IDS Working Paper 122, Institute of

Development Studies, University of Sussex, Brighton.

PROCESS Foundation (1996), *Training on Participatory Rural Appraisal: A Training Manual*, Tagbilaran City, Bohol.

Qureshi, M.E. and Harrison, S.R. (2001), 'A Decision Support Process to Compare Riparian Revegetation Options in Scheu Creek Catchment in North Queensland', *Journal of Environmental Management*, 62(1): 101-112.

RAC (1992), *Multi-Criteria Analysis as a Resource Assessment Tool*, Resource Assessment Commission, *RAC Research Paper No. 6*, Canberra.

Robinson, J. (2000), 'Does MODSS offer an alternative to traditional approaches to natural resource management decision-making?', *Australian Journal of Environmental Management*, 7(3): 170-180.

Saaty, T.L. (1995), *Decisions making for leaders: The analytic hierarchy process for decisions in a complex world*, RWS Publications, Pittsburgh.

Singzon, S.B., Baliña, F.T., Gabunada, F. M. and Morales, N.O. (1993), Tropical Participatory Rural Appraisal about Trees and Tree Planting Activities of Farmers in Matalom, Leyte, Philippines, Farm and Resource Management Institute, ViSCA, Baybay, Leyte, Philippines.

Tykkyäinen, M., Hyttinen, P. and Mononen, A. (1997), 'Theories of regional development and their relevance to the forest sector', *Silva Fennica*, 31(4): 447-459.

Todd, C.R., Loane, I.T. and Ferguson, I.S. (1997), *Potential Impact of a Farm Forestry Industry on the Goulburn Regional Economy*, Trees for Profit Research Centre, Bulletin No. 10, School of Forestry and Resource Conservation, University of Melbourne.

Vanclay, J.K., Muetzelfeldt, R., Haggith, M., and Bousquet, F. (2000), *FLORES: Helping people to realize sustainable futures*, XXI IUFRO World Congress 2000, Forests and Society: The Role of Research, Sub-plenary Sessions, Volume 1, pp. 723-729.

Venn, T.J., Beard R.M. and Harrison S.R. (2000a), 'Modelling Stand Yield of Non-Traditional Timber Species Under Sparse Data', in S.R. Harrison and J.L. Herbohn, eds., *Socio-Economic Evaluation of the Potential for Australian Tree Species in the Philippines,* Australian Centre for International Agricultural Research, Canberra, pp. 55-77.

Venn, T.J., Harrison, S.R. and Herbohn, J.L. (2000b), 'Impediments to Adoption of Australian Tree Species in The Philippines', Ch. 14 in Harrison, S.R. and Herbohn, J.L., *Socio-Economic Evaluation of the Potential for Australian Tree Species in the Philippines,* Australian Centre for International Agricultural Research, Canberra, 167-181.

Vize, S.M. and Creighton, C. (2001), 'Institutional Impediments to Farm Forestry', in S.R. Harrison and J.L. Herbohn, eds., *Sustainable Farm Forestry in the Tropics: Social and Economic Analysis and Policy*, Edward Elgar, Cheltenham, pp. 241-255.

3. Data Collection Methods in Forestry Socio-economic Research

John Herbohn

Data collection is an important component of the research process, yielding facts, figures and views which are subsequently analysed and manipulated in various ways. This module outlines some useful data collection techniques for socio-economic data – including surveys of individuals in a target population, data collection from groups, and scenario development – with reference to the appraisal of forestry projects. The coverage is not exhaustive, but focuses on some of the techniques that have frequent application in the area of forestry and resource management research.

1. DATA COLLECTION IN THE SOCIAL VERSUS PHYSICAL SCIENCES

Data collection methods in the social sciences differ markedly from those used in the biological and physical sciences. Typically in the latter, some form of physical measurement of real-life phenomena is possible under controlled and often repeatable conditions. For instance, the effects of fertilizer application on seedling growth can be assessed under controlled conditions in a greenhouse by measuring key parameters such as length of shoots and roots, biomass accumulation and photosynthetic activity. These types of experiments can be controlled, e.g. the amount of nutrients added can be controlled by adding know concentrations in solution, the amount of light can be controlled by using artificial lighting, and temperature can be controlled by using a constant temperature room or a heating or air-conditioning unit linked to a thermostat and timer. Replication and repeatability in science is usually achievable. This is not the case in social science research.

The key difference between collecting data in the social sciences compared with the biological and physical sciences is 'people'. Social sciences by their very nature deal with people – either directly or indirectly – which poses challenges when collecting data. For example it is not possible to 'measure' peoples attitudes in the same way that that the growth rates of seedlings can be measured. Also, people change in response to past experiences and changes in the environment. It is thus extremely difficult if not impossible to achieve the same replication and repeatability that is the case in the physical and biological sciences.

2. SURVEYS

Surveys are the most common form of data collection in the social sciences. This section will discuss how surveys can be used to collect information from and about individual people, while the following section will deal with the collection of information from groups.

Conducting a survey

The survey *process* may be thought of as both the development and administration of a questionnaire or survey instrument, and the analysis of the survey data. The major steps involved in the survey process are set out in Figure 1. The decisions made in the early stages will affect the choices at the later stages – thus the forward links in Figure 1. For instance the information needs specified at the start will affect the choice of sampling design, the way in which the questionnaire is structured and the selection of data analysis techniques. If there were only forward links in the process then the conducting of a survey could be done one step at a time, completing each step before considering the next. Implicit in this 'single direction' approach is the assumption that there are no limiting factors in later steps. This is seldom, if ever, the case. For instance there are invariably budgetary limitations on data collection or data processing resources. These

limitations restrict the alternatives available at earlier steps; these backward linkages are indicated in Figure 1 by dashed lines running upwards. Backward linkages run from the 'collect data' and 'analyse data' boxes back to the 'develop questionnaire' and 'sampling design' phases. This illustrates that major decisions concerning data collection and analysis should always be considered before selecting a sample and designing a questionnaire.

While the amount of information that could be collected about a project is almost unlimited, limited time and other resources make it necessary to prioritise the information sought. *Information needs* can be categorised into three levels of importance: *(a) absolutely essential*, constituting the reason for the survey (in the case of project appraisals, these data are required for the appraisal to be undertaken), *(b) highly valuable* for making important decisions, *(c) supporting information* which clarifies the picture but is not essential. Care must be taken to group questions logically, and to identify the most important questions to be put to respondents, and to place these appropriately within the questionnaire, e.g. at a point where rapport has been established with the respondent. More intrusive or personal questions are often placed near the end of a questionnaire.

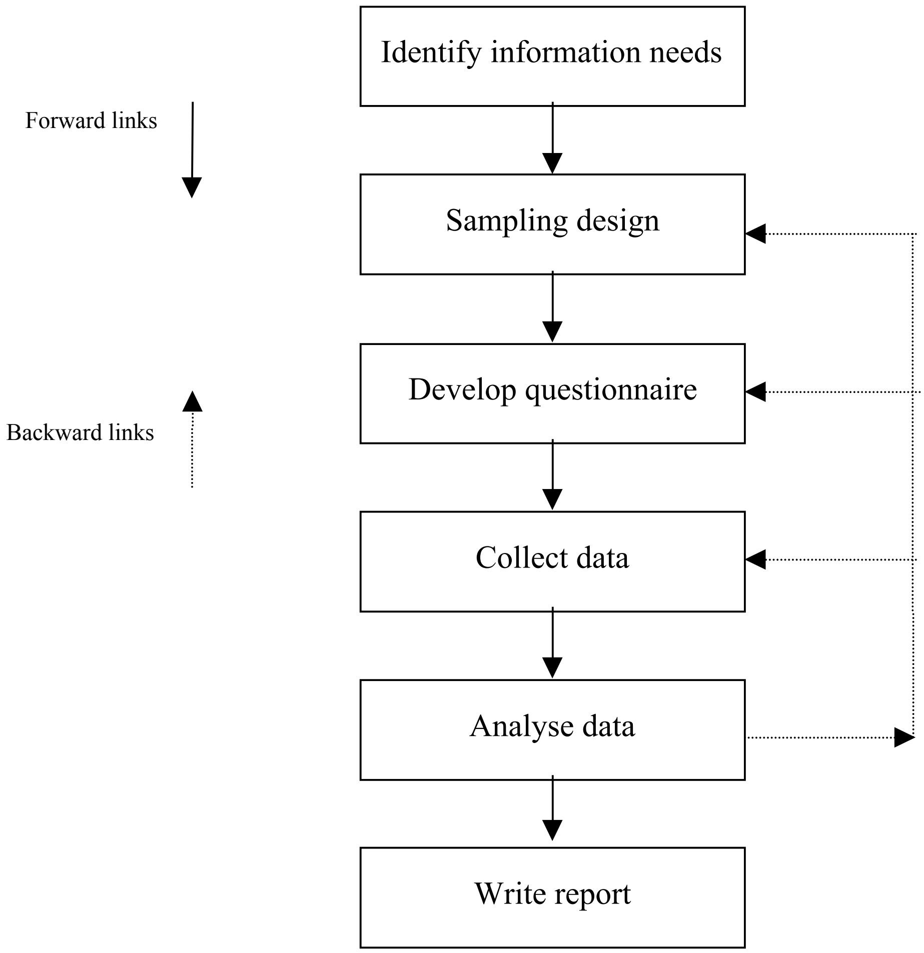


Figure 1. Major steps in the survey and data analysis process
Source: Based on Alreck And Settle (1995, p. 26).

A crucial part of any survey is deciding what group of people is to be surveyed; this group is commonly referred to as the *reference population*. When seeking estimates for input into project appraisals it is critical to ask the people who have the experience, knowledge and skills to be able to provide reliable estimates. There is no point asking people in production to make estimates about likely sales of a new

product; this question is best addressed to either marketing staff or customers. In the case of gathering judgemental estimates used in project appraisals, the population is likely to comprise a small number of experts or semi-experts. In such cases, it may be feasible to distribute questionnaires to all members of the population, i.e. to carry out a *census*. Where the population is of a size that does not permit every member to be contacted, within the budget and timeframe, a choice needs to be made regarding the basic *sampling design*. Here the typical choices are between probability or non-probability sampling. If probability sampling is chosen then further choices need to be made between sampling designs, the typical contenders being simple random sampling, stratified sampling and multistage sampling. As a rule of thumb, the less expert or focussed the population with respect to the parameters being estimated (often corresponding to a large population), the shorter should be the questionnaire. Long questionnaires distributed to groups with little or no interest in the outcomes of the survey will result in a low response rate. Long questionnaires are also more expensive to produce and analyse and are thus highly costly when large numbers are distributed.

Questionnaire development usually proceeds through a number of drafts. As part of this process the instrument may be tested on a small sub-sample in a pilot survey; this usually leads to some revision of questions. It is also critical to ensure the questionnaire is designed to elicit all the information required and that no redundant information is sought. This is best done by reference back to the previously identified information needs.

Implementation of the survey may be through personal interview, telephone interview, drop-off-and-collect or by mail ('snail mail' or email). Personal interviews are generally expensive and time consuming and are suited to the situation where the target or reference population is small and not widely dispersed. Non-response bias is generally not a problem because of the high participation rate generally associated with this method. This is especially so if respondents are being interviewed as part of their employment duties, which is sometimes the case when information is being collected as input into project appraisal. Telephone interviewing can be effective, especially where the information required is straightforward, but is not suitable for the collection of information that is complex and requires detailed thought or calculations. Postal surveys are usually undertaken when large sample numbers are required, from a modest budget. Non-response bias is an issue that needs to be considered no matter what method is used; however, it is especially a concern with postal surveys.

Data analysis is the process through which the survey responses are summarised into *descriptive* statistics such as tables and graphs, and perhaps subjected to *inferential* statistical methods such as chi-squared tests, regression analysis and analysis of variance. In a highly structured questionnaire, highly specific information is sought. Respondents will be required to provide specific estimates, such as estimates of area of trees planted and the date when planting took place or thinning and harvesting are likely to occur. Alternatively, respondents may be required to choose one option from a discrete set of options or to rank a particular statement on a predetermined scale. In such cases, descriptive statistics such as means, medians and standard errors can be easily calculated and used in project appraisal. Open-ended questions within a questionnaire allow respondents the opportunity to answer the question in their own words, and relay their particular perceptions, which can provide insights into specific issues and problems, but also pose challenges when analysing the responses.

Report generation produces a permanent record of the processed data and its interpretation. When the information is being compiled for internal use – often the case in project appraisal – the report, if prepared at all, may be rudimentary and involve simple summary tables and brief discussion of the data.

A number of texts are available which provide more detail on the survey process. An excellent survey research resource is *The Survey Research Handbook* by Alreck and Settle (1995).

3. COLLECTING DATA FROM GROUPS

Survey data collected from individuals is often aggregated to provide insights into groups beliefs and attitudes or to explain social phenomena. In the fields of economics and accounting these data is also often used to develop economic and financial models. However, in sometimes data are collected directly from groups. This section discusses a number of techniques that can be used to collect data from groups as opposed to individuals.

Evidence suggests that forecasts and other estimates produced by groups have greater accuracy than those derived from individuals. Groups also provide more information, although the marginal increase in information content decreases as group size increases. The use of groups also provides an opportunity to gain more information about the range of possible outcome values hence giving an insight into the risk associated with the estimates. From a behavioural perspective, it is also likely that a group responsible for a implementing a project will have greater commitment to it if they are involved in providing estimates of variables used in the financial analysis leading up to a decision to proceed.

When group members are allowed to interact, this may be in a structured or unstructured manner. Group processes, particularly those that are unstructured (also referred to as ‘interacting groups’), have a number of potential shortcomings (Janis and Mann 1977; Lock 1987):

a) *Group think*. In meetings, one idea is often pursued for a considerable period of time and thinking consequently becomes narrow or confined. This often reflects a common information base of group members and a desire for and encouragement of conformity.

b) *Inhibition of contributors*. Within groups there are often power differentials and members of a group may be unwilling to contradict a superior, or even express an opinion. Also, dominant personalities may reduce the willingness of others to contribute.

c) *Premature closure*. There often is a tendency for the group to adopt the first satisfactory option or estimate without fully exploring other options or possibilities.

A number of structured group techniques have been developed which aim to minimise these social and psychological difficulties, some of which are now examined.

Jury of executive opinion

This technique is one of the simplest and most widely used forecasting approaches. In its most basic form it involves simply executives meeting and deciding on the best estimate for the item being forecast. As a precursor to the meeting, it is common to provide background information to executives.

A major drawback of this approach is that it places those making the forecasts in direct contact with each other thus allowing ad hoc and uncontrolled interaction. The potential exists for problems of group interactions (e.g. dominance of individuals, group think) to arise. In particular, the weight attached to the opinion of a particular individual is likely to be determined by the rank (and personality) of that individual in the organisation. The views of executives with the best information or in the best position to make an accurate forecast are not necessarily given sufficient weight.

One variation of the this approach involves the jury (group) periodically submitting estimates in writing, which are then reviewed by the president or some other senior member. The person reviewing the individual forecasts will sometimes make a final assessment based on the opinions expressed. In doing so, this person will also often call on past experience to take into account which executives are biased in which direction and then weight each individual’s estimates accordingly. Alternatively, the individual estimates can be averaged to derive an which is considered representative of the group. This later approach could almost be considered an informal variant of the Delphi method discussed below

The Delphi method

The Delphi method was originally developed by the Rand Corporation in the 1950s to obtain consensus among experts. Since this time it has been refined further and applied to gain information in a wide range of fields. These fields are as diverse as regional economic development, health care policy, sociology, environmental risk, prediction of fruit prices, tourism and recreation, forestry and advanced manufacturing techniques. The Delphi technique may be particularly useful in situations where objective data are scarce.

The Delphi method is designed to elicit estimates from experts within a group or panel without allowing interaction between individuals on the panel, thus avoiding problems with dominant members. Experts do however have the ability to revise their estimates on the basis of group views. Such an option is not available using the traditional survey method. This technique proceeds through a series of data collection rounds.

In a classic Delphi survey, the first round is unstructured, allowing panellists to identify freely and elaborate on the issues that they consider important. These are then consolidated into a single set by the monitors, who then produce a structured questionnaire designed to elicit the views, opinions and judgements of the panellists in a quantitative form. The consolidated list of scenarios is presented to the panellists in round two, at which time they place estimates on key variable such as the time an event will occur. These responses are then summarised and the summary information is presented to the panellists, who are invited to reassess their original opinions in light of anonymous individual responses. In addition, if panellists assessments fall outside the upper or lower quartiles, they may be asked to provide justifications as to why they consider their estimates are more accurate than the median values.

Further rounds of collection of estimates, compiling summary information and inviting revisions continues until there is no further convergence of expert opinion. Experience reveals this usually occurs after two rounds, or at the most four rounds (Janssen 1978).

There are a number of variants on the classical Delphi method. When the issues are well defined, a clearly defined scenario can be developed by the monitoring team. In such circumstances, it is common to replace the unstructured first round with a highly structured set of questions through which specific estimates of parameters are obtained. A statistical summary of all responses is then provided to the panel for the second round, rather than in the third. In such cases, it is common for the Delphi method to include only one or two iterations.

The classic Delphi method is conducted through a combination of a polling procedure and a conference. Communication between conference panellists is however restricted and undertaken through the monitoring team. Even though panellists are at the same physical location, there is no face-to-face contact. A variant is the 'paper' Delphi (sometimes also known as a 'paper and pencil Delphi poll') that is conducted entirely by mail. Another variant is the 'real time' Delphi whereby feedback is provided by computer and final results are usually available at the end of the session.

The quality of forecasts and other estimates provided by Delphi method very much depends on how the technique is applied. The following list of suggestions of how to best apply the Delphi method are primarily those of Parente *et al.* (1984) with a few additions from other sources:

1. The criteria for the selection of panellists (e.g. education, experience) should be carefully determined and clearly communicated.
2. A minimum of 10 panellists after dropout are recommended although it is sometimes suggested that five is sufficient
3. Commitment to serve on the panel should be secured before the first round of estimates is requested. This will improve motivation and ensure a balanced sample if dropout is likely. Time should be taken to explain the Delphi technique and the information provided.

4. A range of estimation problems may be presented, although these should be less than 25 in number. Where appropriate, the main estimates should be broken down into sub-problems. Alternatively, different outcomes might be presented and their likelihood requested. Either way, the estimates will be useless unless the right problems are presented, hence the effort needs to be put into framing the problem. Some prior testing may be appropriate, especially if the Delphi survey is being undertaken through the post.
5. Problem statements should not be longer than 20 words and should use quantitative data (e.g. 50% increase) rather than fuzzy linguistics (e.g. 'considerable increase').
6. The 'rules' for good questionnaire design should be applied to the presentation of problems. These include avoiding compound sentences.
7. If the purpose of the Delphi process is to generate estimation problems then it is suggested that examples of attractive and undesirable scenarios be presented.
8. There is little difference in the manner in which the Delphi approach is designed in the sense that the same steps are involved in the process, regardless of whether it is administered by mail, a networked computer or face-to-face meeting. Factors such as cost, the need for timely information or the availability of experts to attend a face-to-face meeting may determine the appropriate method.
9. The principle of anonymity should be ensured. The organiser's opinions on the estimates should not be communicated to the panellists.
10. The amount and form of the feedback will need to be carefully managed. The number of rounds will depend on the panellists and the manner in which the Delphi survey is conducted (i.e. at what stage a highly structured questionnaire is distributed). The general advice is that more rather than fewer rounds, as well as descriptive feedback, are preferable. Medians should be provided.
11. Extreme responses should be screened for the panellist's expertise. If the expert has relatively low expertise, then the response might be discounted.
12. If the Delphi survey is directed to research applications, a detailed report of the process should be published to allow replication by other researchers at a later time. The range of responses should be published to demonstrate consensus or panellists' reasoning.

The nominal group technique

The nominal group technique (NGT) uses the basic Delphi structure but in face-to-face meetings which allow discussion among participants. A meeting with NGT starts without any interaction, with individuals initially writing down ideas or estimates related to the problem or scenario. Each individual then presents their ideas or estimates, with no discussion until all participants have spoken. Then each idea or estimate is discussed. The process is then repeated. For this reason, NGT is sometimes known as the 'estimate-talk-estimate' procedure. In practical terms, like Delphi, the framing of the questions or the scenario is crucial for the success of the process. Also, ideally, the leader or moderator of the discussion should come from outside the group.

Other group techniques

A number of other group techniques are available. The *Devil's Advocate* and *Dialectical Inquiry* involve individuals or small groups taking a 'devil's advocate' role or using the dialectic approach (presenting multiple views) to explore alternative different options. Both methods are considered to be ways of overcoming the problem of 'group think'.

Lock (1987) outlined a further approach to group judgmental forecasting which draws upon elements of the nominal group technique and *Inquiry Systems*. Inquiry systems according to Lock are simply philosophical systems that underlie different approaches to analyzing or investigating particular phenomena. This approach consists of seven phases:

1. Problem or task definition
2. Pre-collection of estimates of the variable of interest and the reasoning behind the estimate

3. Sharing of the estimates and clarification of the underlying reasoning behind them
4. Discussion of underlying reasoning
5. Encouragement of multiple advocacy (dialectic inquiry)
6. Individual revision of estimates
7. Synthesis of estimates.

This approach recognises the benefits of communication between groups.

4. THE DELPHI TECHNIQUE APPLIED TO APPRAISING FORESTRY PROJECTS

The Delphi technique is now illustrated as a means of collecting information to undertake a financial analysis of forestry projects, based on two 'real-life' Delphi surveys undertaken in northern Australia.

A simple model of appraising forestry investment

A simple model for appraising investment in forestry projects is illustrated in Figure 2. This diagram illustrates the key parameters which need to be estimated for evaluation of forestry projects, viz. harvest volumes and stumpage prices for the various types of timber harvested, and input costs. It is also critical to have estimates of the timing forestry operations and costs and returns throughout the plantation life or 'rotation length'. The estimates made at the time of planting become *forecasts* for deriving cash flows for the various years throughout the plantation. This information can then be entered onto a spreadsheet in which annual net cash flows and financial performance criteria are derived. Performance estimates are typically made on a one-hectare basis, and then aggregated up for plantation size.

For traditional exotic conifer plantations such as radiata or caribbean pine, it is relatively easy to obtain estimates of the various parameters of the model. For example, costs of establishment, continuing maintenance and non-commercial thinning are easily obtained, for example from contractors. Yield estimates, along with final harvest price, are the two key parameters in determining final harvest revenue. For pine plantations there are well developed stand growth models for various site indices based on many years of past growth data that can provide accurate projections of likely yield. The following two examples, on the other hand, apply to situations where non-traditional species are grown, hence little stand growth data are available.

Example 1: Appraising Forestry Projects Involving New Planting Systems

In recent years there has been a move away from traditional silviculture systems involving monocultures of a small number of mostly softwood species. In Australia for example, plantations of native hardwoods including many rainforest species have been established. In the case of native timber species for which little is known about the silviculture, it is extremely difficult to obtain estimates of growth rates that can be accepted with a high degree of confidence.

The Delphi technique is a convenient method of obtaining estimates of expected growth and harvest age of native species for which there exist no growth models based on past performance or physiological characteristics. This was in fact the case recently in tropical Australia, where it was necessary to obtain estimates of growth rates and harvest ages for 31 species (Herbohn *et al*. 1999). In this case, the Delphi method proved to be an effective method to collect plantation productivity data necessary for the financial appraisal.

This project used the Delphi method to provide estimates of (a) mean annual increment or MAI (m^3/ha/yr) and (b) time to harvest (years) of 31 species. Harvest age and MAI are the key biological parameters needed to estimate yield and harvest scheduling for use in financial models. In this case the species for which information was sought had either been widely planted in the area or had been included in a previous Delphi survey.

Opinions were sought from 13 individuals with extensive experience in growing of Australian tropical and sub-tropical rainforest species for either timber production or reforestation. Individuals generally had either extensive field experience or had undertaken research involving native rainforest and tropical eucalypt species.

Panellists were provided with a table listing the 31 selected species and asked to provide estimates of their 'best guess' of optimal rotation period (years) for each species along with estimates of 'shortest time to harvest' and 'longest time to harvest'. Estimates were also requested for the 'best guess' for expected yield (m^3/ha/yr) based on the 'best guess' rotation period along with estimates of 'highest yield expected' and 'lowest expected yield'. In this section, participants were asked to assume that the trees would be planted on relatively fertile basaltic soils, that average annual rainfall would be between 1500-2000 mm, that initial planting density would be around 660 stems per hectare (sph) and suitable thinning regimes would be applied.

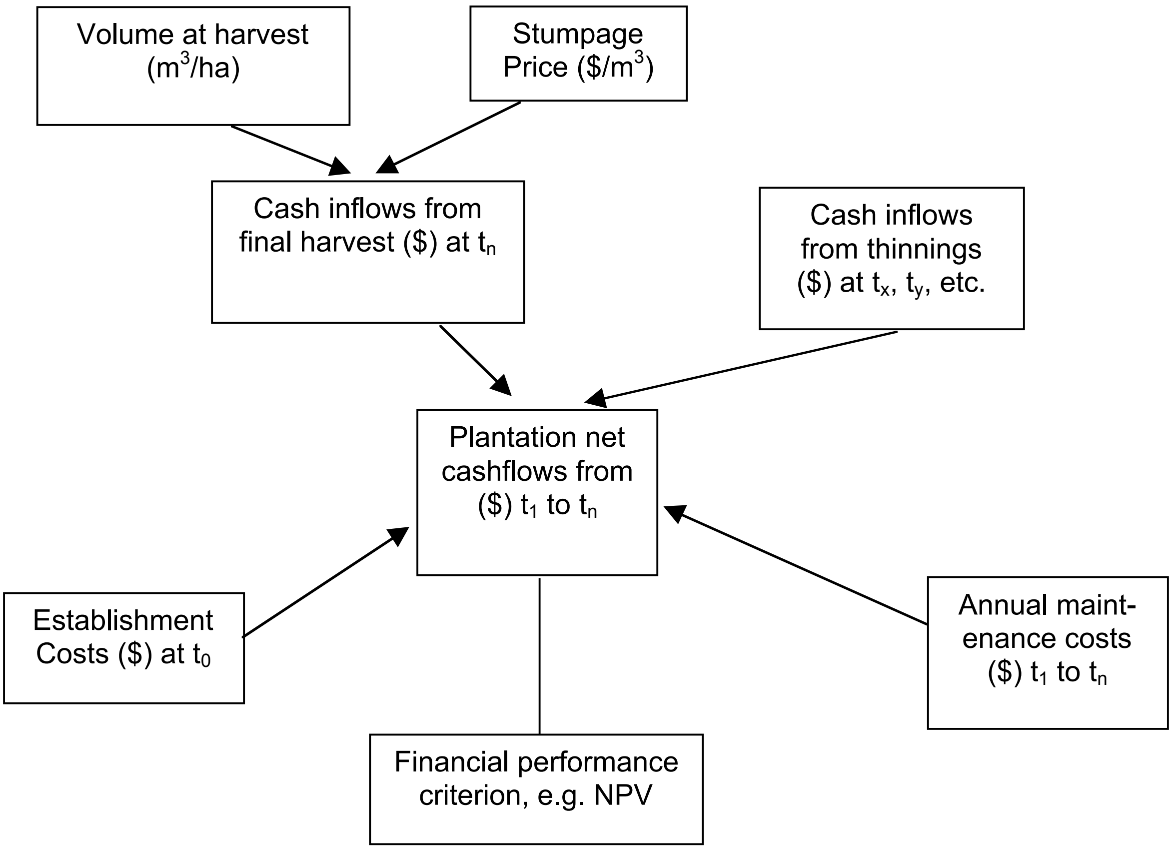


Figure 2. A simple model for appraising investment in forestry projects

Questionnaires were distributed to participants followed by a visit by one of the research team. Responses for the estimates of growth rates and harvest ages of the 31 selected species were then collated and averages calculated. A summary table including the group averages was prepared and distributed to participants along with their original estimates. In this second round of the Delphi survey participants were given the opportunity to review their original estimates of growth rates and harvest ages in light of the group averages and to provide any appropriate revisions or comments. Few revisions were received in this second round and the Delphi process was then terminated. An extract of the survey form used in the first round of the Delphi survey is provided in Figure 3. In the second round, a similar table was compiled with the averages of estimates provided from all panellists, along with the estimates from that particular panellist.

Outcomes of the Delphi survey

The outcome of this Delphi survey was a table of harvest ages and yields, where for each variable the group mean and highest and lowest estimates were recorded.

Species	Common name	Optimal rotation period (years)			Yield based on 'best guess' rotation period (m^3/ha/yr)		
		Best guess harvest age (years)	Shortest time to harvest (years)	Longest time to harvest (years)	Best guess yield	Highest expected yield	Lowest expected yield
Acacia mangium	Mangium						
Acacia melanoxylon	Black-wood						
Agathis robusta	Kauri Pine						
Araucaria cunninghamii	Hoop pine						
Beilschmieda bancroftii	Yellow walnut						
Blepharocarya involucrigera	Rose Butternut						
Cardwellia sublimis	Northern silky oak						
Castano-spermum australe	Black Bean						
Cedrela odorata	West Indian cedar						
Ceratopetalum apetalum	Coach-wood						

Figure 3. Extract of survey form used in Stage 1 of Delphi survey in Example 1

Example 2: Collecting Data for Forestry Projects Involving New Planting Systems

In north Queensland it has become a common practice to plant *Flindersia brayleana* with *Eucalyptus cloeziana* and many potential investors are interested in the possible financial returns from such as plantation. There is however a lack of growth models for this mixture. *Flindersia brayleana* exhibits marked crown shyness (i.e. stops growing when the leaves in its crown touch the leaves in a crown of another tree). There is also a neat relationship between crown diameter and the diameter of the stem. These two characteristics make it very easy to develop a well-structured plantation scenario involving the two species. A Delphi survey was carried out to obtain information that for development of a financial model for the two species mixture (see Herbohn and Harrison 2001).

A planting and harvesting scenario was developed for a 50:50 mixture of *Flindersia brayleyana* and *Eucalyptus cloeziana* (Table 1). Thinning and harvesting regimes are timed to occur just as *F. brayleyana* crowns touch, at which time lock-up of growth would be expected to occur.

Table 1. Planting and harvesting scenario for a *Flindersia brayleyana* and *Eucalyptus cloeziana* mixture

Stage and activity	Density after treatment (stems/ha)	Estimates requested from participants in Delphi survey of reforestation experts
1. Plant alternating rows of maple and eucalyptus at 3m x 5m spacing	660 (330 *F. brayleyana*, 330 *E. cloezianna*)	
2. Thin to waste every second tree., when *F. brayleyana* reaches 18 cm dbh.	340 (170 *F. brayleyana*, 170 *E. cloeziana*)	Age at which *F. brayleyana* expected to reach 18 cm dbh
3. Thin every second eucalypt to waste or for strainer posts. Thin when *F. brayleyana* reaches a dbh of 32 cm.	255 (170 *F. brayleyana*, 85 *E. cloeziana*)	Age when *F. brayleyana* expected to reach 32 cm dbh. Bole dbh, small-end diameter, bole length of *E. cloeziana* at this age
4. Remove every second *F. brayleyana* for strainer posts or small diameter logs	170 (85 *F. brayleyana*, 85 *E. cloeziana*)	
5. Remove remaining eucalyptus for poles.	85 *F. brayleyana*	Ages when *E. cloeziana* expected to reach 5 specified pole dimensions
6. Harvest every second *F. brayleyana* when crowns touch (50cm dbh)	43 *F. brayleyana*	Age at which *F. brayleyana* expected to reach 50 cm dbh Small-end diameter and bole length at this time
7. Harvest remaining *F. brayleyana* (81cm dbh)	Nil	Age at which *F. brayleyana* expected to reach 81 cm dbh. Small-end diameter and bole length at this time

Personal interviews were conducted with five north Queensland forestry experts chosen for their familiarity with the species being modelled. At the commencement of interviews the plantation system was outlined and the requirements for information stated. Panellists were provided with a table, similar to Table 2 but with columns 2 and 3 blank, in which to record their estimates after collation of estimates, outliers were identified and clarification was sought from participants. The final estimates of the parameters that were used as inputs to the financial analysis are provided in Table 2.

In this instance, the Delphi method proved to be a timely and cost-effective means through which to collect the information and forecasts necessary to construct the financial model.

While it is difficult to judge the accuracy and quality of the forecast information obtained in Examples 1 and 2, the Delphi surveys provided information without which the construction of a financial model would have been impossible. The stimulus for under choosing to undertake a Delphi survey was the fact that forecasts of tree growth and harvest age from models based on quantitative growth data will not be available until recent plantings using this species mix reach harvest age, and efforts to develop quantitative models based on physiological and environmental parameters had failed to produce suitable models.

5. SCENARIO PROJECTION

What will the transport infrastructure requirements for forestry in Leyte in 25 years? What will electricity demand be within the province in 10 years and how much electricity will be generated from renewable resources such as the hydropower and fuelwood? What demands will the increasing population place on timber for housing, furniture and charcoal in 20 years time? What is the long-term market prospects of a new product such as plywood requiring large-scale research and development expenditure? All of these questions are important for capital budgeting purposes.

Table 2. Estimates of model parameters for a *Flindersia brayleyana* and *Eucalyptus cloeziana* mixed plantation

Stage and parameter	Parameter estimate from Delphi survey	
	Average	Range
2. Age when *F. brayleyana* expected to reach 18 cm dbh	8.6 years	7 - 10 years
3. Age when *F. brayleyana* expected to reach 32 cm dbh	17.6 years	15 - 20 years
At this age, the following are expected for *E. cloeziana*		
- bole dbh	41.4 cm	35 - 50 cm
- small end diameter	26.2 cm	15 - 35 cm
- bole length	16.4 m	12 - 20 m
5. Age when *E. cloeziana* expected to reach specified dimensions for		
- Pole 1	17.4	10 - 25 years
- Pole 2	17.6	11 - 25 years
- Pole 3	21.0	12 - 30 years
- Pole 4	24.2	14 - 35 years
- Pole 5	25.6	15 - 35 years
6. Age when *F. brayleyana* expected to reach 50 cm dbh	34 years	25 - 40 years
Expected small-end dia. at this age	33 cm	30 - 40 cm
Expected bole length at this age	14.8 m	6 - 20 m
7. Age when *F. brayleyana* expected to reach 81 cm dbh	60 years	50 - 65 years
Expected small-end dia. at this age	57 cm	45 - 60 cm
Expected bole length at this age	16.6 m	10 - 20 m

For instance, infrastructure for transport, electricity and plywood require long lead times to develop. In deciding on whether to proceed with infrastructure and new product investments, it is first necessary to estimate demand for those services and products in the future. Scenario projection provides a convenient technique to do this.

Scenarios have been described as ‘descriptions of alternative hypothetical futures’. Scenarios can be used to describe what potential futures we might expect, depending on whether major events will or won’t come to pass. Generating scenarios can be a useful tool in capital budgeting.

While the term ‘scenario’ is used with various meanings, there are several features which characterise scenarios, viz.

- A scenario is hypothetical – it describes some possible future.
- A scenario is selective – it represents on possible state of some complex future.
- A scenario is bound – it consists of a limited number of states, events, actions and consequences.
- A scenario is connected – its elements are related, that is each element is conditional on or caused by other elements.
- A scenario is assessable – it can be judged with respect to its probability or desirability.

Ducot and Lubben (1980) suggest that scenarios can be distinguished in a number of different ways:

Exploratory vs anticipatory. Exploratory scenarios start with some known or assumed states or events and explore what might result, i.e. they look at what might result. Anticipatory scenarios start with some assumed final state of affairs and look at what preconditions (events or actions) could produce this state of affairs, i.e. they are backward looking.

Descriptive vs normative. Descriptive scenarios present a possible future irrespective of their desirability or otherwise. Normative scenarios take values and goals explicitly into account.
Trend vs peripheral. A trend scenario extrapolates the normal, surprise-free course of events that might be expected if nothing out of the normal was to happen or no particular course of action taken. A peripheral scenario depicts radical, trend-breaking or unprobable developments.

Based on practical experience, Schoemaker (1991) provided some guidelines for dealing with scenarios:

1. Develop an understanding of any issues thought important, especially in terms of their history, to get a feel for the degrees of uncertainty.
2. Identify the major stakeholders who would be interested in these issues. Both those with power and those influenced should be noted to clarify their roles, interests and exact power.
3. Make a list of current trends that might affect the issues. Explain how these trends might impact on the issues.
4. Identify key uncertainties and explain how they matter.
5. Construct two 'forced' scenarios by placing all the positive outcomes in one scenario and all the negative outcomes in another.
6. Assess the plausibility of these 'forced' scenarios. Eliminate impossible combinations. These revised 'forced' scenarios might be called 'learning-only' scenarios.
7. Reassess the stakeholders in the learning scenarios. Identify and study topics for further consideration
8. Develop outline plans based on what has been learned so far. Communicate desired scenarios to responsible managers

Example 3: Using scenario forecasting to estimate demand

As an example of scenario forecasting, suppose the planning department of a company in the electricity generation and wholesale sector wishes to forecast electricity demand in 10 years time, to assist in deciding whether to construct a new hydropower station. The steps in scenario development might proceed as follows:

1. Understanding of issues and history. Develop an understanding of any issues thought important, especially in terms of their history, to get a feel for the degrees of uncertainty. These demands have been growing, but there have been some unrealistically high demand forecasts in the past which have at led to excessive capital outlays, and criticism by environmentalists.
2. Identifying major stakeholders groups. The major stakeholders are other electricity generators, electricity consumers (industrial, commercial and domestic) and consumer advocates (sensitive to price increases), environmental groups (concerned about increasing greenhouse gas emissions).
3. Identifying current trends. Electricity demand is increasing steadily due to increased population, rural electrification and increasing ownership of electrical appliances. A campaign has been undertaken by government to encourage adoption of electricity rather than charcoal for cooking, for air quality reasons. The net impact is likely to be rapidly increased electricity demand until adoption of electrical appliances reaches saturation, then a slow increase as energy efficiency is pursued.
4. Identify key uncertainties. On the supply side, other generators may install new plant. On the demand side, it is possible that new industrial facilities will be constructed, and major rural electrification projects will proceed, but the boom in construction of office blocks could come to an end. These events could lead to major changes in overall generation capacity, and in industrial and commercial demand.
5. Identify two 'forced' scenarios. The most favourable scenario for the generator would be where there is little increase in generation capacity, but a large increase in demand. The most negative outcome would be where other generators build new plant, but expected increase in demand does not eventuate. Demand and supply quantities in megawatt hours would need to be placed on these scenarios,

and price impacts inferred.

6. Experience indicates that both extremes are possibilities, although installation of competing new plant, moderate increase in industrial demand and little increase in commercial demand over the forecast period appears more probable.
7. Reassessment of the stakeholders in the learning scenarios. Identify and study topics for further consideration. Here, further investigation might be undertaken into how research developments in the field of renewable energy might affect household consumption and the capacity of export of energy from co-generation of small electricity producers such as cane mills might change with developing technology.
8. Develop outline plans, and communicate scenarios to management. The 'most probable' forecast and some comments on alternative outcomes would form the core information.

Further information on the use of scenarios in forecasting can be found in Jungermann and Thüring (1987). Vlek and Otten (1987) have provided an excellent coverage of the use of scenarios in the energy industry.

6. CONCLUDING COMMENTS – WHICH TECHNIQUE IS BEST?

In this section a number of data collection techniques have been outlined which may be useful in socio-economic studies for estimating future values of key variable as input into financial appraisals. No matter what technique is selected, it is important to recognise its limitations because these will affect how the technique is applied and the quality of the estimates generated. Furthermore, with any technique that involves the collection of data it is essential to proceed in an orderly and well thought-out manner. Sometimes there is a tendency to collect information first and then worry about how it is to be used. The starting point however should be first to clearly identify what information is needed, decide on the most appropriate technique to collect the data (in context of the resources, time and other limitations) and *only then* commence data collection.

REFERENCES

Alreck, P.L. and Settle, R.B (1995), *The Survey Research Handbook*: *Guidelines and Strategies for Conducting a Survey*, Irwin Professional Publishing, Chicago.

Ducot, C. and Lubben, G.J. (1980), A typology of scenarios, *Futures*, 12: 51-57.

Herbohn, J.L., Harrison, S.R. and Emtage, N. (1999), 'Potential performance of rainforest and eucalypt cabinet timber species in plantations in North Queensland', *Australian Forestry*, 62: 79-87.

Herbohn, J.L. and Harrison, S.R. (2001), 'Financial analysis of a two-species farm forestry mixed stand' in S.R. Harrison and J.L. Herbohn, eds., *Sustainable Farm Forestry in the Tropics: Social and Economic Analysis and Policy*. Edward Elgar, Cheltenham, pp. 39-46.

Janis, I.L. and Mann, L. (1977), *Decision Making*, The Free Press, New York.

Jungermann, H. and Thúring, M. (1987), 'The use of mental models for generating scenarios', in G. Wright and P. Ayton, eds., *Judgmental Forecasting*, Wiley, Chichester, pp. 245-266.

Lock, A, (1987), 'Integrating group judgments in subjective forecasts', in G. Wright and P. Ayton, eds., *Judgmental Forecasting*, Wiley, Chichester, pp. 109-127.

Parente, F.J., Anderson, J.K., Myers, P. and O'Brien, T. (1984), 'An examination of factors contributing to Delphi accuracy', *Journal of Forecasting*, 3: 173-182.

Schoemaker, P.J.H. (1991), 'When and how to use scenario planning: A heuristic approach with illustration' *Journal of Forecasting*, 10: 549-564.

Vlek, C. and Otten, W. (1987), 'Judgmental handling of energy scenarios', in G. Wright and P. Ayton, eds., *Judgmental Forecasting*, Wiley, Chichester, pp. 267-289.

4. Basic Computer Skills and Statistical Methods for Analysis of Survey Data

Nick Emtage, John Herbohn and Steve Harrison

This module provides an introduction to the use of spreadsheet software packages, to enter, organise and report data from attitudinal and behavioural surveys. In particular, application of the Excel spreadsheet for these purposes is illustrated. The data used for illustration purposes drawn from a survey of landholders' attitudes to forest plantation establishment in north Queensland, Australia. To ensure comprehensive and accurate reporting of the responses to a survey, it is necessary to carry out a carefully designed series of procedures. The basic stages are data entry, reduction and transformation, analysis and reporting. Figure 1 illustrates the methodology adopted to analyse a survey of landholders attitudes to tree planting and management.

The specific procedures which are discussed in this module include:

1. data entry (spreadsheet formatting, data encoding, data entry, data categorisation and transformation);
2. data summary (development of descriptive statistics such as means and measure of variance, summary tables, error checking);
3. data categorisation and transformation (re-categorising nominal data, transforming data to fit normal distributions);
4. data analysis (Chi-square analysis, one-way ANOVA's); and
5. data reporting (presentation of results of analyses).

1. DESIGNING OF THE SURVEY INSTRUMENT TO MAXIMISE DATA UTILITY

The steps taken following data entry depend on the project duration and budget and on the researchers' aims, experience, training and skill. There is no 'right' way to analyse data from surveys, although the formats or types of data collected in the survey and the way they are recorded does determine the types of statistical analysis that can be undertaken. Compiling descriptive statistics of the variables in the data set is the first step and many survey reports fail to go beyond this and analyse the relationships between the variables. The depth of data analysis required will determine the further actions which must be undertaken. If analysis of the relationships between variables is planned, some form of data reduction and transformation is typically needed. Different data types are reduced and transformed in different ways, as illustrated in Figure 1.

It is critically important that the survey instrument (i.e. questionnaire) be designed to provide data in a form that is appropriate for entry into the computer and analysis. Important decisions about the analysis and intended uses of the survey data need to be made prior to the design of the questionnaires. The format of the questions used affects the types of analysis that can be later undertaken. Those designing the survey instrument need to understand the limitations of different formats of data. Data types include nominal data, ordinal data, scales and interval data. If data are collected in nominal (i.e. categorical) form, this limits the way that analyses can be undertaken. Data collected in an ordinal form (i.e. ranked observations) allow the use of more powerful statistical analysis techniques and the data can be collapsed into categories of the analyst's choosing should this be required.

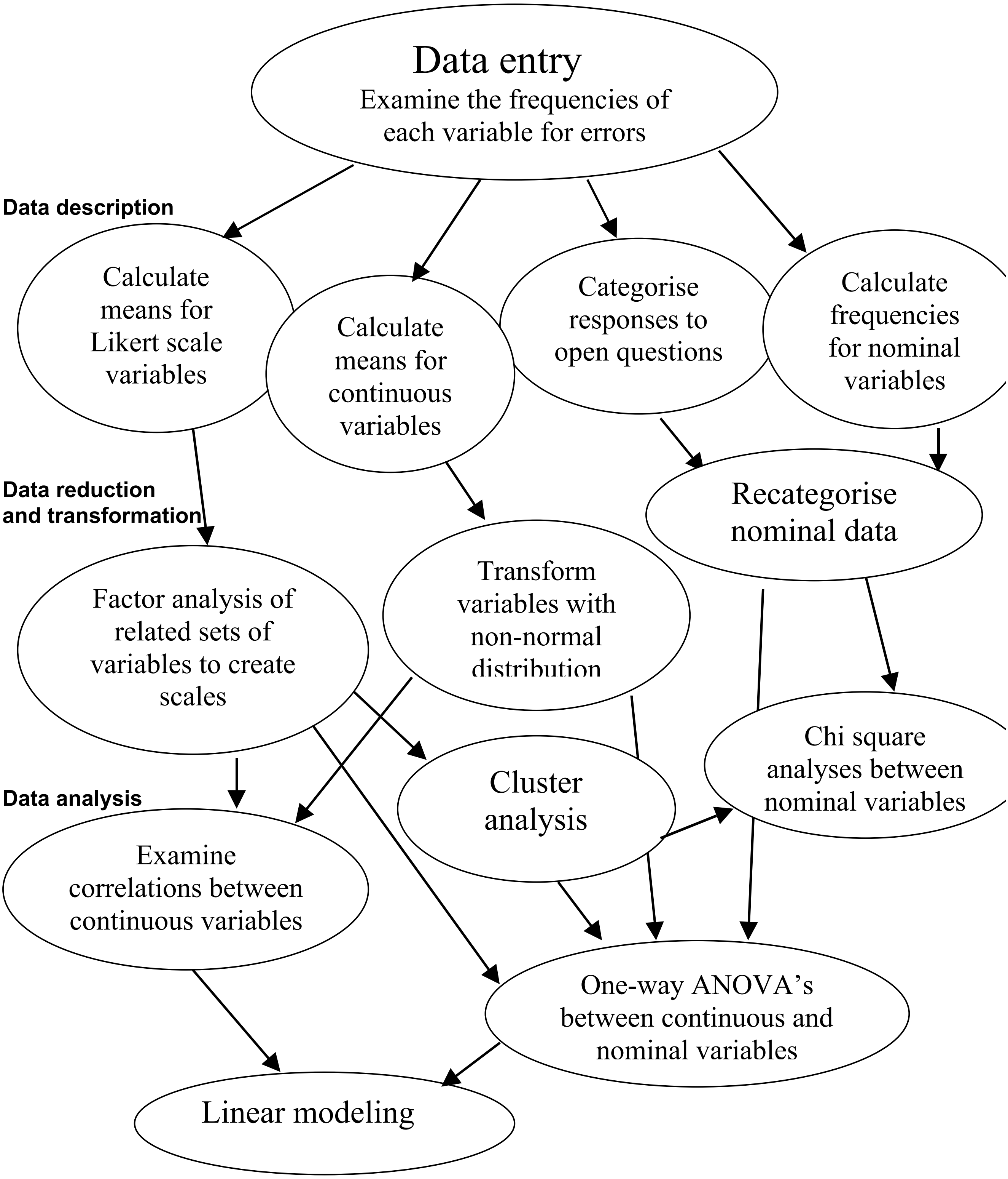


Figure 1. Methodology for analysing the responses to a survey of landholders in the Far North Queensland region of Australia
Source: Emtage *et al.*, in prep.

The desire to collect data in formats that allow greater analytical power has to be balanced against ethical concerns and the need to maximise responses. In Australia many university ethics committees will not approve research that asks for a large amount of detail about an individual. For example, questions about respondents' age are often required to be formatted as class intervals rather than specific number of

years. Such formatting may also be more comfortable for respondents than asking them to state their exact age.

It is important to test the survey instrument to ensure that it is well designed, that questions are clear, and the range of responses can be accurately assessed. It is also important to test the data entry and analysis. This provides the researcher the opportunity to set up the data-entry spreadsheets, and to assess what statistical tests can be legitimately used given the types and formats of data being collected. It also allows the researchers to assess the numbers of responses that may be required to run various statistical tests if the number of categories used for nominal variables is known.

2. DATA ENTRY

A number of factors require consideration at the time of data entry. They include choosing which software package to use for the data analysis, setting up the data entry spreadsheet, and setting up data categorisation and transformations.

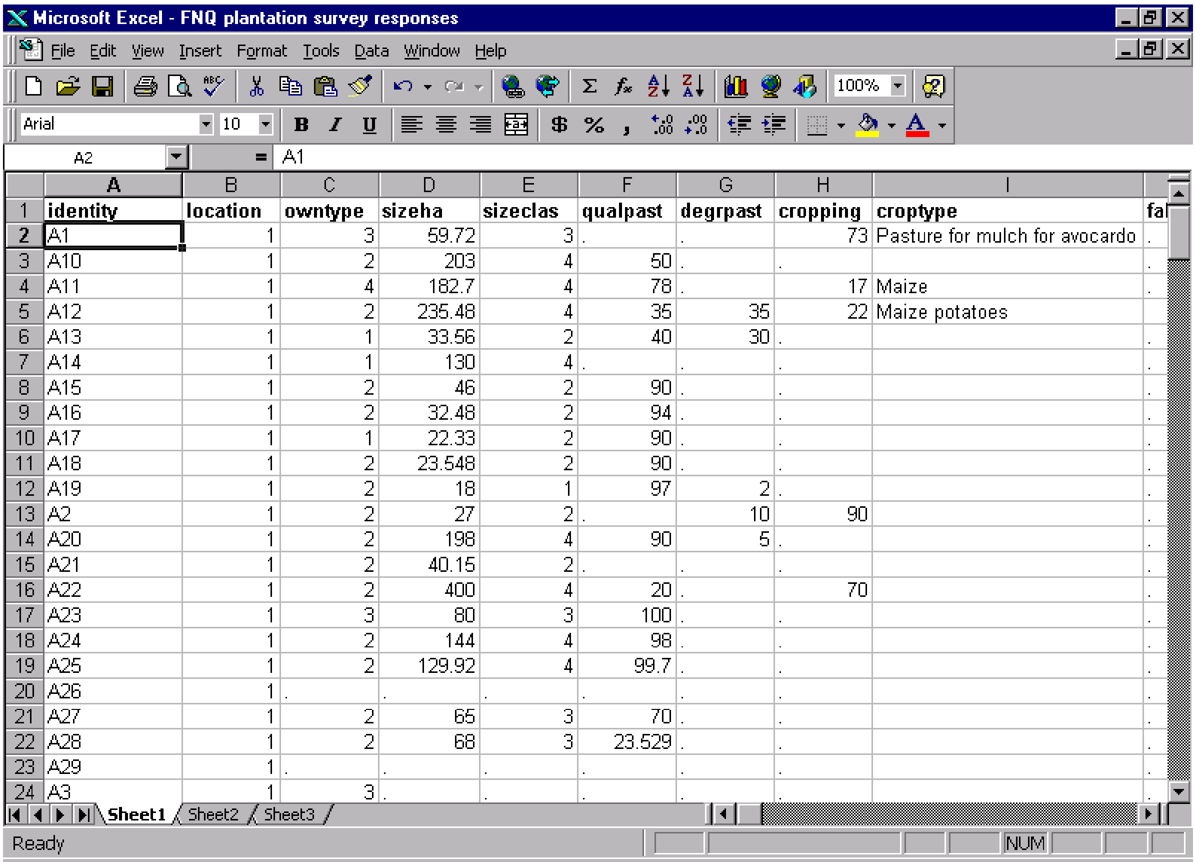
Choosing the software package to use

When entering data from survey responses the researcher needs to consider the types of analysis they wish to undertake and the availability of software packages. If the researcher plans to undertake advanced statistical analysis using multivariate analysis of variance, multiple regressions, factor analysis, cluster analysis or discriminant analyses, then specialist statistical packages such as SPSS (Statistical Package for the Social Sciences) or SAS (Statistical Analysis System) will probably be required. Unless the researcher understands advanced mathematics and statistical theory and can write their own formulae, entering data directly into these specialist programs can save time. If the researcher does not plan to undertake sophisticated statistical analyses or does not have access to such specialist packages then basic analyses can be undertaken using spreadsheet packages such as Microsoft's' Excel or Lotus 1-2-3. The SPSS package allows users to import data directly from Excel if the user later decides to undertake more advanced statistical analyses or the package becomes available. It is recommended that researchers use the specialist packages for all analyses where possible, even the most basic, because of the greater ease of analysis and reporting from statistical programs relative to spreadsheet programs. It should be remembered that all software packages take time to learn. Basic familiarity with large programs such as SPSS and Excel can take months while a high level of expertise may take years of experience to acquire. For the purposes of this module, data entry and analysis is illustrated with reference to Excel spreadsheets, because this package is widely available (as part of the Microsoft Office software) and most researchers have some familiarity with it.

Setting-up the data entry spreadsheet

Just as it is important to know what types of analysis will be attempted when designing the survey instrument, it is also important to keep the intended analysis in mind when entering the data into the software program. As a general principle, a master spreadsheet (and back-up copies!) should be used to enter the data where attempts are made to capture the greatest possible details in the survey responses. For ease of analysis the detail can be summarised or reduced in later copies of the spreadsheet. It is inconvenient to add detail at a later stage and the data entry has to be finished before analyses can commence so it is best to start by entering all available information.

In the north Queensland forestry survey, the survey instrument was a self-administered (i.e. postal) questionnaire. Respondents sent the questionnaires back to the research team using pre-paid and self-addressed envelopes. A master recording spreadsheet was set-up in Excel with the respondents labeled using an identifying code in the first column, and with their responses to each question recorded in subsequent columns (Figure 2).



	A	B	C	D	E	F	G	H	I	
1	identity	location	owntype	sizeha	sizeclas	qualpast	degrpast	cropping	croptype	fa
2	A1	1	3	59.72	3	.	.	73	Pasture for mulch for avocardo	.
3	A10	1	2	203	4	50	.	.		.
4	A11	1	4	182.7	4	78	.	17	Maize	.
5	A12	1	2	235.48	4	35	35	22	Maize potatoes	
6	A13	1	1	33.56	2	40	30	.		.
7	A14	1	1	130	4	.	.	.		.
8	A15	1	2	46	2	90	.	.		.
9	A16	1	2	32.48	2	94	.	.		.
10	A17	1	1	22.33	2	90	.	.		.
11	A18	1	2	23.548	2	90	.	.		.
12	A19	1	2	18	1	97	2	.		.
13	A2	1	2	27	2	.	10	90		.
14	A20	1	2	198	4	90	5	.		.
15	A21	1	2	40.15	2	.	.	.		.
16	A22	1	2	400	4	20	.	70		.
17	A23	1	3	80	3	100	.	.		.
18	A24	1	2	144	4	98	.	.		.
19	A25	1	2	129.92	4	99.7	.	.		.
20	A26	1	.	.	.	.	.	.		.
21	A27	1	2	65	3	70	.	.		.
22	A28	1	2	68	3	23.529	.	.		.
23	A29	1	.	.	.	.	.	.		.
24	A3	1	3	.	.	.	.	.		.

Figure 2. Extract from data entry spreadsheet for north Queensland forestry survey

Data categorisation and transformation

In the example presented in Figure 2, to maintain confidentiality the respondents have been labeled using a code (in column A). Coding is used not only to maintain confidentiality, as in this case, but also to speed up data entry. Note that the responses to some questions are already coded.

For example the responses to the question about the ownership type (which included 'partnerships', 'sole trader', 'business' and others) has already been coded into a numerical format rather than writing the full category title for each respondent. This is easily done when there are a limited number of categories. In column 'I' (croptype) the full text of responses has been entered because this question was framed as an 'open' question. Once all of the responses have been entered the range of responses can be assessed and a decision made about how to collapse or reduce the data. In SPSS, labels can be applied to categories which are then shown in reports of analyses to aid interpretation of the data. An important part of pilot testing a survey instrument is to identify the likely range of responses to such a question in order to determine whether to include a discrete range of responses in the questionnaire (plus an 'other' category), or frame the question in an open format.

An example of the categorisation of continuous data is provided in columns 'D' and 'E' relating to the size of the property operated. In this case the range of responses in column 'D' were examined and size classes were determined and computed as a new variable in column 'E' (i.e. less than 20ha = 1, 20-<50 ha = 2, 50-<100 ha = 3, and >100ha = 4). This is one example of transforming variables to create new variables to assist in summary and analysis of the survey responses. In other cases, transformation of responses may be necessary because of the assumption of normal distribution required for some statistical tests, including one-way ANOVA, as discussed later.

3. DATA SUMMARY

Part of the advantage of using spreadsheets to enter and organise data from surveys is the potential to calculate quickly descriptive statistics of responses to various questions. The specialist statistical software packages such as SPSS are designed for this task and are easier to use than spreadsheet programs such as Excel for this purpose although Excel is relatively simple to use. The development of descriptive statistics by writing formulae into cells is illustrated in Figure 3. Note that the different data types or formats require different summary measures. The calculation of means for categorical variables such as 'location' (column B in Figure 3) is meaningless while the 'count' of the number of responses in each category is valid. It is quicker to type a formula into a cell (e.g. cell B228 in Figure 3) then copy it across the spreadsheet than to enter formulae into each cell individually depending on the data type. Users can go through the columns and delete the irrelevant statistics if they wish to avoid confusion. Organisation of the data for analysis and reporting is necessary. This can be done through categorisation of the sheets in a spreadsheet. Data entry is made onto a 'master' spreadsheet, then copies of this are and are used to carry out data transformation and analyses. The separate sheets in the workbook can be organised to summarise data by topics, organised as summaries of the statistical tests used in analyses, or both can be used. The filing system used to manage the volumes of data generated by surveys and their analysis is up to the researcher.

Some of the summary statistics that can be developed using functions in Excel are illustrated in Figure 4 that shows the 'Paste Function' dialog box. Clicking on the '*fx*' button on the 'standard' toolbar at the top of the screen when Excel is running (as shown in Figures 2 and 3) accesses these functions. The dialog box then prompts users to enter the required parameters for a function. Once the user knows the syntax for these functions they can be typed directly into the formula bar (as shown in Figure 3). An alternative to generating summary statistics using the calculation functions is to use the 'Pivot Table Function' that is available under the 'Data' menu in Excel. This function is discussed further below.

The summary statistics serve three functions. First, they illustrate the types of respondents in terms of their land size, average age, education, land use activities and so on. Second, these averages can be compared to regional or national averages to assess if the respondents to the survey are representative of the broader community (non-response bias tests should also be used). Third, examining the summary statistics helps to identify if there are recording errors in the database. It is easy to make typographic errors that can seriously affect later statistical tests and examination of the database prior to running statistical tests is essential.

Another powerful feature of Excel that can be used to help analyse and report data is the 'macro'. Macros allow users to write their own functions in Visual Basic computer code for specific applications. Like the use of the Excel program generally, it takes time to become familiar with the use of macros and to set-up new code.

If users only need to undertake an operation such as categorising an ordinal variable several times it is probably more efficient to do these tasks manually. If a task is repetitive and needs to be carried out many times it can be more efficient to record and alter a macro to automate the task. Following data entry, macros can be used to automate virtually any of the tasks involved in transforming, analysing and reporting data from a survey. Whether developing macros is more efficient than manually carrying out these tasks depends upon the size of the database being used, the repetition involved in the tasks, and the skills of the researcher as a programmer using Visual Basic.

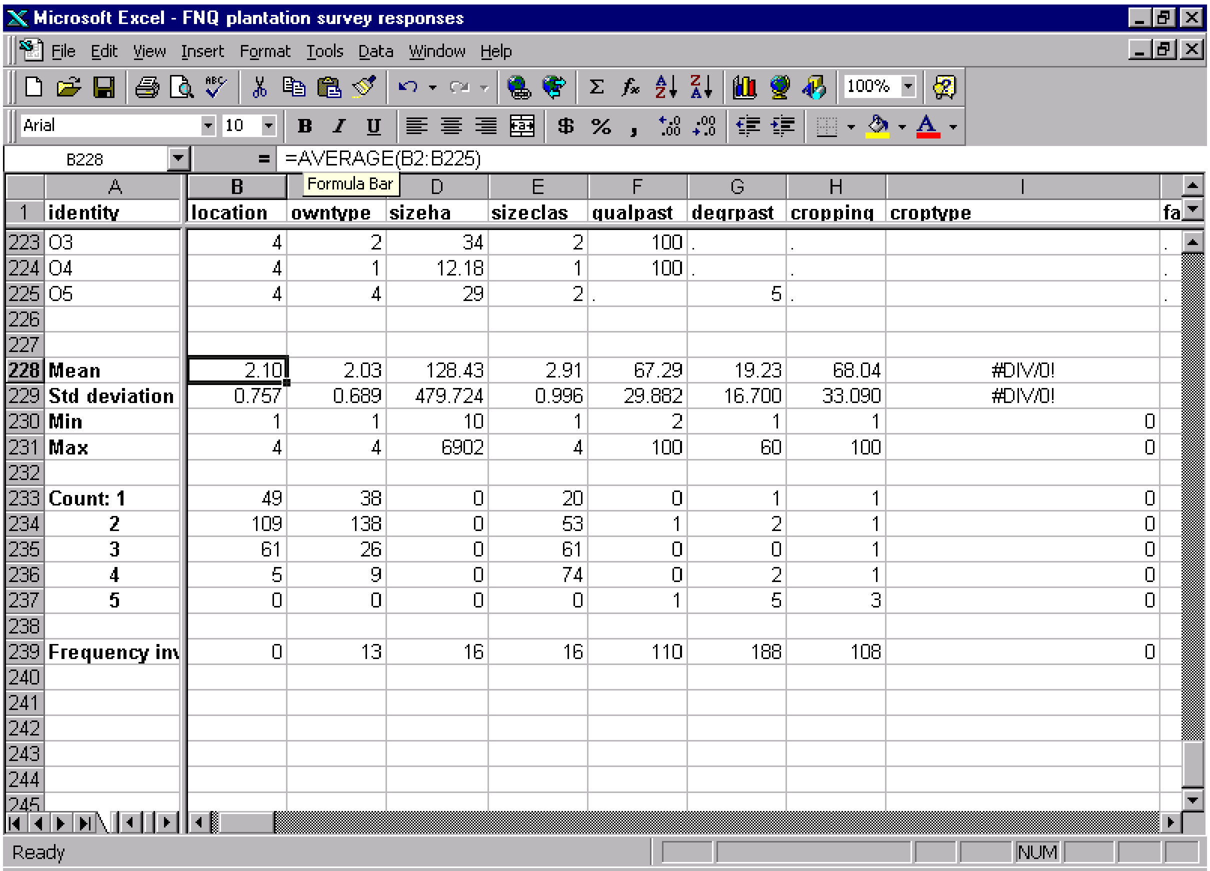


Figure 3. Descriptive statistics developed in Excel

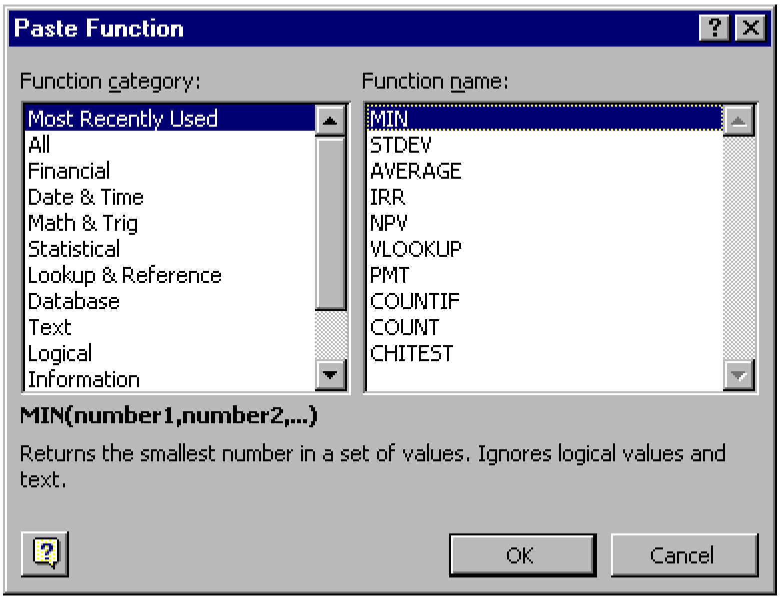


Figure 4. The 'Paste function' dialog box

Once the summary statistics have been computed they can be entered into tables to aid the interpretation of the data. The tables can be organised to contain only related variables, i.e. those related to a particular subject. Two such tables are Tables 1 and

2. It is likely that a reasonable sized survey covering several topics will require the construction of many such tables. Graphs are another way to present data, as discussed in a later section.

In some cases it is useful to present summaries of data using two categories such as land size classes by location as illustrated in Table 3. The 'Pivot Table Report' function in Excel (available under the 'Data' menu) allows users to put together quickly tables that summarise one or more than one variable.

Another Excel function that can be used to construct summary tables for numerical data is the 'descriptive statistics' function. This function is located under 'Data analysis' which is in the 'Tools' menu. The dialog box shown after following the above steps guides users through the use of the function. The table produced is like Table 4.

4. DATA CATEGORISATION AND TRANSFORMATION

Once the responses to the survey have been entered into a database and the database has been examined for errors, the next step toward data analysis involves categorising and transforming the data into formats suitable for analyses. In the case of nominal data, particularly with questions that have been framed in an open format, the researcher often has to re-categorise the initial responses before analyses are possible.

A trade-off is usually necessary between maintaining the details of the responses and being able to analyse and report them.

Table 1. Land uses as a proportion of the total landholding for all respondents (%)

Statistical measure	Quality pasture	Degraded pasture	Cropping	Fallow	Forest	Other
Average	67.29	19.23	68.04	11.85	26.89	12.72
Standard deviation	29.88	16.70	33.09	11.38	28.41	17.17
Minimum	2	1	1	1	0.3	1
Maximum	100	60	100	50	100	100

Table 2. Ratings of importance (on 5-point scale) for various reasons for planting trees by all respondents

Reason for planting trees	Average	Standard deviation	Minimum rating	Maximum rating	n
Other reasons	4.39	0.839	2	5	23
Protect land resources	3.98	1.157	1	5	172
Protect water resources	3.96	1.193	1	5	170
Provide fauna habitat	3.64	1.256	1	5	169
Personal reasons	3.44	1.301	1	5	170
Aesthetic reasons	3.35	1.327	1	5	168
Increase value of land	3.16	1.362	1	5	166
Windbreak	3.15	1.483	1	5	168
Legacy for children	3.13	1.514	1	5	166
To make money	2.66	1.472	1	5	167
Diversification of income	2.39	1.492	1	5	163
Superannuation	2.16	1.483	1	5	164
Fenceposts	1.52	0.975	1	5	161

Table 3. Size classes of respondents by location

Location	10 – 20 ha	20 – 50 ha	50 – 100 ha	>100 ha	Missing	Total
Atherton	6	13	12	13	5	49
Johnstone	1	26	30	44	8	109
Eacham	12	11	19	16	3	61
Unknown	1	3		1		5
Totals	20	53	61	74	16	224

Table 5 presents the results of applying the pivot table function to count responses to an open-ended question that asked landholders what types of crops they grow on their land. It can be seen that a number of the categories are really the same (e.g. Banana and cane; or Cane, bananas, or cane and bananas), but slight differences in the way they have been entered means that the pivot table function reads them as different categories.

There two steps to addressing this problem. The first is to be consistent when entering the responses into the database. A hard (i.e. paper) copy of the questionnaire can be used to record new responses to each question as they are being entered.

This copy can be consulted when recording responses to open-ended questions, or categorising responses to nominal questions that have an 'other' category that is effectively open ended. This ensures that consistent names are given to the same responses. The second step once responses have been entered into the database is to define categories based on examination of a range of responses, like those presented in Table 5.

Table 4. Descriptive statistics for selected land use variables

Statistic	Quality pasture	Degraded pasture	Cropping
Mean	68.2368	23.6667	68.4741
Standard Error	2.69739	3.27375	3.02579
Median	80	20	83.5
Mode	100	30	100
Standard Deviation	28.8003	19.6425	32.5887
Sample Variance	829.457	385.829	1062.03
Kurtosis	-0.754	-0.1582	-0.9025
Skewness	-0.6836	0.86976	-0.744
Range	98	70	99
Minimum	2	1	1
Maximum	100	71	100
Sum	7779	852	7943
Count	114	36	116

Table 5. Initial crop types in the responses database

Crop type	Frequency	Crop type	Frequency
None	107	Cane, bananas	2
Aloe Vera, maize and taro	1	Cane, bananas, nursery	1
Avocados	1	fruit trees	1
Banana and cane	2	Hay	1
banana, pawpaw	1	Maize	2
Bananas	15	Maize Peanuts Potatoes	1
Bananas, pawpaw	1	Maize, potatoes	1
Beans and zucchini	1	Maize, peanuts, vegetables	1
Cane	62	Mangoes	1
Cane & banana	7	Orchid	1
Cane & exotic fruit	1	Pasture seed	1
Cane & pawpaws	2	Peanuts, cane	2
Cane and pawpaws	1	Sorghum, oats and hay crops	1
Cane pawpaw	2	Sorghum, oats, rye and grass for silage	1
cane, bananas	1	Tea, cane	1
Total			223

The definition of categories is up to the researcher and depends upon the number of responses to the questionnaire and the variation in the data. Categorical data are more limiting than ordinal data in terms of the statistical analyses that can be used. One question facing researchers that wish to analyse relationships between variables defined using categorical data is how to establish a series of categories that maintain the diversity in the data yet still have sufficient responses in each category to allow the use of statistical analyses like the chi-squared test and one-way ANOVA. When carrying out chi-square tests, each cell in the table of expected responses should have at least five respondents. If more than 25% of the cells in the expected frequency table do not have five responses the test results may be unreliable.

Several new variables could be created from the data in Table 5. The simplest variable would record the presence or absence of cropping as shown in Table 6. This variable would have the advantage of having many respondents in each category, and the disadvantage of losing a lot of information about the types of crops that are grown.

Another way to classify the data could include some more details about the types and mixtures of crops commonly grown (Table 7).

Table 6. Number of respondents growing crops on their land

If crops grown	Frequency
Crops	121
No crops	103

Table 7. Number of respondents growing crops on their land

If crops grown	Frequency
No crops	107
Cane only	62
Cane and other crops	22
Crops other than cane	33

The resulting classification scheme has four categories and reasonable numbers of respondents in each category. The implications of different classification schemes for categorical data will be further examined in the following section.

5. DATA ANALYSIS

The Excel program contains a number of basic data analysis functions including chi-square tests for independence. An 'add-in' can be loaded with additional statistical functions including t-tests, z-tests, correlation, covariance, regression and ANOVA. In this section the chi-square test is examined.

The relevant application of the chi-square for this discussion is to assess whether there is a relationship between two sets of nominal (categorical) data, known as the *chi-squared test of independence*.

The null hypothesis for this test is that there is no relationship between the two data categories[1]. To run the test in Excel the user has to calculate the expected frequencies of values under the null hypothesis in a table and compare these values with the distribution of observed frequencies. The Pivot Table function makes it easy to compile the table of actual values. An example is provided in Tables 8 and 9. The expected frequencies are calculated by multiplying the row total by the column total then dividing the result by the grand total. Thus the expected frequency of those who have primary school education and have not planted is calculated as (33 x 123)/196 = 20.71.

The chi-square test for independence is performed using the CHITEST function in Excel. The chi-square statistic is calculated as the sum over the rows and columns of: (observed frequency – the expected frequency)^2 / expected frequency. The calculated statistic is then compared to a critical value for the chi-square statistic for the relevant number of degrees of freedom (the product of number of rows less 1 and number of columns less 1). The CHITEST function returns the probability for a chi-square statistic for the relevant number of degrees of freedom. If the probability of the statistic is less than the designated significance level (usually set at 0.05), then the null hypothesis is rejected and it is concluded that there is a relationship between the two variables or categories. In the above example, with the probability of the chi-square statistic of 1.3^{-5} or 0.000013, it is concluded that there is a difference in planting behaviour between those with different levels of formal education. In other words, those with diplomas and degrees are more likely to plant trees than those with primary and secondary education.

As mentioned in the preceding section, the categorisation scheme used to reclassify data for analyses has important implications for the types of statistical tests that can legitimately be carried out.

Difficulties may arise in surveys with relatively small samples if researchers attempt to test relationships between ordinal variables with more than a few categories each.

Consider the example of the different ways of categorising the types of crops grown by landholders in Tables 6 and 7. The data set of responses to the survey does not have sufficient information to legitimately test the relationship between the crop types grown by respondents and their level of formal education (Tables 10 and 11). More than 25% (5/16) of the cells in the table of expected values (Table 11) have a value of less than 5. The probability of obtaining the chi-square statistic in this case is 0.02, which is less than 0.05, but the result should not be reported since the test is invalid.

In the example below there are too many categories in each variable to carry out a chi-square test. The alternative is to reduce the number of categories in one or both of the variables. An example of this procedure is illustrated in Tables 12 and 13.

[1] Technically, this is a test of whether the joint probability distribution is the product of the univariate probability distributions for each of the variables. Further details can be found in Harrison and Tamaschke (1993, pp. 222-224).

Table 8. Actual frequency of respondents who have planted more than 30 trees by education classes

Education category	If planted		Total
	No	Yes	
Primary school	23	10	33
Secondary school	82	31	113
Diploma	12	11	23
Degree	6	21	27
Total	123	73	196

Table 9. Expected frequency of respondents who have planted more than 30 trees by education classes

Education category	If planted		Total
	No	Yes	
Primary school	20.71	12.29	33
Secondary school	70.91	42.09	113
Diploma	14.43	8.57	23
Degree	16.94	10.06	27
Total	122.99	73.01	196

In the second example (Tables 12 and 13,) the reduction in categories of the cropping variable means that there is sufficient responses in each cell to use a chi-square test. For this example the probability of the chi-square statistic returned by the test is less than 0.0001. Thus the statistical decision can be made to reject the null hypothesis, with the practical inference that there are different proportions of the population growing crops when comparing those with different levels of formal education. Inspection of the observed and expected frequencies used in the test tells us that those with lower levels of formal education are more likely to grow crops than those with higher levels of formal education. The combining of categories involves some loss of information about relationships between the variables and thus diminishes our understanding about the relationships.

It can be seen from Table 10 that no respondent with a degree reported growing only sugarcane as a crop. If the researcher thinks that this point is important and worth pursuing then it possible to construct another variable for the types of crops grown by respondents, with three categories.

As the survey has sufficient respondents who report growing sugarcane only this category can be retained, as can the category of respondents who grow no crops. The third category combines those who grow sugarcane and other crops, and those who grow other crops but no sugarcane. The observed frequency table of those with different levels of education by different crop growing categories would then appear as in Table 14, and the expected frequencies would be as in Table 15.

Table 10. Actual frequency of cropping categories by education classes

Cropping category	Education category				Total
	Primary	Secondary	Diploma	Degree	
No crops	12	42	15	21	90
Cane only	14	39	3		56
Cane and …	2	15	1	1	19
Other	5	17	4	5	31
Total	33	113	23	27	196

Table 11. Expected frequency of cropping categories by education classes

Cropping category	Education category				Total
	Primary	Secondary	Diploma	Degree	
No crops	15.2	51.9	10.6	12.4	90
Cane only	9.4	32.3	6.6	7.7	56
Cane + other	3.2	11.0	2.2	2.6	19
Other	5.2	17.9	3.6	4.3	31
Total	33.0	113.1	23.0	27.0	196

Table 12. Actual frequency of crop growing categories by education classes

Education category	Crops	No crops	Total
Primary school	21	12	33
Secondary school	72	41	113
Diploma	8	15	23
Degree	6	21	27
Total	107	89	196

Table 13. Expected frequency of crop growing categories by education classes

Education category	Crops	No crops	Total
Primary school	18.0	15.0	33
Secondary school	61.7	51.3	113
Diploma	12.6	10.4	23
Degree	14.7	12.3	27
Total	107.0	89.0	196

Table 14. Actual frequency of those with different levels of education by different crop growing categories

Education category	No crops	Cane only	Cane and other crops	Total
Primary school	12	14	7	33
Secondary school	42	39	32	113
Diploma	15	3	5	23
Degree	21		6	27
Total	90	56	50	196

Table 15. Expected frequency of those with different levels of education by different crop growing categories

Education category	No crops	Cane only	Cane and other crops	Total
Primary school	15.2	9.4	8.4	33
Secondary school	51.9	32.3	28.8	113
Diploma	10.6	6.6	5.9	23
Degree	12.4	7.7	6.9	27
Total	90.1	56.0	50.0	196

The probability for the chi-square statistic for the data in Tables 14 and 15 is 0.010. As this is less than the critical probability of 0.05, the decision is made to reject the null hypothesis, i.e. there is a significant difference in terms of the types of crops grown by respondents with different levels of formal education. It can thus be concluded that this type of difference exists in the underlying population Comparison of the observed and expected frequencies suggests that the likely source of the difference is the lower than expected frequency of those with degrees growing only cane.

6. DATA REPORTING

The preceding section has illustrated some forms of summary tables used to present data. The way in which data are presented depends upon the type of report being compiled and the types of statistical tests performed. When survey data are analysed, the presentation can occur on a number of levels (as illustrated in Figure 1). Reporting of survey responses should cover:

- responses to survey questions;
- transformation of response data in preparation for data analysis; and
- results of all analyses of relationships between variables prepared from the survey responses.

The first stage of reporting is to summarise responses to each question used in the survey before they are modified. Most survey reports have a section describing on the types of respondents to the survey; tables summarising the data collected about the socio-economic characteristics of the respondents can be used to describe respondents as well as discuss the potential of non-response bias. Where the survey is large – in terms of sample size and number of questions – the researcher may use appendices to report large amounts of data and concentrate on those analyses and descriptions that are most relevant to the research questions. In the case of the examples used in this paper (drawn from a survey of landholders tree planting and management attitudes and behaviour), the initial data should include description of the socio-economic characteristics of respondents. The descriptive sections for a report should be organised to present the responses by topics covered in the survey. The various topics in this case included the reasons landholders plant trees, restrictions to tree planting on their land, their past and intended planting behaviour, their attitudes to tree planting on a regional scale, and their attitudes to past and potential tree planting incentive and assistance schemes.

In the initial descriptive reporting of survey findings, the responses should be reported as an average or mean figure for all respondents. Where the survey has covered clearly different political or geographic areas, or clearly different types of people in socio-economic terms, then the descriptions of responses may be organised to illustrate these differences in the respondents. In the case of the north Queensland survey, three local government areas over two distinct bio-geographic regions were included. Two of the government areas are located in an upland area, and the third is coastal. The differences in the two types of areas arise from differences in their climates, topography and soils, as well as the farm sizes and enterprise types. Initial description of the responses to the survey showed the average responses to the various questions for all respondents and for respondents from each local government area. The presentation of these data also described tests for significant differences in characteristics of respondents in the various local government areas. An example of such information is provided in Table 16.

Using graphs is an excellent way to display data for descriptive purposes or to illustrate the results of analyses. Note that graphs in Excel are called ‘charts’. The type of graph used varies according to the type of data involved and the intentions of the researcher. The *pie chart* format can be used to illustrate the average proportion of land used for different activities as shown in Figure 5.

Where the data are in continuous or ordinal form, line graphs or histograms may be used. Line graphs are particularly useful to aid interpretation of relationships between ordinal variables and to assess if the distribution of the variable is ‘normal’ or at least linear. An example of this is shown in

Figure 6, illustrating the initial distribution of land sizes before they are standardised, with Figure 7 illustrating the distribution of the standardised values.

To obtain the graph shown in Figure 6, the raw data were first copied to a new sheet and sorted according to property size (land area). Examination of the maximum value for the variable showed that one respondent reported a property size of 6902 ha which is clearly an extreme case given that the next largest property size is only 500 ha.

Table 16. Importance placed upon various reasons for planting trees by landholders in the Johnstone, Atherton and Eacham shires

Reason for planting	Rating by shire			Sign. diffs.		Mean rating (all shires)	n	Frequency rated 5 (%)
	J	A	E	LSD	Bon.			
To protect and restore land	3.9	3.9	4.2	ns	ns	4.0	172	42
To protect the local water catchment	3.8	4.0	4.2	ns	ns	4.0	170	42
To attract wildlife and birds	3.5	3.7	3.8	ns	ns	3.6	169	31
Personal interest in trees	3.3	3.4	3.7	ns	ns	3.4	170	26
To improve the look of the property	3.2	3.5	3.6	ns	ns	3.3	170	26
To increase the value of the farm	3.1	3.2	3.2	ns	ns	3.2	166	19
To create windbreaks	2.8	3.4	3.4	A. E. > J	ns	3.1	168	25
Legacy for children or grand children	3.3	2.7	3.2	J > A	ns	3.1	166	26
To make money in the future	2.9	2.5	2.4	ns	ns	2.7	167	15
To diversify farm business	2.6	2.2	2.2	ns	ns	2.4	163	13
Superannuation or retirement fund	2.3	2.1	2.1	ns	ns	2.2	164	13
To provide fence posts	1.5	1.8	1.4	ns	ns	1.5	161	3

Notes: (1 = not important, through to 5 = very important). 'J' = Johnstone, 'A' = Atherton, 'E' = Eacham. Significant differences between means for each shire were tested using least square difference (lsd) and Bonferroni tests (P > 0.05). Significant differences between mean ratings for responses for each question were tested using the Bonferroni test. Overlapping lines indicate means which are not significantly different from each other. The mean rating for all shires includes five responses that could not be classified by shire.

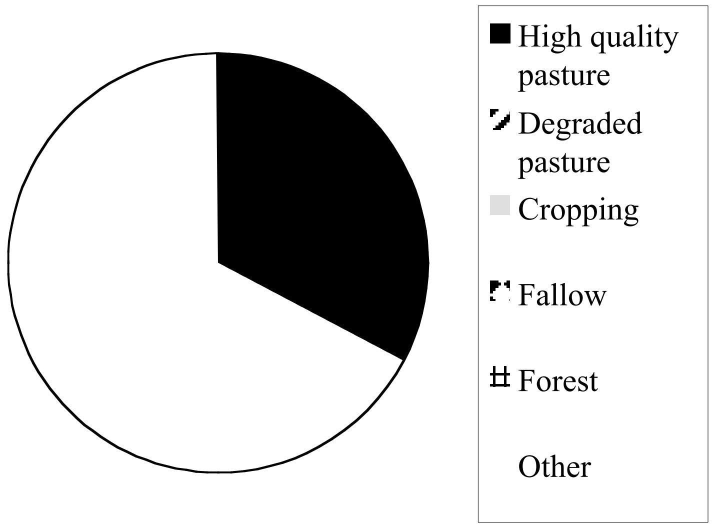


Figure 5. Average proportion of landholding used for various purposes in far north Queensland

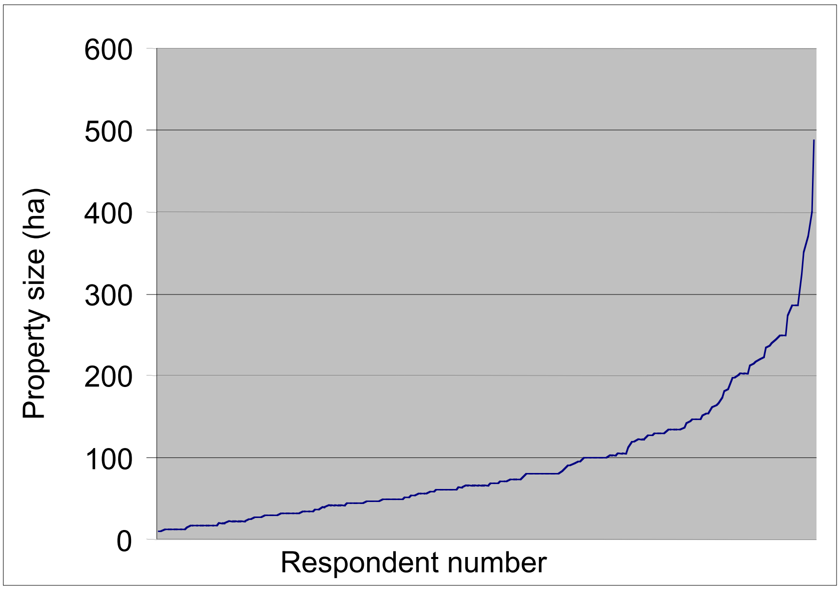


Figure 6. Distribution of values for the variable Landsize

The graph used to illustrate the distribution of the variable therefore dropped the largest value as the graph scale becomes useless when it is included. The shape of the distribution is parabolic indicating that it could be transformed to an approximately linear cumulative distribution using the Log10 function (i.e. which calculates logarithms to the base 10) in Excel. The data for the variable were transformed by taking the Log10 of the initial values and a new variable LogSize was created. The distribution of this new variable is illustrated in Figure 7.

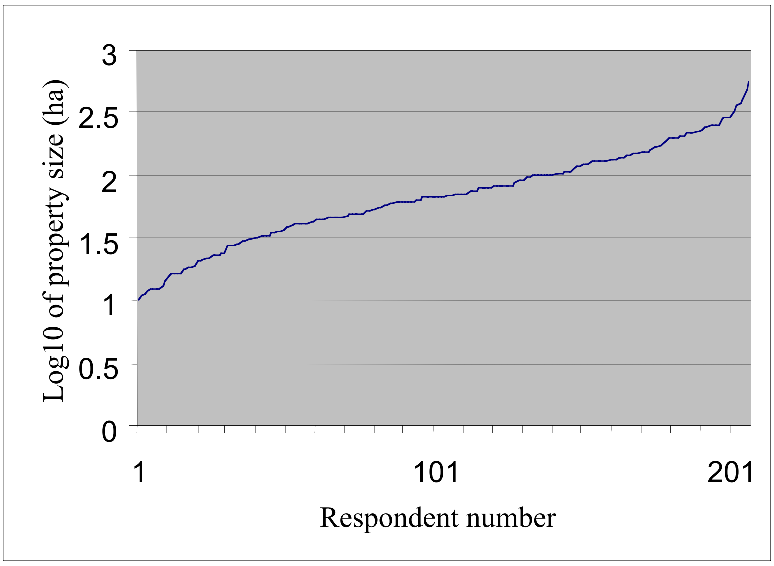


Figure 7. Distribution of values for the variable LogSize

When copying and pasting graphs from Excel to Word (or PowerPoint), open both the Excel file from which the graph is to be taken and the Word file into which it is to be placed. Copy the graph using the 'copy' function under the 'Edit' menu in Excel, then use the 'Paste special' function under the 'Edit' menu in Word to select the format used to save the graph in the Word document. Using the 'picture' format for the graphs creates the smallest file size, but does not maintain a link with the Excel file used to create the graph, and is more difficult to edit than a graph saved as an 'Excel object'.

7. CONCLUDING COMMENTS

Modern statistical packages provide a convenient means to store survey data and powerful facilities of descriptive and statistical analysis. Individual researchers tend to have their favourite data analysis packages, although Microsoft's Excel spreadsheet package and SPSS are widely used. Familiarity with statistical packages requires practice in their use, but some simple steps can be laid down for new users, as set out in this module. It is critical to plan the types of analysis intended when developing the questionnaire for a survey.

REFERENCES

Emtage, N. F., J.L. Herbohn, S.R. Harrison, and D.B. Smorfitt (in prep.), 'Landholders attitudes to farm forestry in far North Queensland: report of a survey of landholders in Eacham, Atherton and Johnstone shires', Rainforest Cooperative Research Centre, James Cook University, Cairns.

Harrison, S.R. and Tamaschke, R.H.U. (1993), *Statistics for Business, Economics and Management*, Prentice-Hall, New York.

5. Assessing the Financial Viability of Farm Forestry: A Tutorial Exercise

Nick Emtage

This module is designed to aid in the understanding of the concepts and techniques used to assess the financial viability of a farm forestry planting, and to develop basis computer skills in assessing the financial viability of farm forestry plantings. The vehicle for these purposes is a simple financial model of farm forestry, from which the net present value and internal rate of return can be derived. The data entry and financial analysis procedure is set up to be followed step by step to derive these financial performance criteria, using the spreadsheet package Excel. This module provides background to operation of the Excel-based Australian Cabinet Timbers Financial Model (ACTFM).

1. FINANCIAL MODELLING OF SMALLHOLDER FORESTRY

The Australian Cabinet Timbers Financial Model (ACTFM) was developed by a team of researchers from the Rainforest Cooperative Research Centre, at the University of Queensland and James Cook University (see references). It is designed to allow landholders and extension officers to assess the potential viability of single and mixed species cabinet timber plantations in far north Queensland. The model was developed using the Microsoft Excel spreadsheet package, and requires this software to run. A Visual Basic version of the model (the Australian Farm Forestry Financial Model, AFFFM) has also been developed, under a project funded by the Rural Industries Research and Development Corporation (RIRDC). This extends the analysis to take into account the whole-farm financial situation (i.e. costs and returns to agriculture, management of native forests and plantations, and utilization of farm and loan finance), Both financial models – the ACTFM and AFFFM – are currently available free of charge through the Rainforest CRC, although further development and testing is continuing for the latter model.

To help understand how these models operate, increase understanding of the concepts used, and to assist workshop participants in developing their own model, an exercise has been developed will involves *the construction of a basic financial model*. For this exercise,

- All the information needed to develop the model is listed in Table 1.
- It is recommended that workshop participants work in teams of two or three. Preferably these teams should include one experienced computer user and one inexperienced user each. To facilitate learning, the inexperienced user should operate the computer.
- The various steps to carry out model development outlined below are to be followed. Data should be entered in the same row and column location as in the instructions; otherwise, it will become difficult to follow the cell references in the instructions.

2. THE FORESTRY SCENARIO TO BE EVALUATED

The case study for this exercise is a simple farm forestry scenario limited to a single species, with only one commercial harvest, and a single price paid for all timber on the stump (stumpage price). The following *costs* and *timing of management activities* of Table 1 are to be used.

Table 1. Plantation costs and revenues and other input data

Establishment and maint. costs ($/ha)	Year of cost	Amount
Establishment cost	0	$2995
Maintenance costs		
Post plant weed control	1	$1,310.00
Post plant weed control	2	$812.00
Post plant weed control	3	$213.00
First prune (plus certification)	4	$880.00
Second prune (plus certification)	8	$648.60
Third prune (plus certification)	12	$864.40
Thinning	8	$501.00
First harvest marking and inventory	44	$80.00
Annual costs		
Insurance	$30	
Fencing and road maintenance	$25	
Total	$55	
Project revenues and other parameters		
Mean annual increment (MAI) = 19 cubic metres/ha/year		
Revenue = MAI * harvest age * stumpage price		
Harvest age = 44 years		
Price = $85/cubic metre (stumpage)		
Discount rate = 6%		

The exercise will involve setting up a spreadsheet containing these data, deriving annual cash flows, and calculating the net present value (NPVC) and internal rate of return (IRR). The model will then be used to examine the effect on NPV and IRR of changing cost and revenue parameters and the discount rate.

3. STEPS IN SETTING UP THE SPREADSHEET MODEL

The following steps are to be followed to set up the spreadsheet model for assessing the potential financial returns from farm forestry:

1. *Start the computer and open the Excel spreadsheet program* from the 'Start' menu at the bottom left of the screen. (Click 'Start', hold the left button down and move the mouse pointer up to 'Programs', then highlight 'Excel' before releasing the mouse button). A new workbook will appear, as in Figure 1.
2. Insert the floppy disk supplied into the 'A' drive. *Open the file named 'Financial exercise'* from the floppy disk. (If this disk is not available, the data of Figure 2 will need to be typed into the spreadsheet.) The computer screen will then appear as in Figure 2. (Use the mouse to access the 'File' menu, holding the left mouse button down highlight 'Open…' then release the button. Use the 'Look in' drop-down text box to highlight the 'A' drive. Release the mouse button, then double-click the name of the file with the left mouse button and the file should open).
3. *Enter the costs* for each year in the maintenance costs column (column F). Establishment costs go into year 0 (cell F2). The costs for each year could be typed directly into the cells. However, the make the maintenance costs column update automatically when changes are made to the cost parameters in column C, it is necessary to enter cell addresses rather than numerical values. This can be achieved by typing the entry '=C4' (without the inverted commas) in cell F2, '=C8' in cell F3, '=C9' in cell F4, and so on. A simpler approach is to first click on a cell in column F (e.g. F2) then type '='. Next click on the cell in column C which relates to that year (here C4). Press 'Enter' and the cell should now show the value (e.g. $2,995) (see Figure 3). The

same procedure can be carried out for each of the other maintenance costs (for years 1-4, 8, 12 and 44).

4. Click on cell G2 again and position the mouse pointer over the square at the bottom right corner of the cell. The pointer should change from a double lined cross shape to a single line cross. When this happens, hold down the left mouse button and drag the pointer straight down the column so that the cells are within a shaded frame. Go to cell G46 then release the mouse button. Now the value of annual costs should appear in G column for each year.

5. *Add the total costs* for each year in column H. First, click on the top cell in the 'total costs' column (H2). Type '=', then click on the adjacent cell for maintenance costs (F2). Next type '+', then click on cell G2 to add the total annual costs for the year. Press enter. Next click on cell H2 again and position the mouse pointer over the square at the bottom right corner of the cell. The pointer should change from a double lined cross shape to a single line cross. When this happens, hold down the left mouse button and drag the pointer straight down the column so that the cells are within a shaded frame. Go to cell H46 then release the mouse button. Now the total value of costs for each year should appear in H column. Note that omitting the '$' sign in the reference to a cell allows the program to change the formula as you copy down the column so that it always refers to the adjacent two cells rather than the cells referred to in the original formula.

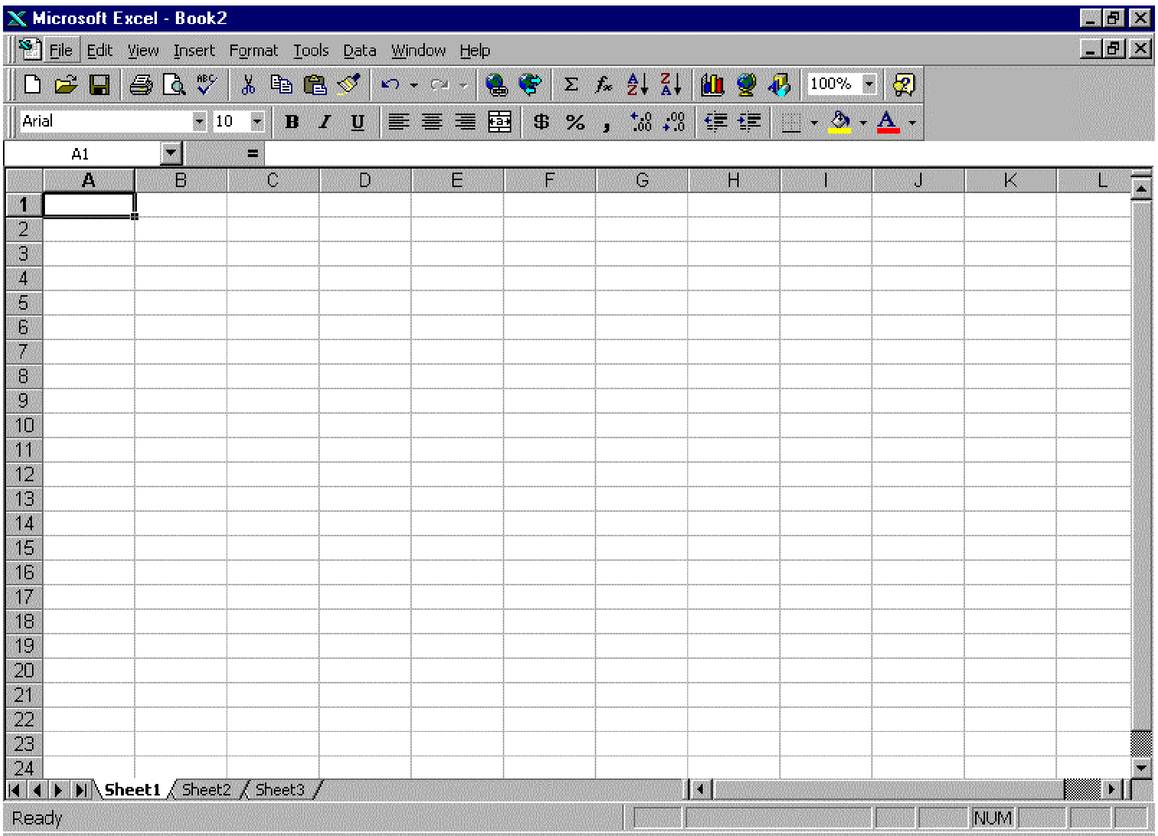
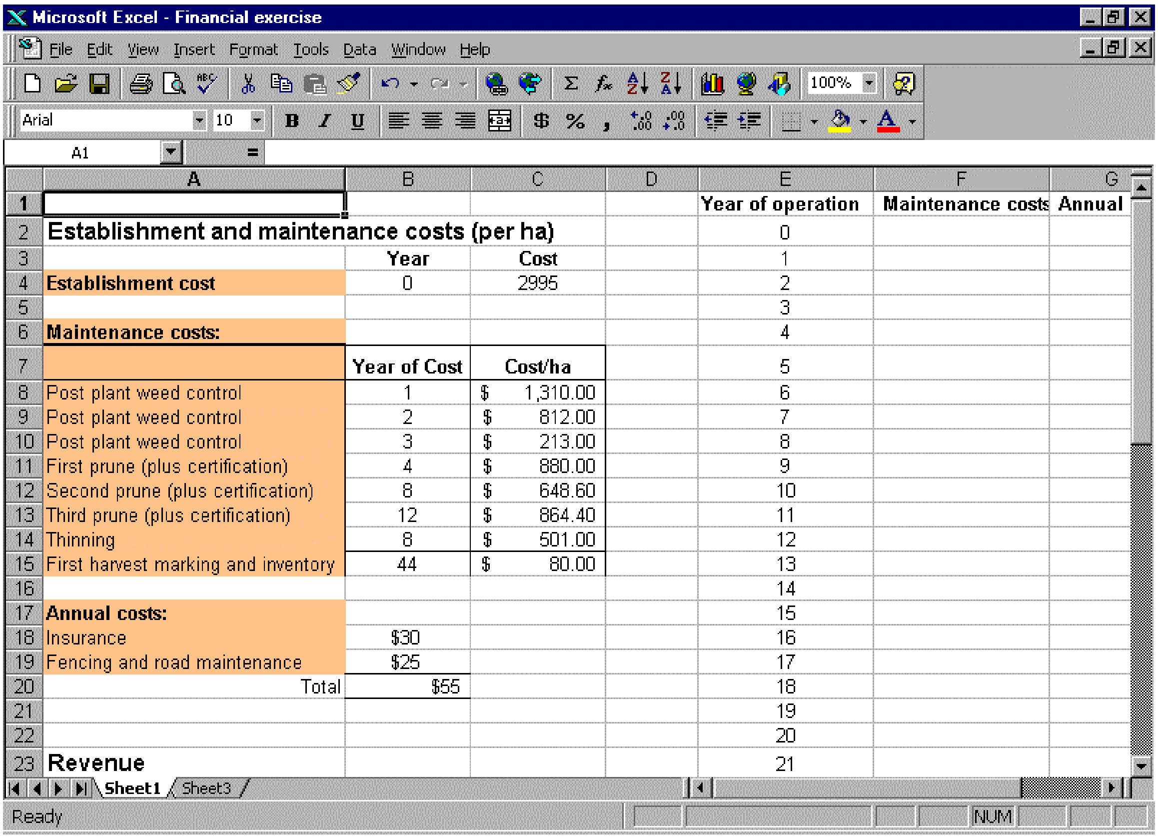


Figure 1. New workbook in Excel



	A	B	C	D	E	F	G
1					Year of operation	Maintenance costs	Annual
2	Establishment and maintenance costs (per ha)				0		
3		Year	Cost		1		
4	Establishment cost	0	2995		2		
5					3		
6	Maintenance costs:				4		
7		Year of Cost	Cost/ha		5		
8	Post plant weed control	1	$ 1,310.00		6		
9	Post plant weed control	2	$ 812.00		7		
10	Post plant weed control	3	$ 213.00		8		
11	First prune (plus certification)	4	$ 880.00		9		
12	Second prune (plus certification)	8	$ 648.60		10		
13	Third prune (plus certification)	12	$ 864.40		11		
14	Thinning	8	$ 501.00		12		
15	First harvest marking and inventory	44	$ 80.00		13		
16					14		
17	Annual costs:				15		
18	Insurance	$30			16		
19	Fencing and road maintenance	$25			17		
20	Total	$55			18		
21					19		
22					20		
23	Revenue				21		

Figure 2. The financial exercise file

6. *Calculate the revenues* from the sale of timber and enter this amount into the relevant cell in the 'revenue' column (in cell I46). Timber revenue is the product of timber price, mean annual increment and harvest age, here B30 multiplied by B31 multiplied by B32. Note that when we nominate cell addresses for multiplication, we are really referring to cell contents. Also, the asterisk is the multiplication sign adopted (to avoid confusion with the letter 'x'). Once the cursor is clicked on cell H46, the calculation '=B30*B31*B32' could be typed directly into this cell. Alternatively, we could type '=', click on cell B30, type '*', click on cell B31, type '*', click on cell B32, and then press Enter. Since cell references rather than actual parameter values have been entered into cell I46, the revenue will be updated automatically if a change is made to any of the three parameters (timber price, MAI and harvest age).

7. *Calculate the net cashflows for each year.* To do this, click on cell J2. Type '=', then click on cell I2. Next type '-', then click on cell H2. Press 'Enter', and the cashflow for year 0 will be shown in cell J2 (see Figure 5). To copy this formula to the rest of the cells in the cashflows column, click on cell J2 again. Next place the mouse pointer over the square on the bottom right corner of the cell and the pointer should change from a double lined cross shape to a single line cross. When this happens, hold down the left mouse button and drag the pointer straight down the column so that the cells are within a shaded frame. Go to cell J46 then release the mouse button, and the formula will be copied. The values for the scenario should appear as in Figure 5.

8. *Calculate the net present value (NPV)* of the scenario. This could be entered in say cell J47, with a caption typed to the cell to the left saying 'NPV ='. In cell J47, type '=NPV(B33,J3:J46) + J2'. Here B33 is the discount rate. (Note that is the percentage sign is not included, the number 0.06 has to be typed into the cell; a 6 would be interpreted as 600%.) The formula calls up the NPV function

already available in the spreadsheet to derive the discounted sum of the cash flows. Note that the range of cells in the *argument* within the NPV function corresponds to years 1 to 44, and that the cash flow for year zero is placed outside the brackets for the NPV function. This is necessary because the NPV function regards the first cash flow in the specified range as corresponding to year 1, and discounts this value back by one year.

The cursor should be now flashing in the 'rate' box. Click on the cell containing the discount rate (B35) or type 'B35' and a reference to this cell this will appear in the 'rate' box (Figure 7).

Next click the mouse pointer in the 'Value 1' box. Select the on the first cell in the Cashflows column (J2). It will be surrounded by a flashing dashed line. Put the mouse pointer in the middle of the cell, hold down the left button and drag the pointer to cell J46. Release the mouse button and the cell reference J2:J46 will appear in the 'Value 1' box. Last, click on the 'OK' button, and the NPV formula is now active in cell B37.

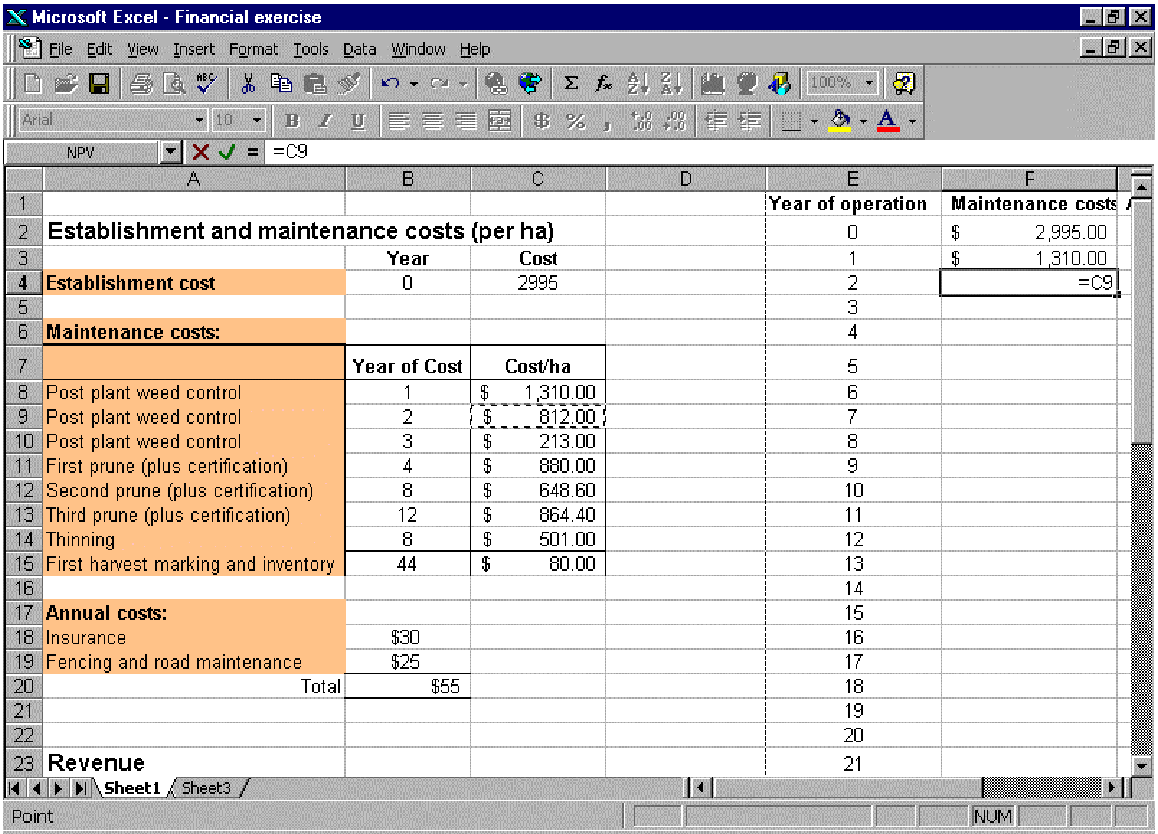


Figure 3. Entering figures into the 'Maintenance costs' column

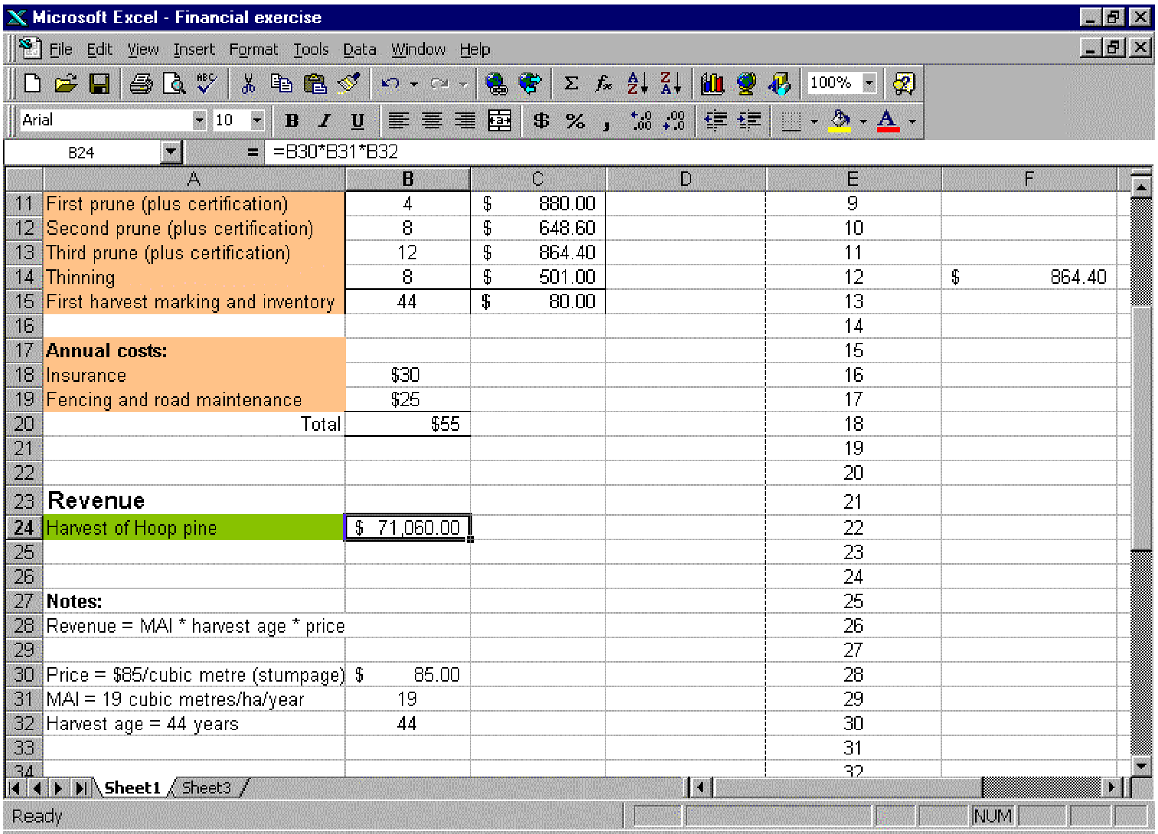


Figure 4. Revenues calculated for the scenario

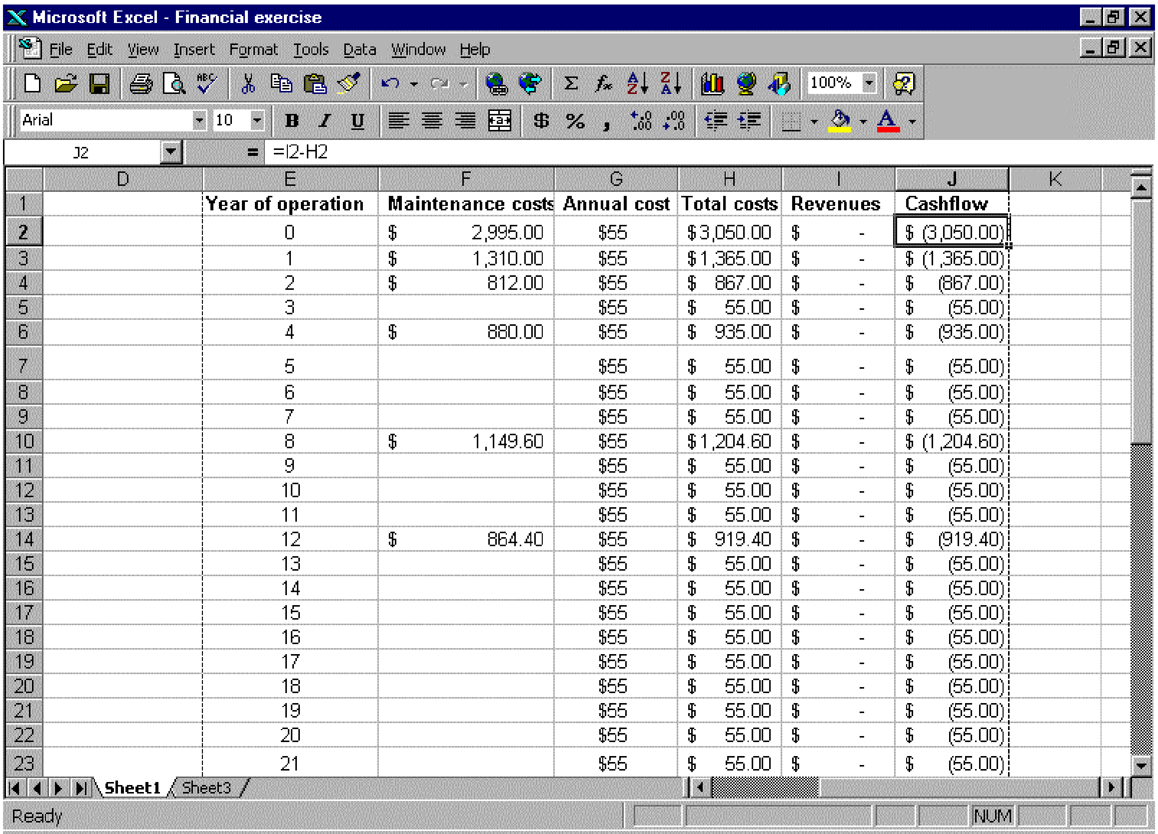


Figure 5. Completed cash flows for the scenario

9. *Calculate the internal rate of return (IRR) for the investment scenario.* This could again be obtained by using the formula paste approach but selecting IRR instead of NPV. An initial guess of the IRR is required, and a value of '0.05' (for 5%) is suggested. More simply, the IRR can be obtained by simply typing '=IRR(J2:J46)' in the chosen cell. The financial model is now complete!

4. ASKING 'WHAT IF?' QUESTIONS AND PERFORMING SENSITIVITY ANALYSIS

Once a spreadsheet has been developed, it is a powerful tool for examining the impact of variations in assumptions or parameter levels on project performance. In particular, we are able to pose questions such as '*What* will the NPV be *if* the discount rate is reduced to 5%?' Arranging such questions may be formalised into *sensitivity analysis* (exploring the sensitivity of financial performance to levels of various parameters) and *breakeven analysis* (determining the levels of parameter values such that the project just covers costs, i.e. such that the NPV is zero). As an exercise:

1. determine how the NPV changes for discount rates between 4% and 7%, in steps of 1%.

2. test how the NPV and IRR change when the timber yield is (MAI) is reduced to 15 m^3/ha and increased to 23 m^3/ha.

3. determine what timber price would be required for the project to break even, at a 6% discount rate.

REFERENCES

Emtage, N.F., Herbohn, J.L. and Harrison, S.R. (2001), Australian Farm Forestry Financial Model (Visual Basic version), The University of Queensland, Brisbane.

Herbohn, J.L., Harrison, S.R. and Emtage, N.F. (1999), 'Potential performance of rainforest and eucalypt cabinet timber in plantations in North Queensland', *Australian Forestry*, 62(1): 79-87.

Herbohn J.L. and Harrison. S.R. (2001), 'Financial Analysis of a Two-Species Farm Forestry Mixed Stand', in S.R. Harrison and J.L. Herbohn, eds., *Sustainable Farm Forestry in the Tropics*, Edward Elgar, Cheltenham, pp. 37-46.

Herbohn, J.L., Harrison, S.R., Emtage, N. and Smorfitt, D.B. (2001), 'The Australian Cabinet Timber Financial Model: A model for mixed species cabinet timber plantations in North Queensland', in A. Grodecki, ed., *Managing and Growing Trees: Farm Forestry and Vegetation Management,* Kooralbyn, 19-21 October 1998, Department of Natural Resources and Greening Australia, Brisbane, pp. 397-405.

6. An Introduction to Statistical Package for the Social Sciences

Nick Emtage and Stephen Duthy

This module provides an introduction to statistical analysis, particularly in regard to survey data. Some of the features of the Statistical Package for the Social Sciences (SPSS) are then explained, with reference to a farm forestry survey. Of necessity, this is a brief overview to the highly complex and powerful SPSS package.

1. INTRODUCTION

Computer based statistical packages are an important tool for researchers in the social sciences. The prospect of using statistics is sometimes either repugnant or simply frightening for people, yet most researchers recognise the potential utility of statistical analysis to aid them to describe, analyse, interpret and report their data. The mathematics behind statistical analysis can be daunting for those who have little formal training in either mathematics or the use of statistics. The development of specialist statistical analysis packages has greatly reduced the mathematical challenge of undertaking many analyses. It should be emphasised, however, that these packages have not reduced the need for researchers to understand the assumptions behind statistical analyses, and to be able to interpret their results. The packages have however reduced the need for researchers to be able to undertake many of the calculations that are required for statistical analyses. In this way they allow researchers to concentrate on understanding the assumptions behind the various methods, as well as the potential applications and limitations of various statistical tests.

Statistical software packages have, like other software packages, changed greatly since the advent of the personal computer a little over 20 years ago. Some of the authors still remember programming mainframe computers with paper cards. Holes were punched into the cards and these were then fed into the computer. Computers in those days were scarce, especially the big ones with four megabyte memory! Needless to say that statistical tests were difficult to perform unless the user had an advanced understanding of the mathematics required. More recent computer software packages are reasonably easy to use for people with some familiarity with computers. Most of the packages have features such as drop-down menus, 'tree' structure diagrams and on-line help systems. This said, it should be remembered that the packages discussed here are large and highly complex. While they are considerably easier to use today than they were even 10 years ago, like other large software packages, familiarity and ease of use are only developed through practice with the package. A user can become functionally proficient with a package such as Excel and Word after several weeks, use but development of a high level of expertise can take many months or even several years.

When choosing which package to use for statistical analyses a number of factors must be considered. These include the availability of a package, its cost, the functions it can perform, familiarity with the package and the availability of an expert statistician to assist with the analysis process. As discussed above, the packages take time and effort to learn, and many researchers prefer to continue using a particular package once they learned how to use it. Other factors may affect this however. Availability of a package is an important factor in deciding whether to use it or not. If an institution has already obtained the rights to use a particular package, it may be the only choice available. Buying copies of the latest versions of the specialist statistical packages is expensive, as is the cost of maintaining the license to use the package. If an institution already has a package that

can provide the functions required then the researcher may be forced to use that package despite preferences for other software because of limited funds.

Where expert statisticians are available to assist with data analyses then the preferred package of the expert is likely to be the one used. As discussed in other modules it is important to discuss research projects with expert statisticians *during their design* to ensure that the data collected will be in a format that allows the use of the desired analysis techniques. It is also important at this stage to discuss the packages available to the researcher and the time available to access a computer for data entry, analysis and reporting. Where access to the computers with the statistical software is limited it may be possible for the researcher to enter the data into a spreadsheet program like Excel and then transfer the data set to the statistics package in order to carry out the analyses. In this case it is important to have some understanding of the formatting required by the statistical package to be used so as to avoid unnecessary reformatting of the data in the statistical package. Where possible data should be entered directly into the statistical package to avoid the potential need to reformat the data.

2. THE STATISTICAL PROGRAM FOR SOCIAL SCIENTISTS (SPSS)

The SPSS Corporation first produced the SPSS software package in the early 1980's and has recently released version 11.0. It is presently one of the most commonly used statistical packages in Australian research institutions and is available at all Australian universities. The advantages of the package are its relative ease of use, its familiarity to many statistical experts and its functionality. One of SPSS's major disadvantages is its cost. The SPSS corporation appears to be progressively breaking up the program into different sections that can be purchased separately. For Australian students an individual users' license (one year) costs approximately $A100 for a 'base' student version and $A350 for a 'graduate pack' licensed for 5 years (as of March 2002). The different versions have varying analytical functions and different capacities in terms of the number of cases and variables that can be used. An institutional license costs even more, depending on the number of expected users. The different packages have licenses that also differ. In most cases licenses are set up to expire automatically after a limited period after which the package can no longer be used. The package is developed for a number of operating systems including Windows and Unix. Information about SPSS products is available on-line at www.SPSS.com.

Organisation of the SPSS package

The set-up of the version 10.0 package (used for illustration here) is organised into two main sections, for defining and entering data and for output. When defining and entering data, users can move between the 'variable' and 'data' 'views' by clicking on the tabs at the bottom of the screen. The third 'output' section opens in a separate window and displays the results of the statistical analyses. The 'output' data are saved as a separate file to the data set.

In the 'variable view' (Figure 1) the users sets up the data entry and analysis cells by naming and defining the variables included in the data set. Users are required to use names for the variables of eight or fewer characters. Names must begin with an alphabetic character. Longer descriptions of the variables can be added using the 'Labels' dialog box (Figure 2). A quick way to define the variable format (including the variable type, the number of characters used and labels) if a number of variables have a similar format is to copy the attributes of a variable then paste them into other variable fields.

Once the variables to be recorded have been named and defined the user can access the 'data view' to enter in the values for each variable. The SPSS data view looks similar to a spreadsheet program. The variables are organised as columns with each row as a single 'case' in the data set containing values for the variables relating to that case. It is common practice to use codes to enter data into the package and labels can be used to describe values where needed. For example, codes may be used to record the types of agriculture practiced on a landholding, or respondent's

educational levels. The defined labels will appear, by clicking the drop-down list arrow on the right side of the cell, and the user can select the relevant value (Figure 3).

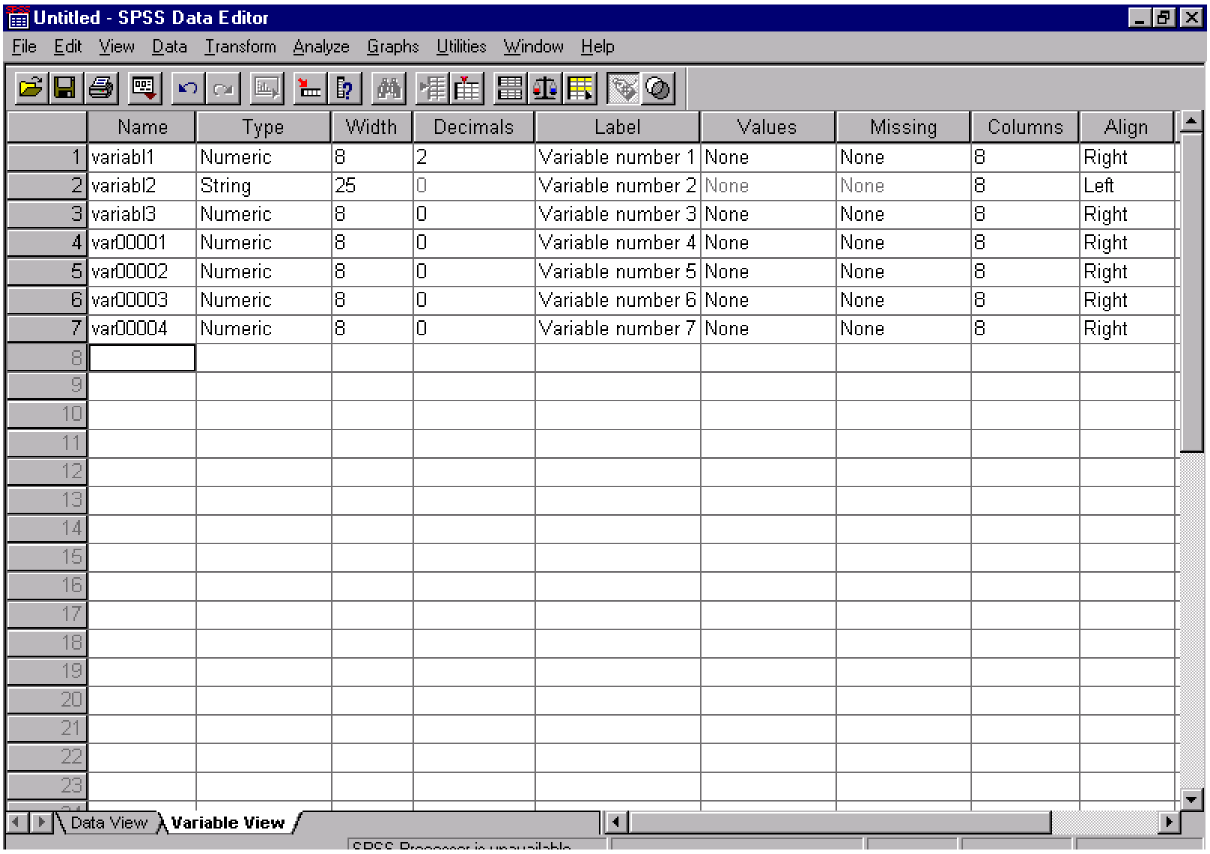


Figure 1. Variable view in SPSS 10.0

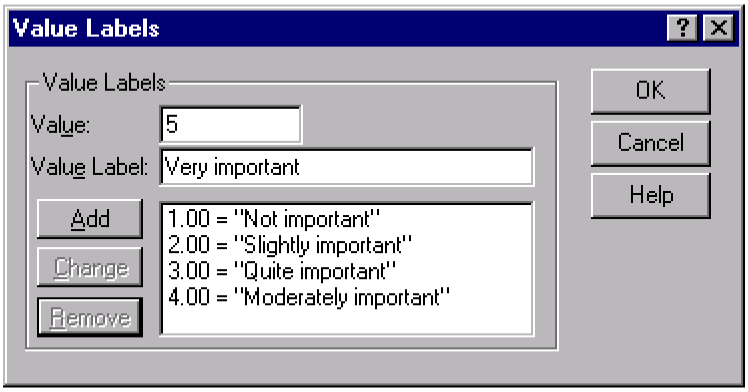


Figure 2. Defining variable labels using the Value labels dialog box

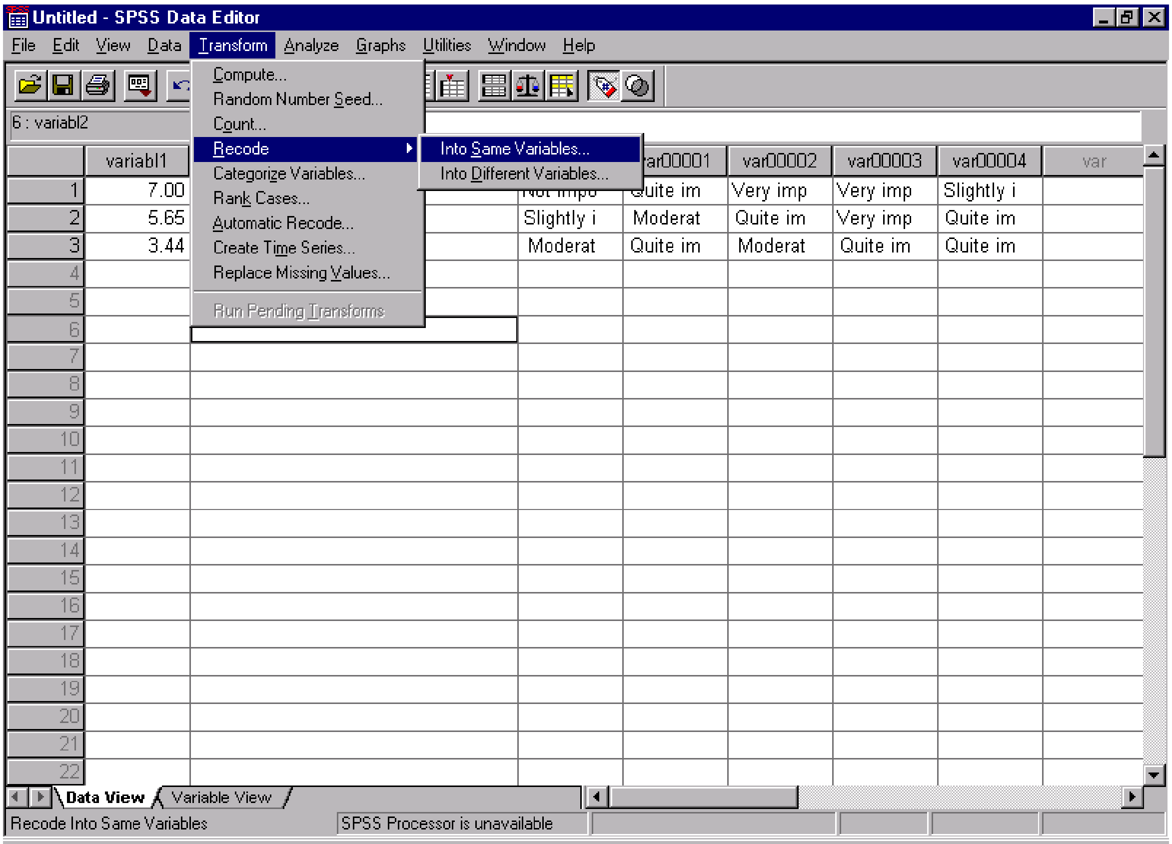


Figure 3. Entering data into the SPSS

This is handy when there is a large number of possible responses, and thus codes, for a variable, and the user cannot remember all of them. The user can choose to have the codes or the labels displayed in the data view by selecting the 'Value labels' option under the 'View' menu.

Data analysis using SPSS

The SPSS 'student pack' has a wide range of analytical functions, from basic descriptive statistics to advanced general linear modeling capabilities. Specific functions are also included to allow the transformation of variables as preparation for different tests (e.g. for creating standardised or logarithmic values, or the calculation of scales from a number of variables) (Figure 4). The use of these functions allows researchers to calculate quickly new variables based on the values of other variables, test variations in category schemes used to classify responses to 'open ended' questions, and collapse categories where necessary.

Once the data are entered into the SPSS program it is important to check the database for typographic errors that may affect the results of statistical analyses. One means of achieving this is to examine the frequencies of categorical (nominal) data, and descriptive statistics of numeric (ordinal, scale or interval) data. All of the analytical functions available in SPSS can be accessed using the 'Analyse' menu (Figure 6). If the 'Descriptive statistics' then the 'Frequencies' options are selected, the dialog box illustrated in Figure 5 appears. This dialog box enables users to select the variables for which frequencies are computed as well as control the types and, to a limited extent, the formatting of displays of the analyses.

If calculation of descriptive statistics is required, users should select 'Descriptive statistics' and the 'Descriptives' options under the Analysis menu to reveal the 'Descriptives' dialog box (Figure 7).

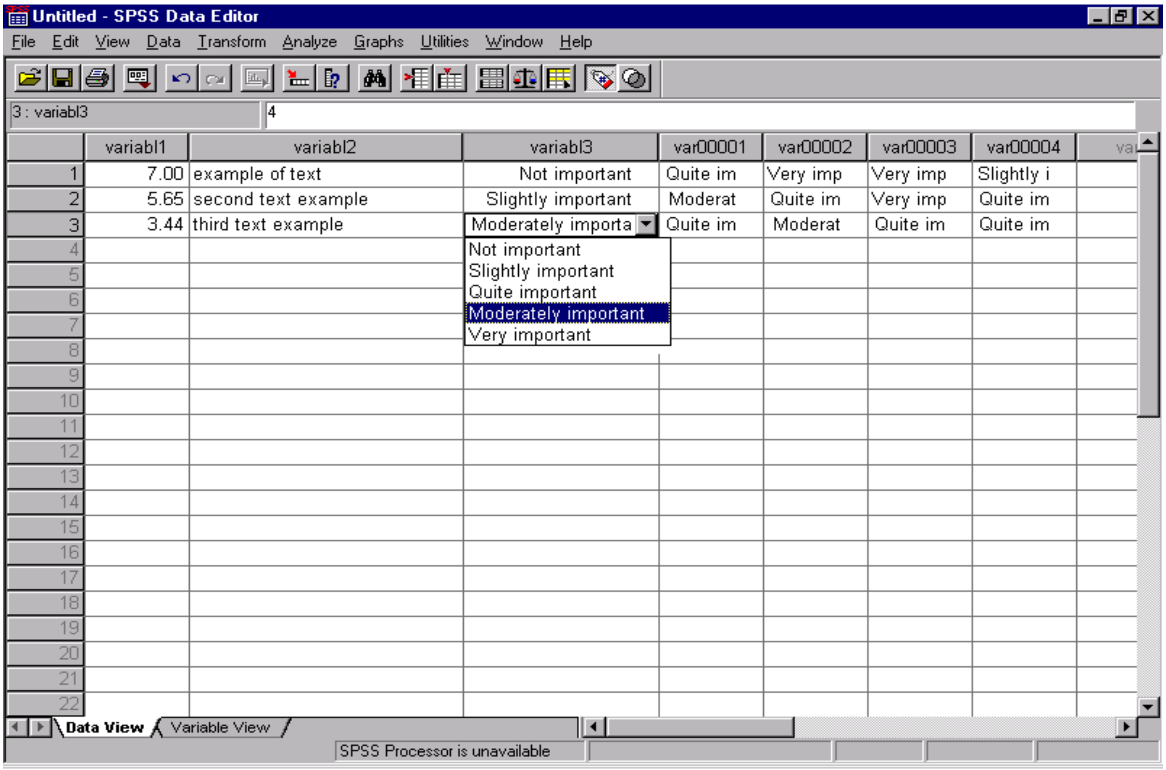


Figure 4. Data transformation functions in SPSS

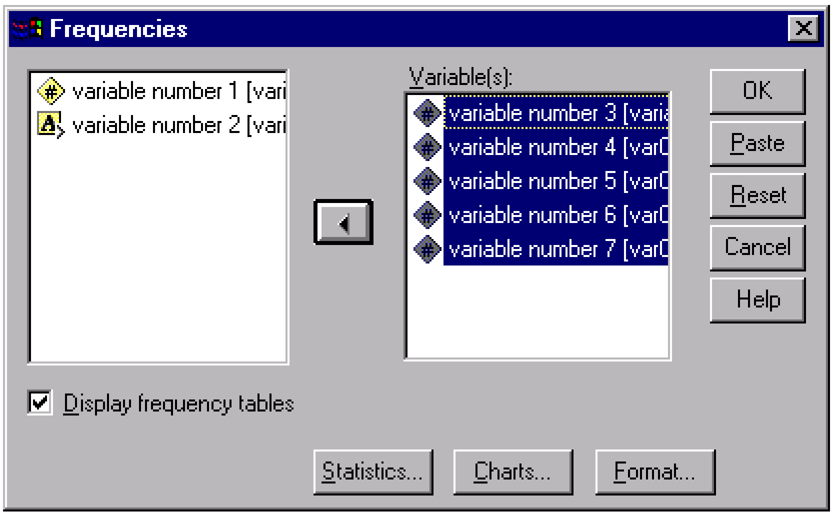


Figure 5. Frequencies dialog box

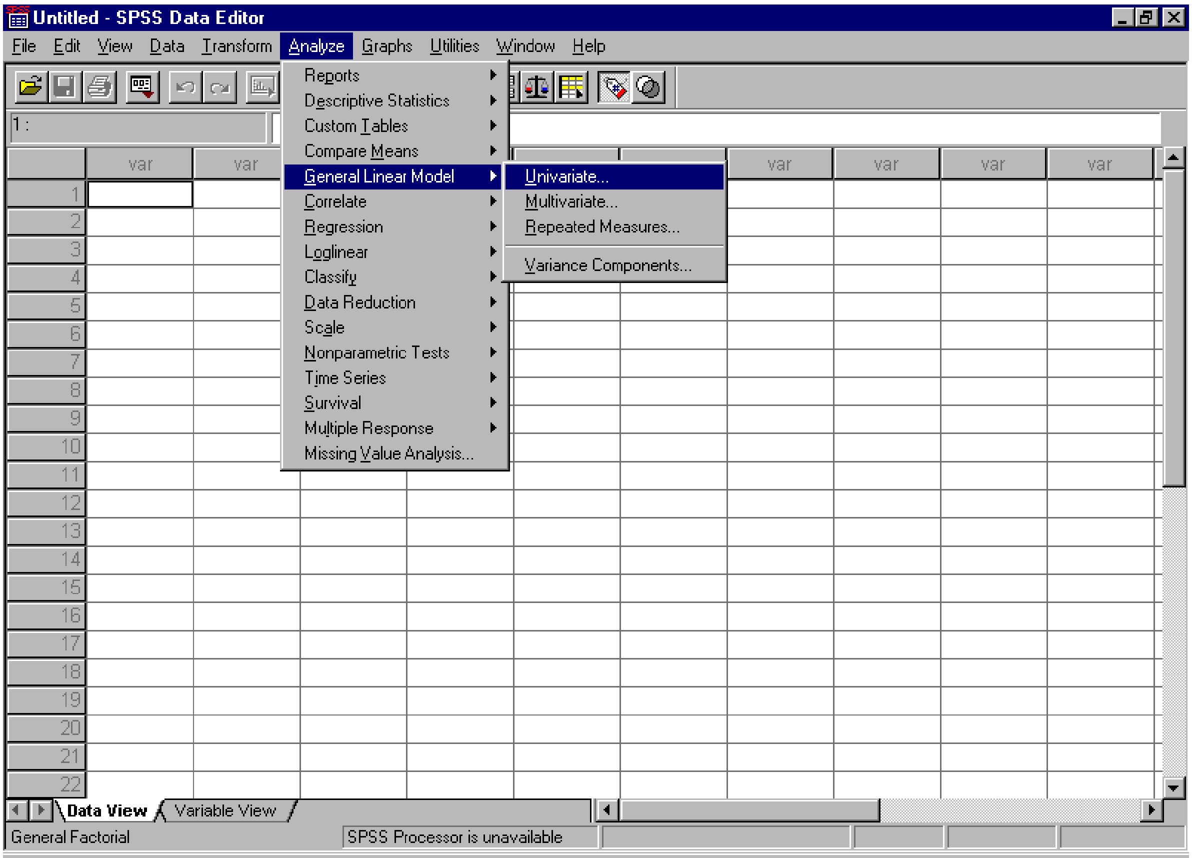


Figure 6. Analysis options available in SPSS

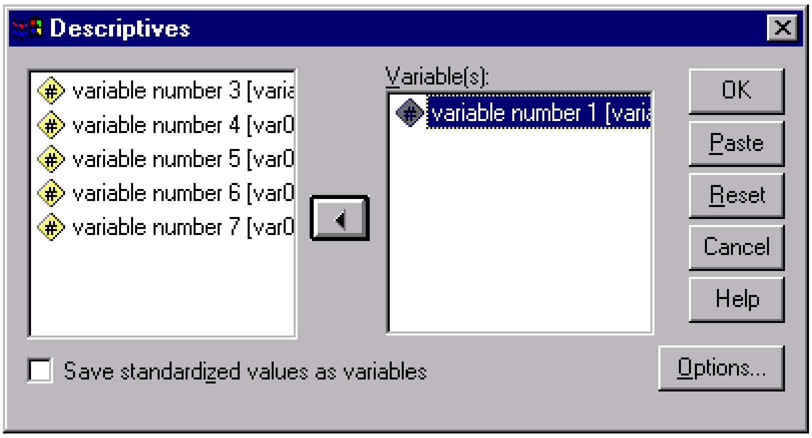


Figure 7. Descriptives function dialog box

Once the 'Descriptives' dialog box is shown, the variables to be included in the analyses are selected from the list on the left side of the box (Figure 7), and transferred to the list on the right side of the box (labeled 'Variables' in Figure 7) using the arrow in the centre of the box. The types of descriptive statistics that will be Calculated using this function can be selected by clicking on the 'Options…' button (Figure 7). This reveals the 'Options' dialog box for the Descriptives function (Figure 8).

Other analytical functions included in the SPSS student pack (Version 10) include chi-square tests, correlations, regressions, principal components analyses, ANOVA, cluster analyses, general linear modeling and more.

Whilst this paper does not attempt to provide the reader with statistical skills, the flowchart in Figure 9 may act as a guide for the reader to access quickly those functions in SPSS that will best serve their statistical analysis needs.

The analytical functions are adequate for all but the most advanced researchers or those requiring highly specific analyses. Most advanced or specific applications can be met as well, with SPSS open to manipulation via user compiled 'Sax Basic' computer code (also known as 'scripts' in SPSS). This is similar to the use of the Visual Basic programming language to develop and execute macros in Microsoft Excel. The Sax Basic language is compatible with Visual Basic for Applications.

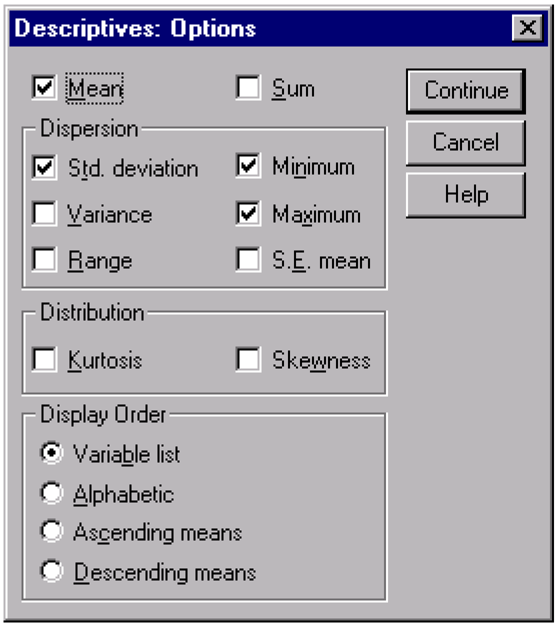


Figure 8. Options dialog box for the 'Descriptives' function dialog box

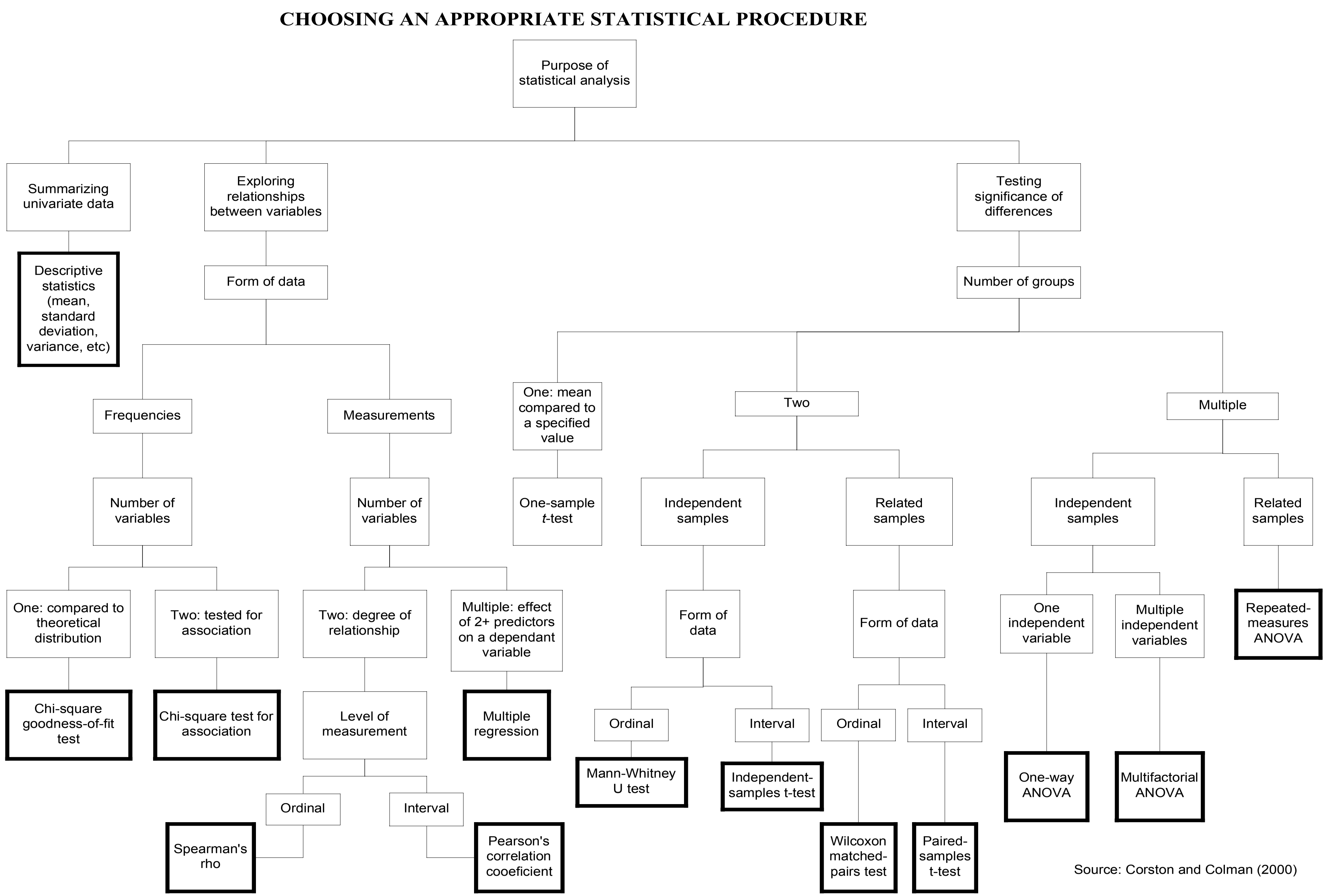


Figure 9. Choosing an appropriate statistical procedure

Using SPSS to Describe Data

Whilst computer-based statistical packages provide a high degree of functionality with regard to data analysis, they also provide a number of highly useful tools for the description and presentation of summaries of the dataset.

These functions include Descriptives and Frequencies as explained earlier and Crosstabs, also found under the Descriptive Statistics menu, and Basic Tables, General Tables, Multiple Response Tables and Tables of Frequencies all located under the Custom Tables menu item (Figure 10). It is often useful to undertake one or more of these processes before commencing data analysis to identify any weaknesses in the dataset such as poorly represented groups within the sample that may limit the statistical validity of some forms of analysis. Crosstabs are also an efficient way of presenting data summaries in research and project reports.

The charting functions available in SPSS also provide a number of techniques for the initial exploration and the presentation of data. Scatter Plots (Figures 11 and 12) can be used to identify quickly the presence and nature of any correlations between variables while Histograms (Figures 13 and 14) can be used to present a graphical representation of the shape of the distribution of the data for important variables.

There is a reasonable amount of literature available to assist users of the SPSS package produced by the SPSS Corporation and by independent authors. The tutorial and help facilities for the package are comprehensive, generally easy to understand and include the on-line Statistics Coach and Syntax Guide.

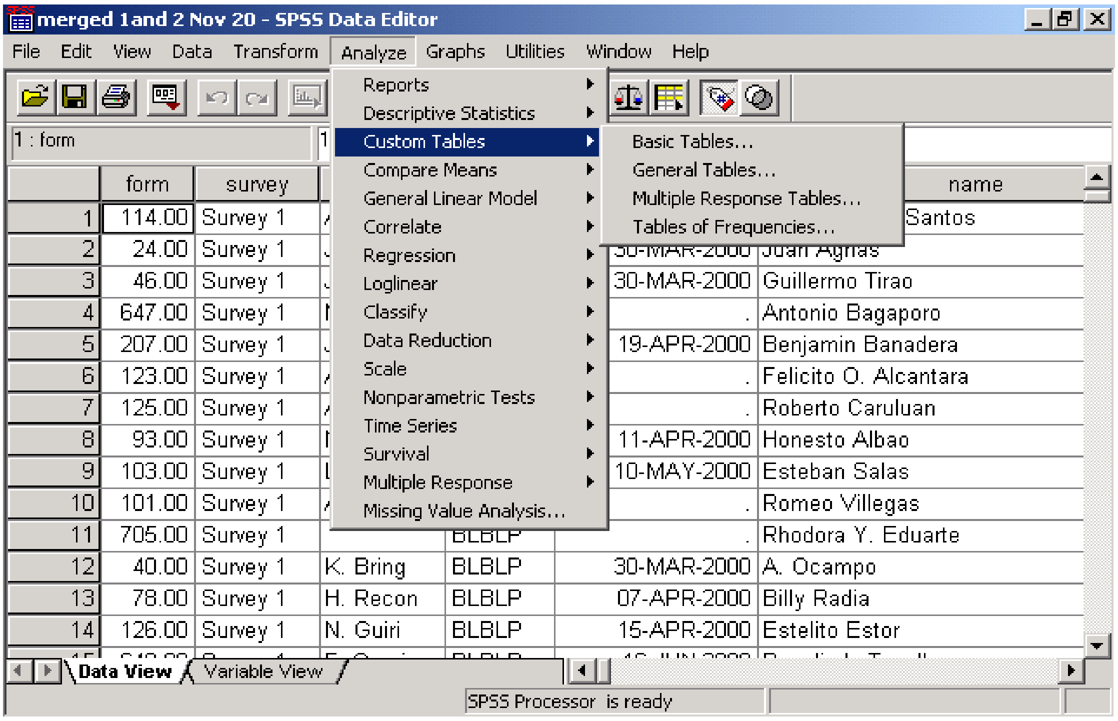


Figure 10. Custom Tables drop-down menu selections

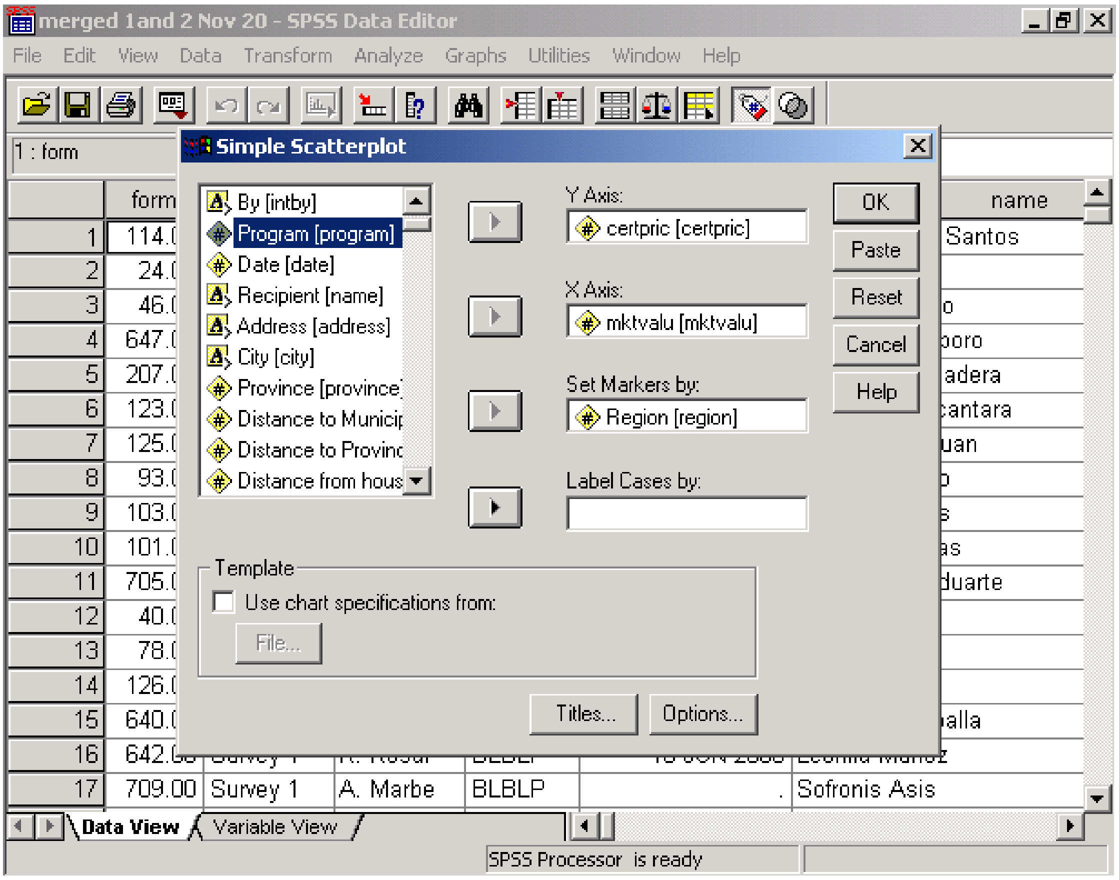


Figure 11. Design options for a Simple Scatterplot

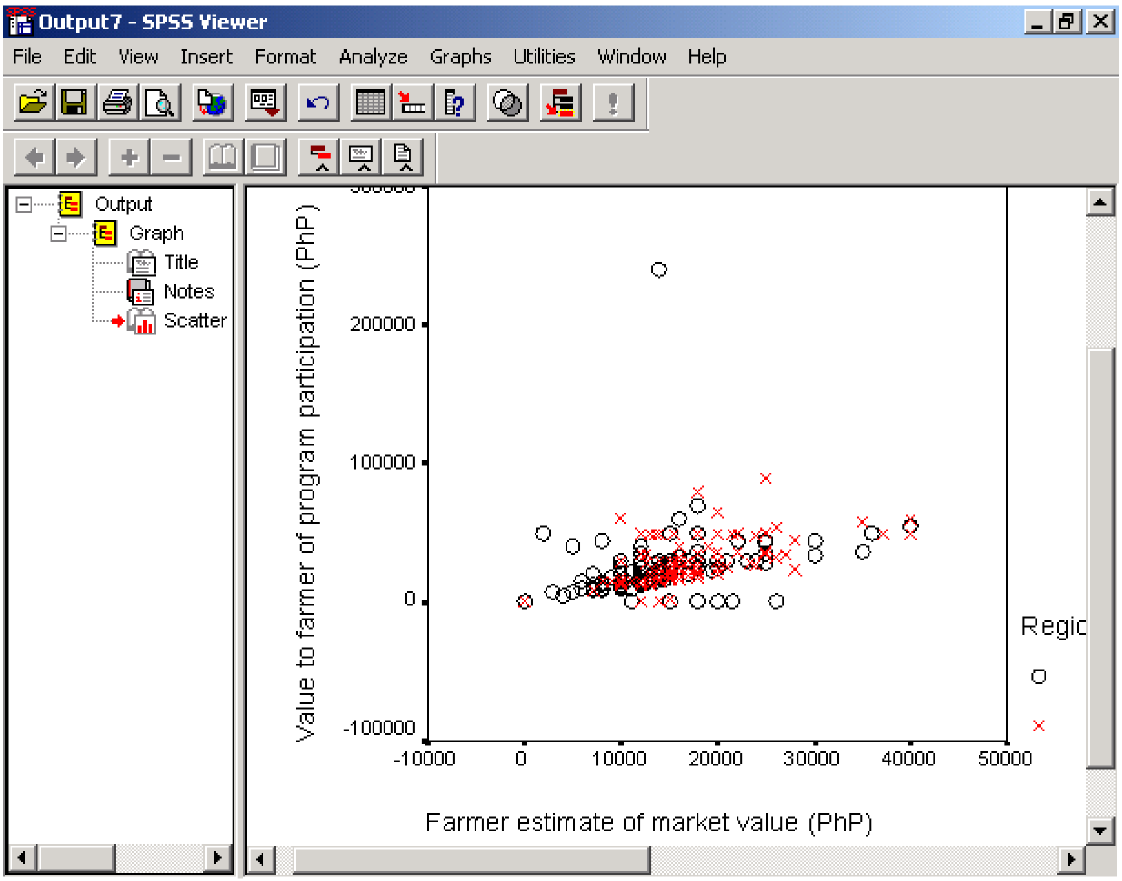


Figure 12. Simple Scatterplot displayed in the Output Viewer

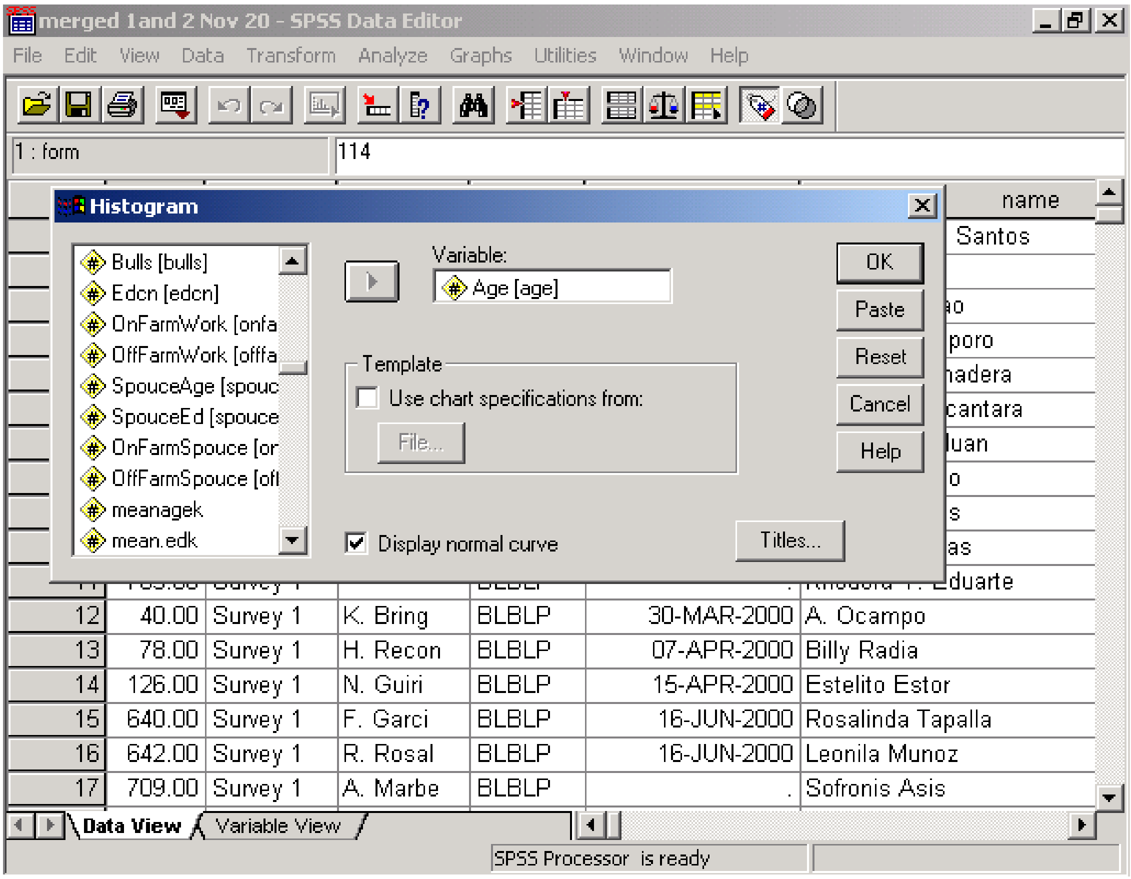


Figure 13. Design options for a histogram

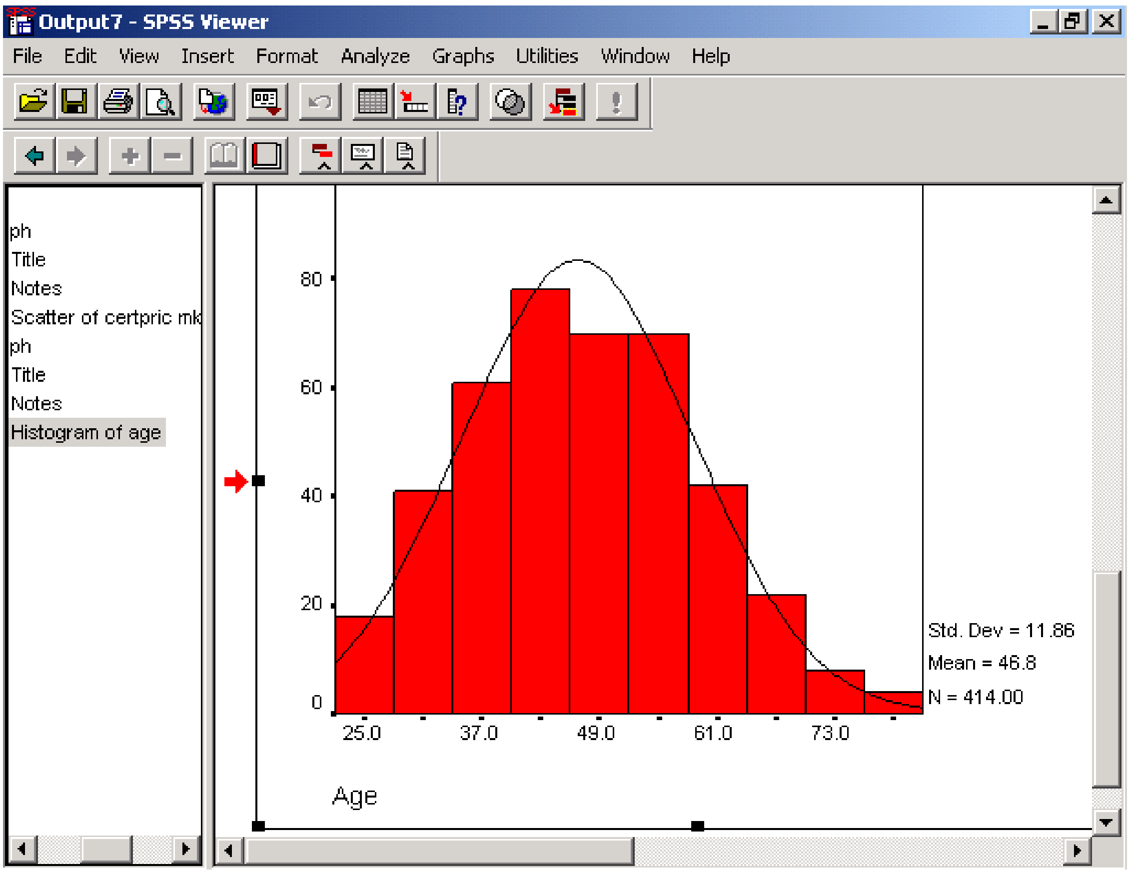


Figure 14. Histogram displayed in the Output Viewer

3. CONCLUDING COMMENTS

Researchers frequently collect large quantities of data, from surveys, experiments and other forms of observation. A statistical computing package provides a convenient means to store these data, and derive descriptive and inferential statistics. The Statistical Package for the Social Sciences (SPSS) is a widely used general-purpose survey analysis package, and hence a useful one to master. It is necessary to allow some learning time to become familiar with this package, and annual license fees can be a disincentive.

REFERENCES

Corston, R. and Colman, A. (2000), *A Crash Course in SPSS for Windows*, Blackwell, Oxford.

7. A Personal View of Statistical Packages for Linear Regression

Jerome Vanclay

This module presents some personal observations and views on statistical packages for forestry research, with particular attention on suitability for regression analysis, and important activity in development of systems models. There are many packages able to assist statistical analyses of data, designed for different purposes and audiences. This module illustrates some of the features that are particularly important when choosing a package for linear regression analyses. Among these is the graphical abilities of a package, especially the ease with which data (and residuals) can be plotted, because a model may be judged by the residuals rather than by the R-squared statistic. However, preferences for particular features and approaches are personal, so provided that candidates have the functionality one needs, a package may be chosen on the basis of personal preference. If you don't like the one you have, or find that it cannot handle your data or your analysis, look for another package. There are many well designed and tested statistical packages available, so it should always be possible to find one that suits your needs and budget.

1. INTRODUCTION

Two-variable and multivariate analysis are important steps in fitting relationships for use in systems models. Many statistical packages for use on personal computers are available with regression capabilities, and there is great variation in range of capabilities, ease of use and cost. This is a personal overview of statistical packages used by the author, emphasizing the utility of the package for fitting curves to data using linear regression. It is not a comprehensive review, and does not consider expensive packages such as SPSS, SAS, and S-plus. Instead, it looks mainly at the free or cheap packages that do not require an annual license fee. Some basic concepts in regression analysis are first introduced, and then a number of packages with regression capabilities are reviewed – specifically Excel, CurveExpert, GLIM, ARC and ViSta.

2. SOME BASIC CONCEPTS IN STATISTICS

Let's begin by re-examining the principles underlying curve fitting with regression analysis. Figure 1 is a graph with six data points. I'm going to ask you to draw the single straight line that best describes the trend evident in these points. Your line should be the free-hand equivalent of the least-squares approach used in regression analysis. Understanding the position of this line is fundamental to an understanding of regression analysis. Go ahead, and draw your line.

Examining the data graphically, and developing a mental image of the best fit are important first steps in model-building. As Forest Young (2001) puts it in poetry,

first, you see your data
for what they seem to be
then, you ask them for the truth
- are you what you seem to me?

you see with broad expanse
yet ask with narrowed power
you see and ask and see
and ask and see … and ask …

with brush you paint the possibilities
with pen you scribe the probabilities
for in pictures we find insight
while in numbers we find strength

Forrest W. Young, 2001

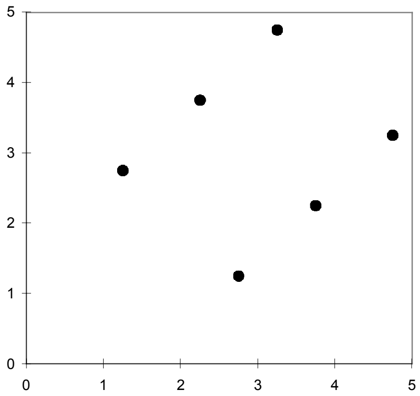


Figure 1. A scatter plot.

Now compare your line with those illustrated in Figure 2. Many people draw the first principal component (dotted line) rather than the least squares fit (solid line). They are minimizing the overall distance between the points and their line, but that's not what linear regression does. Linear regression assumes that the X-values (the numbers on the horizontal axis) are known with certainty, and that any errors are associated only with the Y-values. Thus linear regression minimizes the vertical distances between the points and the lines (Figure 3). In fact, it minimizes the square of those distances – that's why it's called the *least squares* fit, and why one of the measures of goodness-of-fit is the residual mean square (the average of the vertical lines squared). How did your line compare with those in Figure 2?

The vertical lines in Figure 3 are known as residuals. Examining the residuals is a useful way to judge the quality of a fit. And because the regression technique is based on the square of these residuals, you should take special note of large residuals, which may have a strong influence on the shape and position of the fitted line. Consult any standard text on regression analysis for a fuller explanation. The text by Cook and Weisberg (1994) is a helpful place to begin.

Many people judge the quality of a fit using a statistic known as R-squared. An R-squared of zero indicates that the fitted line is no better than a simple average, and an R-squared of one reveals a perfect fit. When R-squared is close to one, it may indicate a

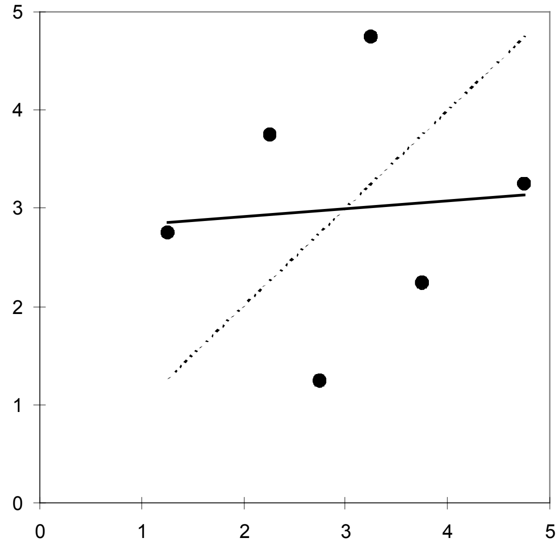


Figure 2. First principal component (dotted line) and least-squares fit (solid line) to the six data points (redrawn from Vanclay 1994, Figure 6.3)

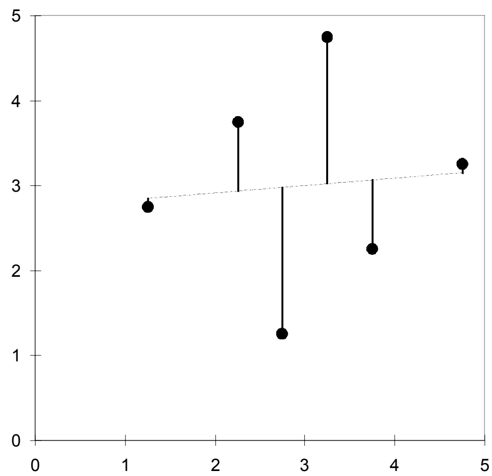


Figure 3. Regression analysis chooses the dotted line in such a way as to minimize the sum of the squares of the solid vertical lines

good fit, but it does not indicate if the fit is good enough.

Anscombe (1973) created four data sets that reveal some limitations of R-squared and emphasise the need to examine data graphically (Figure 4). The four data sets that he formulated have exactly the same linear regression estimates (Y = 3.0 + 0.5 X), exactly the same residual mean square (13.75) and exactly the same R-squared (0.667). However, despite these similar values, the quality of the fit varies greatly

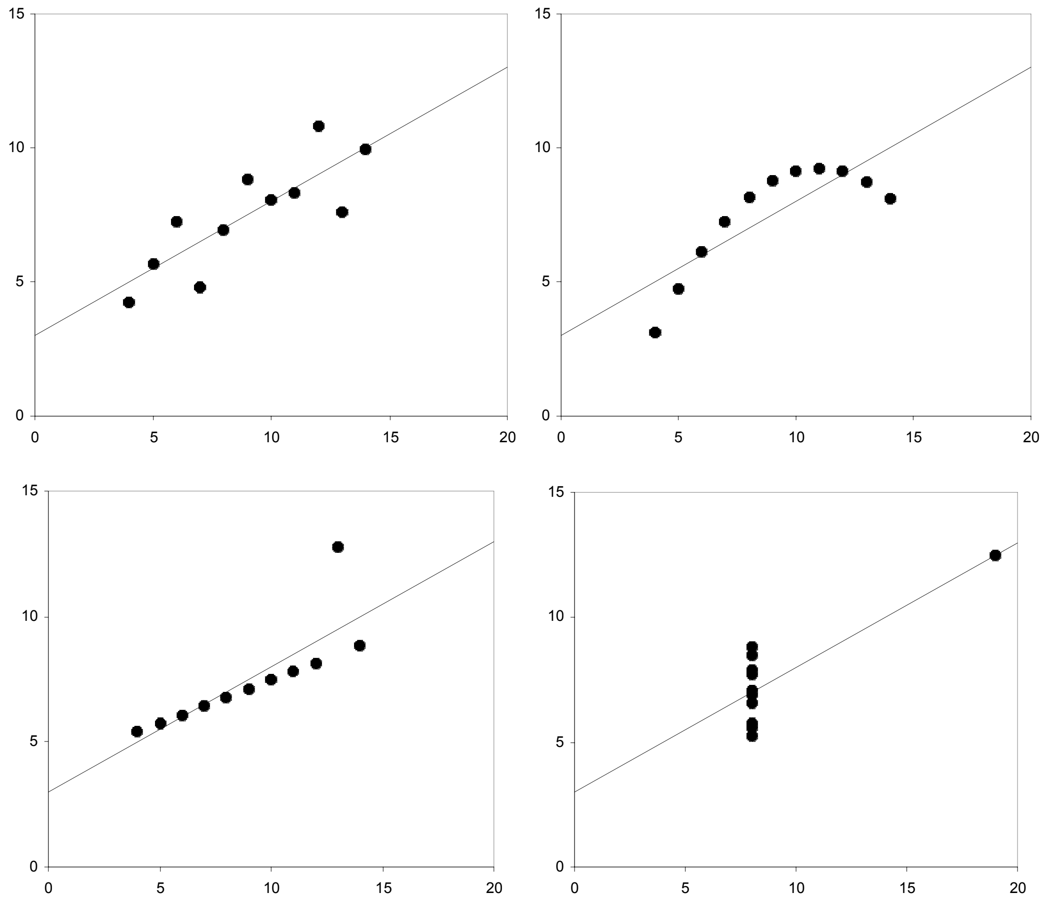


Figure 4. Four data sets constructed by Anscombe (1973) in which the R-squared is exactly the same (0.667), but in which the quality of the fit varies greatly

between the four data sets (Figure 4). They reveal 'pure error' (top left), an outlier (bottom left), use of the wrong model (a quadratic term may be needed; top right), and a case where the estimated trend relies entirely on a single point with high leverage (bottom right). Without further information, it is impossible to judge which of these models is suitable, but all except the first case warrant careful investigation.

A fuller discussion of this illustration was given by Anscombe (1973) and Weisberg (1985), but the importance of plotting both the data and the model remains obvious.

Because it is so important to plot the data and examine the points visually, I am going to consider only those statistical packages that have strong graphics capabilities. As Forrest Young (2001) says in his poem, it is '*… in pictures we find insight, while in numbers we find strength*'.

Five packages will be illustrated, and used to fit a simple linear regression (Y = a + b X) to a small set of data. The data are 10 pairs of numbers, with X = 1, 2, 3, …, 10 and $Y = X^2$. Fitting a straight line isn't particularly appropriate, but the R-squared is close to one (0.95). We contrast the insights we gain into this analysis as we examine the five packages – Excel, CurveExpert, GLIM, ARC and ViSta.

3. MICROSOFT EXCEL

MicroSoft Excel is part of the MicroSoft Office suite of programs, and is available on many computers. It's not my favourite package, but it is handy and sometimes may be the only statistical software available.

To perform regression analysis with Excel, click on 'Tools' and on 'Data Analysis'. Use the slider to find 'Regression', select and

click ‘OK’, and the computer screen should look like Figure 5. If the Tools menu doesn’t include ‘Data Analysis’, it may be necessary to select the add-in: Click on ‘Tools’ and on ‘Add-Ins’, and make sure that the ‘Analysis TookPak’ is selected. If you cannot find the ‘Add-Ins’, ask your computer support people to install the full version of Excel.

When performing a regression analysis with Excel, be sure to plot the raw data with the fitted model, and examine the residuals (Choose both the ‘Residual Plots’ and ‘Line Fit Plots’ in the regression dialogue box, Figure 5), and do not rely only on the R-squared and the F-test to judge the adequacy of a model.

Statistical procedures in Excel are not always reliable. McCullough and Wilson (1999) reported that the Statistical Reference Dataset of the American National Institute of Standards and Technology reveals problems with Excel’s univariate summary statistics, analysis of variance, linear regression, non-linear regression and random number generation. The problems appear to be present in Excel 4, Excel 95, Excel 97 and Excel 2000. Thus it may be advisable not to use Excel for complex or critical analyses.

4. CURVE EXPERT

CurveExpert has limited capabilities, fitting curves to (X, Y) pairs of data, but it is an easy way to fit an equation to data. CurveExpert automatically fits and compares 35 built-in regression models and up to 15 additional user-defined models comprising up to 19 parameters. These include both linear and non-linear regression as well as splines. There is no limit on the number of data points. Good documentation is available on the Help menu.

Data entry and analysis involves simply typing pairs of data (X, Y) into the built-in spreadsheet, and pressing control-f, and CurveExpert will establish the line of best fit, graph it, and display the corresponding equation (Figure 6). Graphs can be customized and copied onto the clipboard for use in other Windows applications.

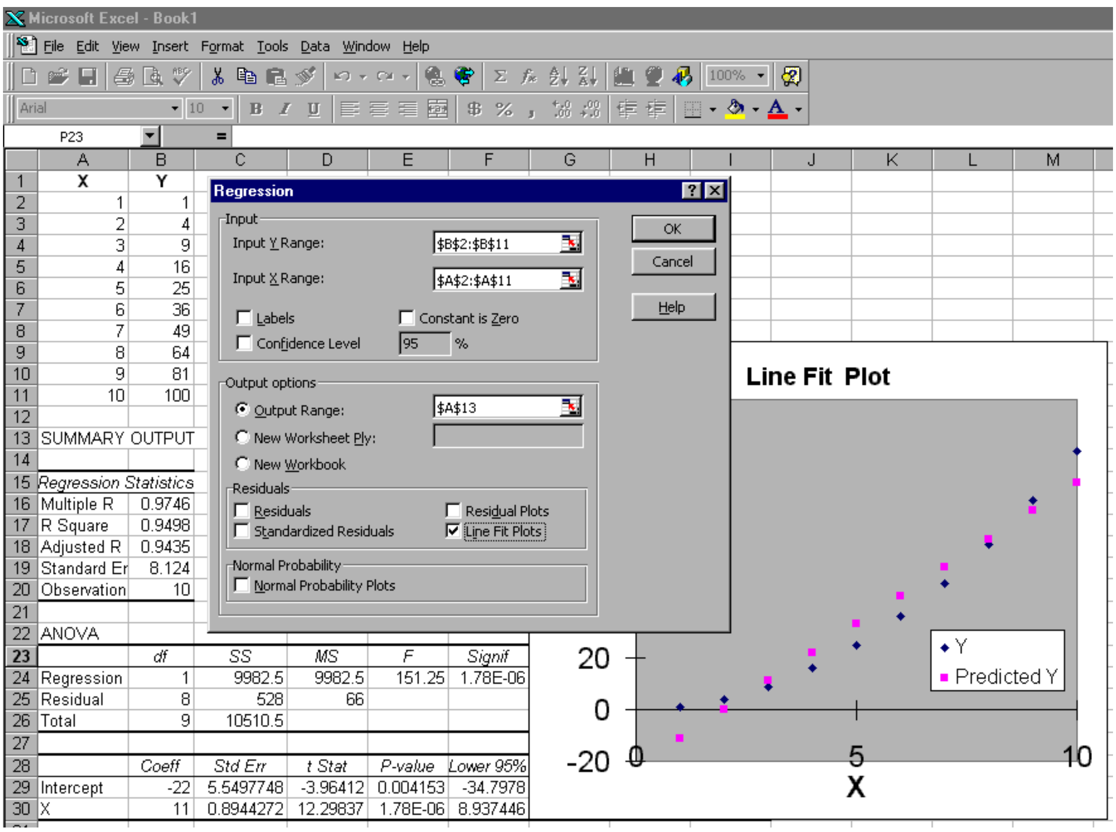


Figure 5. Image of a computer screen during a regression analysis with Excel

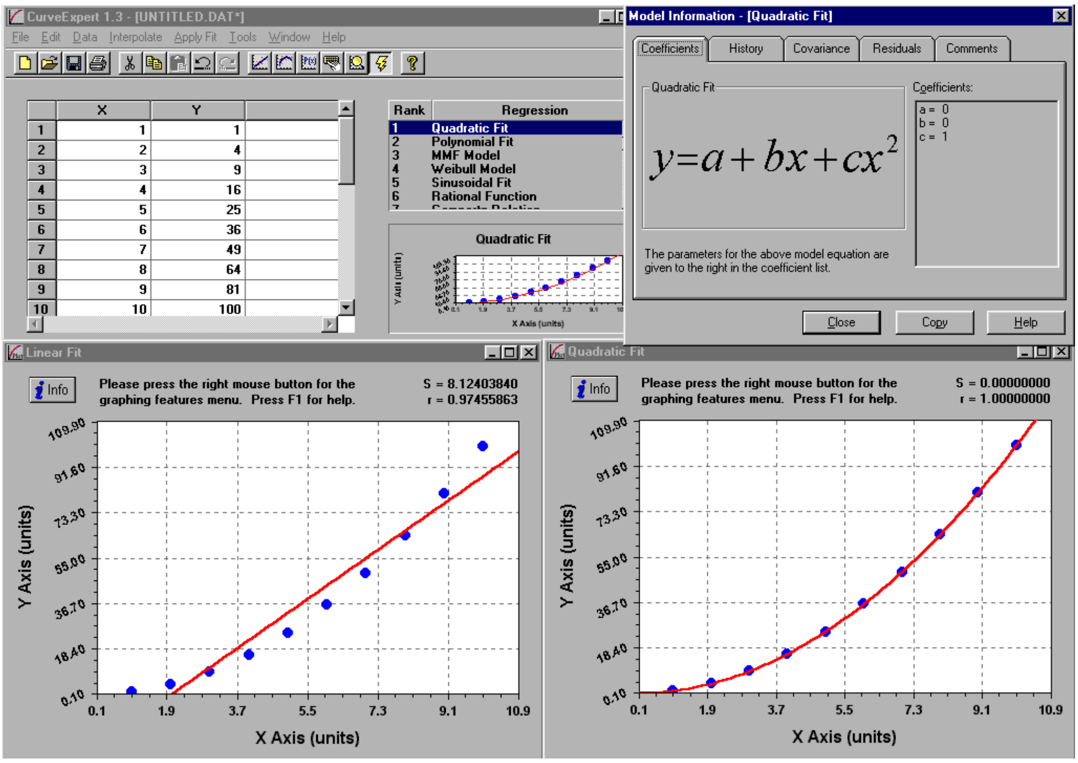


Figure 6. Image of computer screen while fitting a curve with CurveExpert

The major limitation of CurveExpert is that it can use only pairs of data. This is adequate if you want to build a simple yield table based only on stand age, but means that you cannot easily include site index or other explanatory variables.

CurveExpert is shareware developed by Daniel Hyams. It is available freely on the internet at http://www.ebicom.net/~dhyams/cvxpt.htm. Users are requested to pay a once-only registration fee of US$40.

5. GLIM

GLIM (Generalized Linear Interactive Modelling, release 4, Aitkin *et al*. 1989) is an old favourite, with which I do most of my analyses. It's compact – the whole system fits on a couple of diskettes. And it is powerful – many analyses can be completed with just a few commands (e.g. Figure 7, in which the data are simulated and analysis completed in one line). But it is not easy to use, so I rarely recommend it to others. There are no pull-down menus, and when GLIM is started, the user is greeted with a simple prompt, as it awaits a command. However, it is flexible and powerful, and is the only package that I can use to complete some analyses (e.g. Phylogenic regression, Grafen 1989). The model fitting illustrated below is achieved with the following commands:

```
$yvariate Y
$fit X
$display estimates
$calculate r=%yv-%fv
$plot r %fv
```

The commands (prefixed with a '$') can be abbreviated, often to one or two characters. The equation r=%yv-%fv calculates the residuals by taking the difference between the y-variables (%yv) and the fitted values (%fv). The graph is a plot of the residuals and the fitted values, and is a good way to judge the quality of a fit

GLIM is marketed by the Numerical Algorithms Group in Oxford, UK (NAG Ltd, infodesk@nag.co.uk). When I purchased my copy (several years ago), there was a once-only fee of about £100. Crawley (1993) published a helpful guide to the use of GLIM.

6. ARC

Arc (Cook and Weisberg 1999) is a revision of the program *R-code* (Cook and Weisberg 1994). *R-code* greatly influenced my approach to regression analysis, and *Arc* has become my favourite regression analysis package. It has many innovative tools to allow insightful graphical scrutiny of data.

I quickly found *Arc* intuitive and easy to use. It has pull-down menus, but the user may need to consult the text (Cook and Weisberg 1999) to get the most from the package. Figure 8 shows the analysis of the data used in Figure 6 in *Arc*, but the real strengths of *Arc* become apparent only with more complex sets of data.

ARC is copyrighted software, but may be distributed and used free of charge. It is available from the web at www.stat.umn.edu/arc.

7. VISTA

I've recently discovered ViSta, the Visual Statistics System (Young 2001), and am still learning it, so I'm not well qualified to present it. But I would like to attention to this package, because I think that it is a useful system and it has many innovative features. It encourages graphical examination of data, and has an innovative way to document what has been done, and to suggest what can be done with the data (Figure 9).

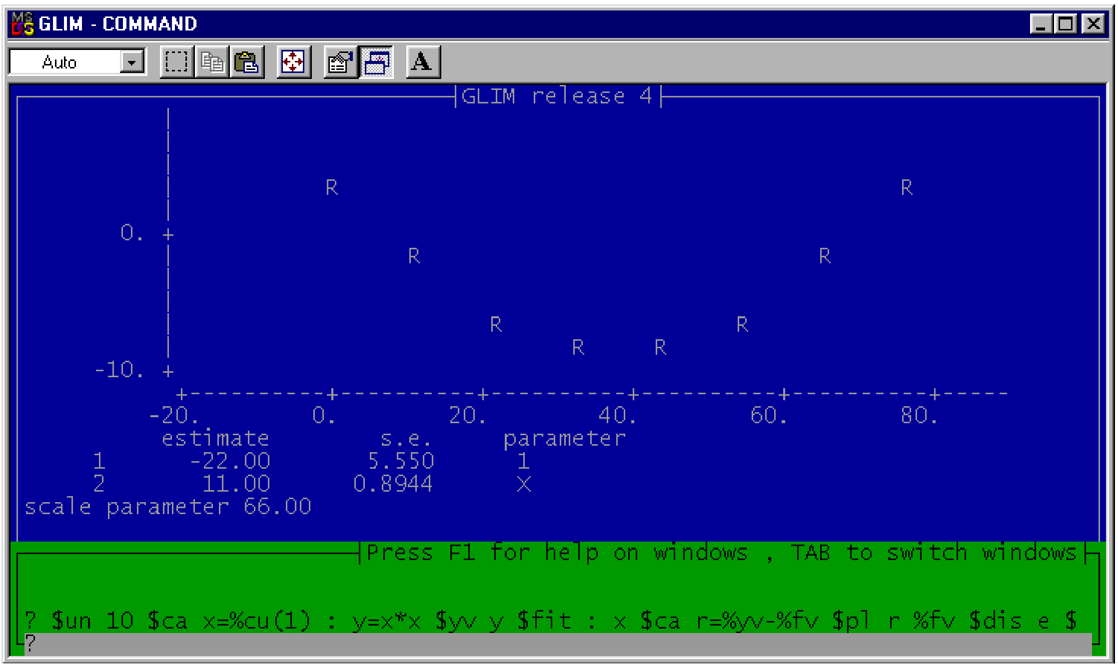


Figure 7. Image of computer screen while using the statistics package GLIM (release 4)

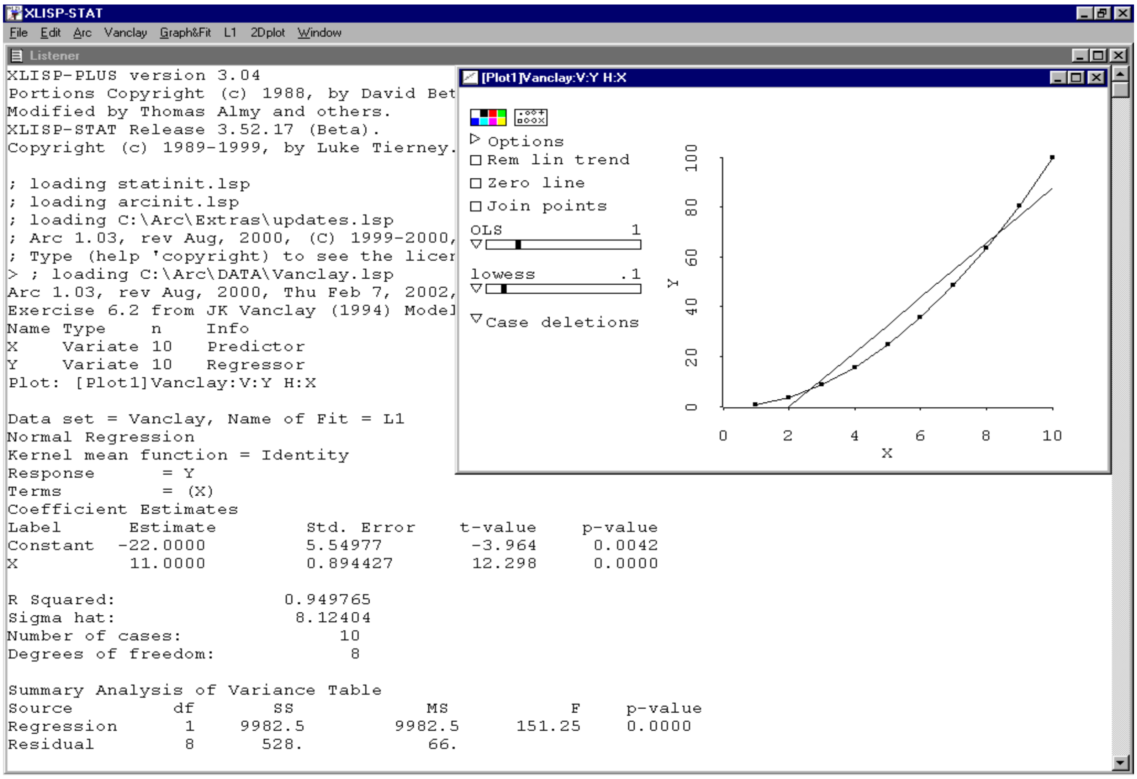


Figure 8. Image of computer screen while using the statistical package ARC

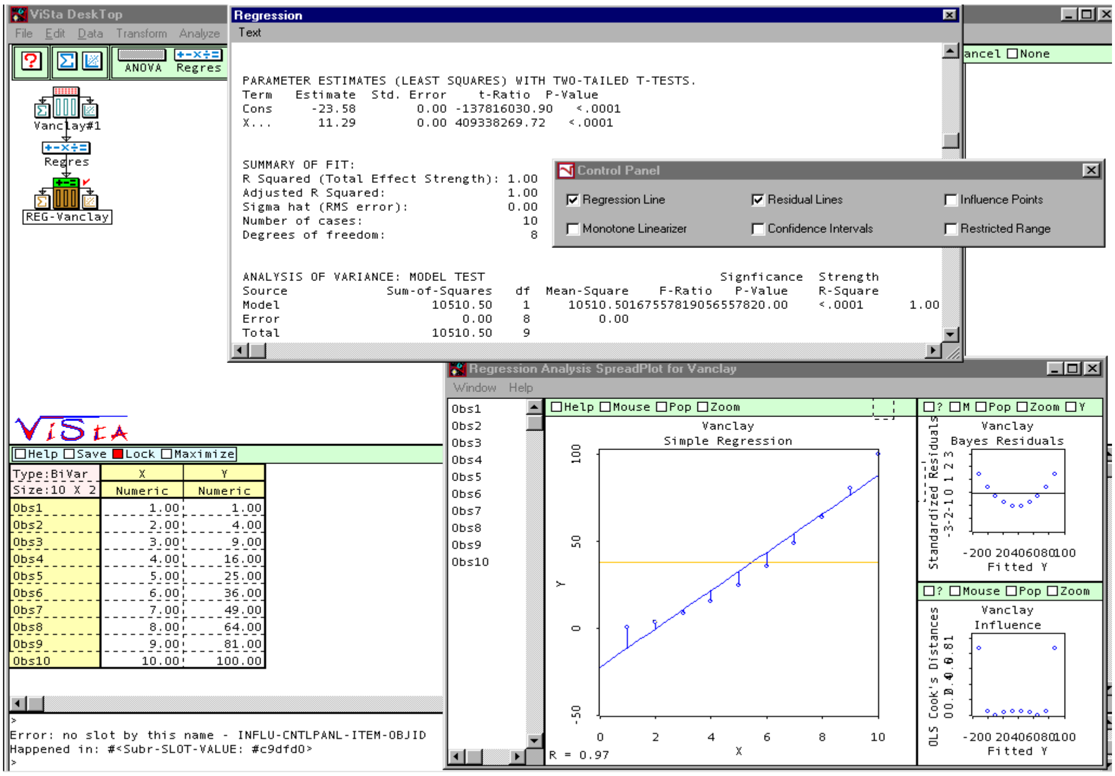


Figure 9. Image of computer screen while using Vista

Note: Note symbols to show what has been done in the analysis (top left), and automatic calculation and display of residuals and influences (bottom right) for every analysis.

ViSta is free, and can be downloaded from the web at www.visualstats.org.

8. CONCLUDING COMMENTS

I am not going to recommend any particular one of these packages – they all have strengths and weaknesses. I like the utility of Excel for data input, the ease of fitting equations CurveExpert, the power of GLIM, the graphics abilities of Arc, and the innovation in ViSta. But Excel has bugs, CurveExpert has limited capabilities, GLIM isn't easy for non-statisticians, Arc doesn't offer sufficient help (unless the text is also obtained), and ViSta takes time to learn. Preferences for particular features and approaches are personal, so provided that candidates have the functionality you need, choose a package that you like. If you don't like the one you have, or find that it cannot handle your data or your analysis, look for another package. There are many good packages out there, and I'm sure that you'll find one that suits your needs and budget.

REFERENCES

Aitkin, M., Anderson, D., Francis, B. and Hinde, J. (1989), *Statistical Modelling in GLIM*, Oxford Statistical Science Series 4, Oxford Science Publications, Clarendon Press, Oxford.

Anscombe, F.J. (1973), 'Graphs in statistical analysis', *American Statistician* 27(1): 17-21.

Cook, R.D. and Weisberg, S. (1994), *An Introduction to Regression Graphic*s, Wiley, New York.

Cook, R.D. and Weisberg, S. (1999), *Applied Regression Including Computing and Graphic*s, Wiley, New York.

Crawley, M.J. (1993), *GLIM for Ecologists*, Blackwell, Oxford.

Grafen, A. (1989), 'The phylogenetic regression', *Philosophical Transactions of the Royal Society, London, Series B*, 205, 581-98.

McCullough, B.D. and Wilson, B. (1999), 'On the accuracy of statistical procedures in Microsoft Excel 97', *Computational Statistics and Data Analysis*, 31(1): 27-37.

Vanclay, J.K. (1994), *Modelling Forest Growth and Yield: Applications to Mixed Tropical Forests*, CAB International, Wallingford.

Weisberg, S. (1985), *Applied Linear Regression*, 2nd. edn., Wiley, New York.

Young, F.W. (2001), How to use ViSta: the Visual Statistics System, The Visual Statistics Project, Psychometric Laboratory, University of North Carolina.

8. Simulation Philosophy and Methods

Steve Harrison

In general, to 'simulate' means to mimic or capture the essence of something, without attaining reality. For management-oriented applications, the something is an identified *system* under the control of management – for example of a *bioeconomic* system – and the essence is captured by way of a symbolic or algebraic *model*. Simulation consists of the steps of developing the model to represent a *real system*, and then performing *experiments* using this model to predict how the real system would behave under a range of *management policies*. The objective of simulation may be to increase understanding of the behaviour of the system, or to compare various policies for management of the system. While many quantitative techniques take a well-recognized form, simulation differs in its great flexibility, variety of applications and variations in form. These features, while highly valuable for modelling complex systems, make this a difficult methodology to explain and to comprehend. In fact, simulation has been described as 'more art than science'. Proficiency with this technique cannot be gained easily in the classroom, but rather requires considerable practical experience. This module presents the basic concepts of simulation and the steps fundamental to simulation studies. The module first defines the nature of simulation and the philosophy behind this technique. Modelling concepts and elements of models are then discussed. The typical steps when using simulation are next outlined. A simple example of developing and applying a model is presented to aid discussion. Some comments are made about validation of simulation models and about design of simulation experiments.

1. WHAT IS SIMULATION?

Simulation consists of a number of disparate concepts and techniques, with different terminologies adopted by different disciplines. Hence it is useful to establish the terminology which will be adopted in this discussion of the application of simulation to bioeconomic systems.

Proponents of simulation usually subscribe to what is called the *systems approach*. According to Shannon (1975), a system is 'a group of objects united by some form of interaction or interdependence to perform a specified function'. In other words, any system consists of a number of *interrelated* and *interacting* parts; further, these parts should not be studied in isolation but rather in the context of the overall system and its complex interdependencies. The whole is more than just the sum of the parts. Any change to one part of the system may cause unexpected changes elsewhere. Scientists, engineers, managers and so on have increasingly realized that it is necessary to take this *holistic view* of the systems they design and operate.

As the body of scientific knowledge has increased, there has been a tendency for greater specialization of research, with a loss in overall perspective and loss of communication between researchers in different disciplines. This 'spread of specialized deafness' has led to the study of more narrowly defined systems. This is not to suggest that *reductionist* research into narrowly defined systems does not play an important role in advancement of knowledge. However, from a management point of view, the system of interest is usually the level of aggregation at which planning and control decisions are made, which is often an overall business firm or a particular project being undertaken by a firm.

Various terminologies have been adopted in presentations of simulation methodology. In particular, the terms *systems analysis*, *systems research* and simulation have been used interchangeably. The term 'systems research' is typically applied to describe all the steps in the study of an organized system. 'Systems analysis' was originally used in this broad context, but has recently become applied more often to just one step in systems research, viz. that of identifying

the boundaries, elements and interrelationships of a system prior to modelling it. The term 'simulation' is sometimes applied to the overall procedure of modelling and experimentation, and sometimes to the experimentation stage only. In this module the term 'simulation' will be used in the broader context, as a synonym for systems research.

Under the systems approach, a model is developed which represents as closely as possible the essential structure and performance of the real system, in terms of the specific behavioural features the researcher wishes to examine. This model once constructed and tested for reliability is then used to simulate or mimic how the actual system would behave under particular circumstances, by conducting experiments on a computer using the model. In these experiments, measures of system performance are generated when levels decision variables are set at a number of levels.

Simulation tends to be used where a solvable model is not available (i.e. would not represent the complex structure of the real system adequately). The simulation experiments may be likened to observing how the real system would perform if particular management policies were to be adopted, except that the real system is not interfered with, real resources are not used (apart from computing resources), and time is greatly compressed. Computer simulation experiments can thus provide a great deal of information about how an actual system would behave, under a host of different policies and environments, and this may provide a greatly improved understanding of the system and how to manage it.[1]

Practitioners of the systems approach need to have a good understanding of the overall system, and the ability and willingness to consult with experts on various components of the system. In fact, systems research is often conducted by groups or *multidisciplinary teams* rather than individuals, since these can take on board a broader range of expertise on various aspects of the system.[2]

2. THE NATURE OF MODELS

The term 'model' has wide everyday use. Physical or *iconic* models are typically scaled down versions of real systems. Some physical models are not true-to-scale replicas, but still convey information about a system. In a broad sense, any map is a model of a territory, and any timetable is a model of the operation of a transport system. Hence, any form of model building could be thought of as a form of simulation. However, the term has become associated with a particular approach to decision support. From a forestry research perspective, the models we are interested in are usually algebraic models of bioeconomic systems.

Nowadays, people are familiar with building algebraic models, through the widespread use of spreadsheets. Any spreadsheet – including one to derive the net present value (NPV) of a project – is in effect an algebraic model. The spreadsheet allows experimentation with the model, such as asking 'what if' type questions through varying a model input and checking the performance estimates.

[1] A particular type of simulation is that of various forms of *games*. There is a long history of war games, both using model ships or soldiers on a table, or using actual troops but dummy ammunition. In recent years, enormous resources have been put into development of computer games, and some of the computer graphics in these have become incredibly good simulations of real world scenes and activities. Another gaming application is that of *management games* for training in business, and macroeconomic games for training in economics. For these games to be considered genuine by participants, a credible model must be present which mimics the behaviour of the underlying business or economic system for which the training is being conducted.

[2] The alternative is to have studies conducted by *transdisciplinary* individuals, i.e. researchers with multiple skills across a range of disciplines. The obvious difficulty here is for individuals to have sufficient depth across a range of disciplines to adopt a systems approach.

The need for a model

Sometimes it is possible to gain experience about the operation of a system by imposing various managements upon the system itself. Agricultural researchers have traditionally used research stations and on-farm trials to evaluate pasture species, cropping systems, irrigation practices and so on. Many business managers would admit to learning how to operate their business in the 'school of hard knocks'. However, learning by one's mistakes can be a slow, costly and unpredictable way of gaining experience. Conducting experiments on a model rather than the real system may be preferable on a number of grounds:

(i) the real system might not yet exist. For example, a company may wish to expand or diversify its activities, perhaps investing a large amount of capital. It would be useful to have an idea of potential outcomes of this venture, and their likelihoods, before committing capital to it.

(ii) performing experiments on the real system may be unacceptably time consuming or expensive. This would be the case, for example, if management wished to know the cash flows resulting over the next 10 years if forestry plantations were to be located in either of two alternative sites or either of two different species or mixtures were adopted. Once a model is constructed, stand growth and economic performance can be generated and compared on the computer for the two sites or species, in a matter of seconds of real time.

(iii) performing experiments may be inimical to the real system. For example, suppose the directors of a forestry company wished to know whether the company could survive under various levels of borrowed finance. It would certainly be a risky proposition for the company to take out a very large loan and see if the debt could be serviced. However, no harm would be done to the company if a model were used to experiment with different gearing ratios, and to determine what debt level leads to an acceptably low probability of 'simular' bankrupcy or takeover.

(iv) The exercise of developing a model can be in itself a valuable learning experience about a system. A model provides a framework for systematically organizing existing information, and for assembling the (often subjective) knowledge of experts. The explicit nature of models makes ideas and assumptions more transparent, and places these under the scrutiny of more people. This in part explains why models tend to be built through a series of prototypes, evolving over time as questionable assumptions and gaps in knowledge are exposed, more questions are asked and more information is assembled.

Types of models

Typically, a bioeconomic system is modeled as a set of equations or relationships between variables, although often the algebra of the model is implied rather than made explicit. Various classifications of symbolic models have been advanced. For example, models may be grouped as

(i) biological, economic or bioeconomic
(ii) discrete or continuous
(iii) single-period or multiperiod (or static or dynamic)
(iv) deterministic or stochastic.

A great deal of modelling has been carried out in the natural sciences, including forestry. For example, models have been developed of physiological processes within trees, and of the growth of individual trees and stands of trees. A *biological* model of a forestry stand could take into account soil moisture, nutrient uptake, biomass growth, partitioning of nutrients between plant parts, and so on. The time step would normally be a period less than a year, perhaps as short as a day. The performance measure might be total biomass or bole biomass produced. A *financial* or *economic* model of a forestry stand would typically take account of expenditure and of product sales, on a yearly basis, and seek to estimate the net present value of the forestry investment.[3] Components of a biological and an economic model may be combined to

[3] Some differences exist between financial and economic models, as discussed in Modules 15 and 16.

produce a *bioeconomic* model, which predicts profitability from the forestry investment, but taking into account the production processes rather than simply taking final yield estimates.

Models in which time changes in a *continuous* fashion are used extensively in mathematics and engineering, frequently taking the form of sets of simultaneous differential equations. Most management-oriented models are of *discrete* form, with variables taking time steps of say one week or one year.

Static models make no allowance for changes in values of variables over time. They are thus not suited to examining long-term production processes such as forestry systems. Multiperiod or dynamic models include specific representation of changes in levels of variables over time, such as growth, partition and decay of biological components and accumulation of funds or assets.

In a *deterministic* model, all variables take single point or best-estimate values, and all relationships between variables are assumed to be known with certainty. Stochastic models on the other hand incorporate probability distributions for one or more variables, or uncertain error terms in relationships between variables. Thus a forestry bioeconomic model could take account of variable rainfall, crop hazards (e.g. pests, wildfires) and timber prices. With a stochastic model, the measure of performance could include both expected payoff and a measure of uncertainty. It is generally considered that a refinement in modelling introduces a need to incorporate uncertainty. That is, as the modeling is carried out in increasing detail, it becomes necessary to introduce of stochastic variables.

3. ELEMENTS OF MODELS

It is useful to introduce some further terminology to assist in the discussion of models. A convenient classification (drawing on Naylor *et al.* 1966) is to divide models into the following elements:

(i) *components* or building blocks. For example, in a forestry model these could include trees, soil and weather, or at a more aggregate level nursery, plantation stands, workforce, and machinery and equipment. By definition, management is outside but exerting planning and control functions over the system (and hence not a component of the model).

(ii) *variables*, the levels of which may be determined outside the system (exogenous) or within it (endogenous), or which may describe the state of the system (status variables). Exogenous variables may be under the control of management (*decision variables*) or not controllable (*environmental variables*). For a forestry operation, silvicultural treatments and associated labour and material inputs typically are controllable, while prices of labour and other inputs and of products generally are non-controllable. In simulation experiments, exogenous variables form the inputs which drive the model. Those variables under the control of management form the decision or policy instruments, the levels of which are adjusted in simulation experiments. Non-controllable exogenous variables may be modelled as either fixed values (or predetermined time series) or as probability distributions from which random values are generated. The desirability of particular management policies is assessed in terms of one or more model outputs or levels of endogenous *performance* variables. In a forestry model, these might include, mean annual increment (MAI), and NPV from a plantation. *Status* variables record the state of the system at each period in time. Examples for a plantation include mean tree height and girth at various ages, and also financial variables such as cash and debts. The growth rate of trees will depend on growth in previous time periods, sometimes referred to as a 'feedback loop'.

(iii) *functional relationships*. These indicate how variables are interrelated, e.g. by linear or non-linear equations, and with or without time lags. Included are *identities*, which are true by definition (e.g. value = price x quantity sold) and *operating characteristics* which have to be estimated statistically or subjectively. The latter might include the relationship between tree height and age. Functional relationships give the system its unique behaviour, and needless to say the reliability of any systems model depends

vitally on how accurately the relationships are identified and estimated.

(iv) *parameters*: these are the coefficients of the operating characteristics, values of which can only by estimated to within given confidence levels.

Any systems model may be summarized by the following symbolic relationship:

$$Z = f(X,Y,S,A)$$

where Z is a set of performance variables
X is a set of policy variables
Y is a set of environmental variables
S is a set of initial levels or status variables (including initial resource levels)
A is a set of parameter values, and
f signifies that a functional relationship exists between the variables in the various sets (i.e. f represents the model).

4. THE STEPS IN A SIMULATION STUDY

Carrying out a simulation study is like carrying out any other quantitative study (though sometimes a bit more complicated). The so-called *scientific method* is employed, which means a series of steps are performed in a logical fashion to achieve the overall task. The terminology for simulation steps varies between experts, but the following is a workable classification.

(i) identification of the problem
(ii) analysis of the system
(iii) synthesis of the model
(iv) programming the model to a computer
(v) testing the model
(vi) experimentation with the model
(vii) interpretation of results, and reporting to the relevant authority.

These steps are performed in the sequence listed, except that there is usually some cycling between them, e.g. testing the model may reveal the need to modify the structure then revise the computer program. Each of the individual steps will now be discussed briefly.

Identification of the problem

It is most important to identify clearly the study objectives in terms of the research or managerial problem which is to be examined. The nature of the model to be developed will depend on the problem to be analysed. Is the objective to understand the system or to prescribe management policies? If the latter, who is responsible for the system, and what are their goals, and what is wrong with present policies?

Analysis of the system

Once the problem is identified, the ‘systems analysis’ stage can be performed, in which the boundaries of the system, the relevant variables and their interrelationships are identified. This may involve drawing various charts or diagrams of the system.

Systems synthesis

The next step, of ‘systems synthesis’, consists of expressing the relationships between variables in symbolic form, and estimating the parameters of these relationships. Because of the high degree of flexibility possible, it is difficult to lay down rules for this major step, though some guidelines can be given. Where possible, statistical techniques should be used to estimate relationships between variables. Distributions or random variables can be obtained by testing the goodness-of-fit of alternative probability models using say the chi-squared test. Where historical data are scarce or are not considered relevant to future behaviour of the system, subjective estimation by people regarded as having *expert knowledge* about the system may be preferable. It is usually recommended to start with a relatively simple model, and gradually extend and refine it. To the extent that sub-systems are sufficiently independent, the model should be constructed in the form of a number of relatively self-contained *modules*, allowing these to be programmed and tested separately. Existing models of similar systems should be examined for relevance, since it may be possible to obtain ideas or even adapt modules from them.

Programming to a computer

Once a prototype version of the model is constructed, computer programming can commence, using a spreadsheet package or – if the model is to complex for this –

using a computer programming language such as Visual Basic, FORTRAN or C, or a specialist simulation modelling language such as Simile.

Testing the model

Having constructed a working model, it is necessary to test whether this model is an adequate representation of the real system. There is far from general agreement on how systems models should be tested, but a workable approach is to divide testing into *verification*, *validation* and *sensitivity analysis*. Verification is the process of testing whether the model takes its intended structure, i.e. whether the model is free of logical errors and whether the computer program performs as intended. Validation examines the broader question of whether the intended structure truely represents the real system. Validation efforts often lead to further refinement of the model.

Once a model has been validated as far as practicable, the effect of remaining errors on parameter estimates may be assessed through sensitivity analysis. If the purpose of the model is to identify optimal management policies, errors in estimates of performance due to inaccurate parameter values may not be of concern unless they lead to identification of inferior policies as optimal. That is, we are not concerned with the predictive ability of the model in absolute terms so much as the model's ability to correctly rank alternative management policies. For this reason, it is desirable to include a sensitivity analysis with respect to optimal values of decision variables. Sensitivity analysis usually involves adjusting parameter values by small amounts, and calculating various sensitivity criteria. Sensitivity may be expressed quantitatively in terms of 'elasticity' of performance with respect to parameter levels, or elasticity of optimal management policies with respect to parameter values. High sensitivities (elasticities) give cause for concern about the reliability of a model.

Performing experiments

Once sufficient confidence has been gained in a model, a variety of simulation experiments may be conducted. Various designs can be used for these experiments, as discussed later in the module. Many of the principles of experimental design as applied in the physical sciences are relevant to simulation experiments. Management policies may be specified in terms of a single policy variable or a number of variables. In practice, optimal (or at least desirable) levels of a number of policy variables often have to be determined simultaneously. For example, a forest manager may wish to know what fertilizer strategy, pruning and thinning regime and harvest schedule maximizes returns from a plantation. In the language of experimental design, the decision variables which are identified are known as *experimental factors*, and any combination of levels of these factors (i.e. any management policy) is known as a *treatment*. Measures of performance of the system are known as *response* variables. A computer run in which a number of treatments are evaluated is known as a simulation *experiment*. If random variability is built into the model (i.e. if the model is stochastic), then it is necessary to evaluate each treatment or policy under a number of different environments, i.e. to include *replication* treatments in the simulation experiment.

Experiments conducted on a computer also have important differences from real-world experiments. Three main sources of difference arise:

(i) compression of time. Because of the speed of computing and the low cost of computer time, it is usually possible to include a larger number of treatments and a greater degree of replication.

(ii) sequential processing. Traditionally, each of the treatments in a real-system experiment is evaluated at the same time. For example, in a plantation fertilizer experiment, the complete experimental design is decided, with all plots planted at the same time, and fertilizer applied to each on the same day, and girth and height meansurements made at the same times. On the other hand, because a computer is a sequential processor, treatments are evaluated sequentially in a computer simulation experiment. This means that the performance level for the first treatment is

known before the second treatment is evaluated, and the performance for the first and second treatments are known before evaluating the third, and so on. Sequential processing opens the opportunity to use information gained from earlier treatments to direct factor levels in later treatments, within the same experiment. 'Optimum-seeking' experimental designs are discussed later in the module.

(iii) control over experimental variability. In a computer simulation experiment, the variability in the 'environment' is under the control of the researcher. A *random number generator* is used to produce numbers between zero and one, and these are transformed to random observations from the distributions specified for the random variables. If the random number generator is given the same *seed* for each treatment then the treatments are evaluated under the same sequences of random numbers, i.e. under the same environments. This reduces random variability in response levels between treatments relative to independent seeding (i.e. not re-seeding the random number generator). The result is greater power to detect differences between treatments for a given sample size (number of replicates).

Analysis and interpretation of computer output

Simulation experiments often generate reams of computer printout, and this output must be distilled and interpreted to a form useable by managers in a decision-support role. Since experiments conducted using stochastic simulation models do not provide exact results, estimated performance levels should be thought of in a confidence interval context.

5. EXAMPLE OF AN INVENTORY SIMULATION

Some of the concepts introduced above will now be demonstrated with reference to a simple inventory model.

Example 1

A seedling nursery faces uncertain demand for Mahogany seedlings. Recent experience suggests that demand can be approximated by a uniform distribution with a range of 5,000 to 10,000 seedlings in autumn, and 3,000 to 5,000 seedlings in each of winter, spring and summer. Seedling production cost is \$300/1000 seedlings and the sale price is 60c/seedling. Seedlings are grown to be ready for planting in autumn, but may be retained until the following summer (after which they must be discarded), the holding cost being \$120/1000 seedlings for each quarter.

Simulate quarterly marketing of Mahogany seedlings over four years, for a policy of growing 20,000 seedlings per year.

The spreadsheet model

The simulation model may be developed as an Excel spreadsheet, as in Table 1. The level of the decision variable (autumn inventory level) and parameter values (prices, costs, demand limits) are listed at the top of this sheet. For convenience, the year runs from autumn (the preferred planting time) through to summer. Each quarter is represented by a row in the simulation, and an estimate of net revenue is obtained for each quarter. Quarterly demand is obtained by the formula

lower demand level + random number x (demand range)

where the random number is from a uniform distribution in the range zero to one, obtained by the spreadsheet function RAND() (see Appendix A). Quarterly sales are obtained as

MIN(inventory level, demand level)

and any quantity left unsold is transferred to inventory in the following quarter, except that no inventory is carried forward from summer. Quarterly sales revenue is obtained as sales quantity multiplied by price net of production cost. Quarterly net revenue (or loss) is obtained as sales revenue less holding costs. Average annual revenue is obtained by summing quarterly net revenues over the 16 quarters and dividing by four.

In this example, the mean annual net revenue is about \$3,980. Various inventory production policies (treatments) could be

compared, by changing the 'Annual production' level, and re-running the simulation to determine the mean annual net revenue. In this way, the optimal production level could be determined. Because of the random element in demand, it would be preferable to increase the number of years over which sales are simulated, relative to the four years in this illustration. While While separate random numbers would be generated for each inventory policy simulated, leading of confounding of effects of inventory policy and demand environment, this can be overcome by 'freezing' the sequence of random numbers.

Table 1. Spreadsheet model for inventory simulation example

Parameters of the simulation model

Production cost ($/1000)	300
Sale price ($/seedling)	0.6
Holding cost ($/1000 seedlings/quarter)	80
Autumn demand - lower limit (1000)	5
Autumn demand - upper limit (1000)	10
Off-season demand - lower limit (1000)	3
Off-season demand - upper limit (1000)	5

Simulation of quarterly sales and net revenue

Year	Quarter	Inventory (1000)	Random number	Demand (1000)	Sales (1000)	Holding cost ($)	Residue (1000)	Sales rev. ($)	Net rev ($)
1	Autumn	20.00	0.8501	9.25	9.25	0	10.75	2775	2775
1	Winter	10.75	0.4579	3.92	3.92	859.94	6.83	1175	315
1	Spring	6.83	0.6244	4.25	4.25	546.67	2.58	1275	728
1	Summer	2.58	0.3745	3.75	2.58	206.77	0.00	775	569
2	Autumn	20.00	0.3107	6.55	6.55	0	13.45	1966	1966
2	Winter	13.45	0.5514	4.10	4.10	1075.72	9.34	1231	155
2	Spring	9.34	0.5205	4.04	4.04	747.50	5.30	1212	465
2	Summer	5.30	0.4510	3.90	3.90	424.21	1.40	1171	746
3	Autumn	20.00	0.2979	6.49	6.49	0	13.51	1947	1947
3	Winter	13.51	0.6859	4.37	4.37	1080.85	9.14	1312	231
3	Spring	9.14	0.6435	4.29	4.29	731.10	4.85	1286	555
3	Summer	4.85	0.8500	4.70	4.70	388.14	0.15	1410	1022
4	Autumn	20.00	0.7927	8.96	8.96	0	11.04	2689	2689
4	Winter	11.04	0.8044	4.61	4.61	882.90	6.43	1383	500
4	Spring	6.43	0.7076	4.42	4.42	514.19	2.01	1325	810
4	Summer	2.01	0.3933	3.79	2.01	160.97	0.00	604	443
Mean annual net revenue									3979

6. VALIDATION OF SIMULATION MODELS

As noted above, validation is a critical task in development of a simulation model. While validation may be applied to the assumptions of the model, and to the various sub-models, in practice tests are usually concerned with the reasonableness of outputs or predictions of the overall model. In this regard, it is often recommended that statistical tests be used to compare output of the model with that of the real system, where both have been generated under the same management policies and environments. These tests examine hypotheses of the general form

H_0: outputs of the real system conform to those of the model
H_1: outputs of the real system differ from those of the model.

The output of complex (dynamic, stochastic) simulation models typically consist of time series for performance, which can include various statistical contributions (e.g. trend, seasonal effects). Since the two series of outputs may conform with respect to some properties and differ with respect to others, a comparison needs to be made of the various parameters (e.g. means, variances, autocorrelations and trends), and of overall distributions. On the face of it, these tests should allow a thorough check of all the statistical properties of the two output series. However, in practice a number of problems arise with statistical validation.

Often, little output is available from the real system with which to compare output from a model. All available data tend to be used in model construction, yet it is desirable for the data for testing purposes to be independent of that used when constructing the model.

Important differences arise between traditional applications of hypothesis tests and their role in validation of systems models. The motive behind traditional statistical testing is usually to demonstrate that a particular null hypothesis is false; the null hypothesis is simply set up as a 'straw man'. For example, when comparing two sample means the null hypothesis may state that the two underlying population means are equal, while the alternative hypothesis may state that they are unequal, i.e.

H_0: $\mu_1 = \mu_2$ (mean predicted output = mean real system output)

H_1: $\mu_1 \neq \mu_2$

and it may be possible to 'prove' that H_0 is false in the sense of demonstrating that the probability of μ_1 equalling μ_2 is negligible. When applying a validation test, the motive is usually to accredit the model, i.e. prove H_0 is true. Comparison of the t-statistic with critical values provides an indication of the probability that H_0 is false. If this probability is low, the model can be declared invalid. But if the probability is high (i.e. above the chosen significance level) then all that may be concluded is that the model has not been demonstrated to be invalid, beyond reasonable doubt. To state that the model must therefore be valid is at best a tenuous statistical inference.

Traditional hypothesis testing is designed to limit the frequency of type 1 errors. The cost of a type 1 error in terms of declaring a valid model as invalid may not be great; unnecessary marginal refinements may be made to the model. On the other hand, the cost of accepting as valid an invalid model may be substantial, both in terms of loss of credibility of the modeller if his or her creation is later found to be misleading, and in dollar terms for users from making incorrect management decisions based on information generated by the model. For a given sample size, the lower the significance level the greater the incidence of type 2 errors (accepting as valid models which are in fact invalid). In other words, to make the test procedure more stringent in terms of ability to reject invalid models, it is necessary to increase the significance level, say to 20%.

A further problem with statistical validation is that the assumptions underlying the tests may not be appropriate. For example, consider the t-test on mean model and real-system outputs. A paired-sample *t*-test is likely to be more appropriate than a test of independent samples, when the outputs have been generated under the same environments and managements. This paired-sample test assumes that successive differences are independent; in reality output series are often positively autocorrelated over time, with the result that the variance of differences between model and real-system outputs may be seriously underestimated. A modified form of *t*-test may be used which takes account of autocorrelations between paired differences for various time lags. Even this test requires the (relatively weak) assumption of second-order stationarity in differences, i.e. the correlation between differences depends only on the number of time periods between them, and is constant over time.

As illustrated by the above discussion, the application of traditional statistical tests to model validation is not a simple matter. It is no wonder that most model testing has been subjective in nature, such as graphical comparison of time series of model and real-system output, and assessment of the

reasonableness of model output by people judged as experts with respect to the system.

Another reason for concern about statistical validation procedures is that they are being applied to a moving target. Simulation models typically evolve over time. A need for a model is identified, and a prototype model is developed and implemented to fulfil this need. Often, limitations of the model are recognized, and the model is further refined. As well as applying the model for its initial design purposes, other uses are often discovered, and extensions of adaptions of the model are undertaken to meet these other purposes. As a consequence, it is generally conceded that there is no clear finish to the testing of a systems model. Rather, confidence is gradually built up in a model over time as tests are performed, new information is obtained, and new versions of the model are produced.

7. THE DESIGN OF SIMULATION EXPERIMENTS

Suppose a simulation experiment involves two decision variables or experimental factors. Combining each level of one factor with each level of the other leads to *full factorial* design. The full factorial design is conceptually simple and well suited for simulation experiments in which there are only two or three factors, each taking only a few levels. In general, if there are *m* factors, each taking *n* levels, then there are *mn* distinct treatments. If *m* is more than two or three, the number of treatments can increase dramatically, and the experiment may become unmanageably large, even when the number of replicates is small. Alternatives to the full factorial design, such as the *fractional factorial* and *central composite* designs, allow for some reduction in treatment numbers for given m and n.

An alternative to designs which are specified in advance of the simulation experiment, is to include a set of rules in the computer program to select factor levels during the experiment on the basis of information gained from earlier treatments. The simplest of these designs is *steepest ascent* (or steepest descent). This is an iterative procedure in which the slope of the response surface is estimated with respect to each experimental factor, then all factor levels are adjusted so as to achieve the most rapid increase in performance. More sophisticated *hill climbing designs* may be adopted. These will involve programming the procedure for allocation of treatments on the basis of progressive performance during the experiment, or using sub-routines which have been developed by others for this purpose, which are available from a variety of sources.

Yet another alternative is *random search*, in which treatments are chosen simply by selecting the level of each factor at random. It is necessary to place upper and lower bounds on factor levels in order to confine the search area, based on prior knowledge of the system and exploratory computer runs. Factor levels are then sampled from uniform distributions over these ranges. If the number of treatments evaluated is sufficiently large, then there is a high probability that near-optimal policies will be identified. Often, computer output is obtained only for those treatments for which performance exceeds a specified threshold level. Random search is relatively easy to program on a computer. This experimental design procedure is particularly useful when the levels of policy variables are discrete, and cannot be approximated satisfactorily by continuous variables. This design is also used where activities form complementary or mutually exclusive sets. More sophisticated random search routines include provision for adjustment of probabilities or *heuristic learning* during the experiment; this may involve narrowing the search ranges or departing from uniform distributions.

8. SIMULATION IN PERSPECTIVE

Computer simulation relies on a overall systems philosophy, which asserts that components of a system should not be viewed in isolation and in a reductionist manner, but rather in terms of the complex interrelationships and interactions between variables. An attempt is made to analyse carefully all the components of the system and their interrelationships, and to represent these in a simplified and abstract model. This model is used to conduct

experiments in which the behaviour of the system is simulated for various environments and management policies.

Simulation has proved to be a powerful and versatile approach to improving the understanding of systems and determining near-optimal management policies. A number of advantages are afforded over analytical techniques for problem solving, especially with regard to the handling of uncertainty, multiple goals, non-linear relationships and other real-world complexities. On the other hand, some difficulties arise in simulation studies. The successful application of this technique requires considerable experience with modelling, an understanding of statistical techniques and their limitations, and substantial human and computing resources. It is not usually possible to make use of a recognized model structure or existing computer program, and particular attention must be directed to validation of simulation models. A variety of designs may be exploited in simulation experiments, including optimum-seeking designs which take account of the special features of computer-based experimentation.

REFERENCES

Dayananda, D., Irons, R., Harrison, S., Herbohn, J. and Rowland, P. (2002), *Capital Budgeting: Financial Appraisal of Investment Projects*, Cambridge University Press, Cambridge.

Shannon, R.E. (1975), *Systems Simulation: the Art and the Science*, Prentice-Hall, Englewood Cliffs.

APPENDIX A: GENERATION OF RANDOM VARIATES

A critical step in any stochastic simulation model is to generate values of random (stochastic) variables or *variates*. A wide variety of forms of probability distributions has been recognized as reflecting the behaviour of particular random variables. The starting point for producing (or *generating*) values from any form of probability distribution is to use a random number from a known distribution. In practice, this is usually a number from a uniform distribution over the range zero to one, for example as provided in the Excel function '= RAND( )'.[4]

Generating uniform variates over a specified range

Perhaps the simplest form of *continuous* probability distribution to use in simulation models is the uniform distribution. Here *a* is the smallest value and *b* the largest value that the variable is expected to take. Since the area under any probability curve of function must be unity, and the range of values of the variable is *b-a*, the height of the curve must be *1/(b-a).*

A value y from a uniform distribution can be generated by taking a random number *r* then applying the expression

$$y = a + r(b-a),$$

where the random number is be obtained using the spreadsheet function '=RAND( )', i.e. the lower limit plus the random number times the distance between the upper and lower numbers. For the first demand estimate in Table 1, this becomes

$$y = 5 + 0.8501\ (10-5) = 5 + 4.2505 = 9.25$$

To obtain a series of values on the random variable, simply repeat this process a number of times (on a computer).

Generating values from a discrete distribution

In essence, this involves expressing the distribution in cumulative form, and then associating random numbers with cumulative probability ranges. For example, suppose timber price (in $\$/m^3$) can be

[4] The procedure used on the computer is typically to divide a very large number by another large divisor, and to take the remainder as a fraction of the divisor (hence yielding a number between zero and one). As well, the remainder is multiplied by a third large number, and the product divided by the original divisor. By repeating this process, a series of random numbers can be generated. The numbers are sometimes called 'pseudorandom' because given the same initial three large numbers (and same computer accuracy), the same series will always be produced.

represented by the discrete values 30, 40 and 55, with probabilities 0.3, 0.5 and 0.2. Cumulative probabilities are then 0.3, 0.8 and 1.0, and ranges of random numbers as assigned 0 to 0.3000, 0.3000 to 0.8000 and 0.8000 to 1.0000.

Now suppose the sequence of random numbers 0.2488, 0.8324, and so on is obtained. The first of these numbers falls in the first cumulative probability range hence a timber price of $30/m^3$ is generated. The second number falls in the range 0.8000 to 1.0000 hence a timber price of $55/m^3$ is generated. Proceeding in this way, a sequence of price 'observations' can be generated for which the relative frequencies approximate the discrete probabilities, providing the sample is sufficiently large.

Generating normal variates

The normal distribution is widely recognized in statistical methods as a commonly occurring or approximated distribution. There are several methods for generating random normal variates. The simplest is to take the sum of 12 random numbers from a uniform 0-1 distribution (e.g. a computer random number generator), subtract six, then multiply by the target standard deviation σ and add the target mean μ. The reason this works will not be explained here, but is associated with the Central Limit Theorem.

Generating values from a triangular distribution

A convenient distribution for fitting subjectively to random variables is the triangular distribution, defined in terms of the most pessimistic, most likely (modal) and most optimistic values. For example, for timber price these might be 30, 40 and 55 (in $/m^3$). If these points are called *a*, *b* and *c*, and a distance parameter is defined as

$$d = (b-a)/(c-a),$$

then using a random number r a value from this distribution *y* can be generated as:

if $r \leq d$ then $y = a + \sqrt{r\,(c-a)(b-a)}$

if $r > d$ then $y = c - \sqrt{(1-r)(c-a)(c-b)}.$

Procedures or 'recipes' can be developed along similar lines for sampling from a number of other continuous or discrete forms of probability distributions. Some computer packages in fact have these sampling procedures built in, such as the risk simulation package @RISK.

9. Introduction to Simile

Jerry Vanclay[1]

This module provides an introduction to Simile, a powerful modelling language with some innovative features. Simile is being developed by Dr Robert Muetzelfeldt, at the University of Edinburgh, with suggestions and testing by a number of users. The package is still under development, so it is getting better all the time, but it is already a sophisticated and stable package. Documentation is available at http://www.ierm.ed.ac.uk/simile, including some that was prepared when Simile was known by its previous name, AME (Agroforestry Modelling Environment). The latest version of the software can also be downloaded without charge from this Web site, for PCs with Windows 9x or computers running Linux. Dr Muetzelfeldt has long had an interest in efficient representation of ecological information, and his development of Simile was due in part because of his concern that with many models, the documentation, diagrams and computer implementation diverge. So he set out to build a modelling platform where the diagram was the model and the documentation, and he's close to achieving this goal. He's also provided the possibility to create a model that runs on a computer without the need for computer code or mathematical equations. However, there's no free lunch, and it is necessary to learn some of the standard notation used in systems dynamics to be able to use this package. This module provides introduces the systems dynamics notation, and then provides a worked example of developing a Simile model to represent a personal bank account. Some of the unique features of simile are then examined with reference to a simple forestry model. Finally, two other examples of forestry models are presented briefly.

1. SYSTEMS DYNAMICS NOTATION

Although an effort will be made to keep the jargon to a minimum, the user does need to learn some basic systems dynamics vocabulary to get started. In particular, a common understanding is needed of the following terms: compartment, flow, cloud, variable, influence and equation.

A *compartment*

- is represented by a rectangle
- represents amount of some substance (or occasionally non-substance things like height or position)
- should be labelled as object:substance, e.g. tree volume with units cu_m/ha (or no./ha, kgs, etc)
- unlike real compartments, can go negative, and has infinite capacity
- cannot receive an influence arrow – it changes only as a result of flows
- rate of change is the net effect of all inflows minus all outflows
- two connected compartments must have same substance, same units
- can only contain one substance.

A *flow*

- usually corresponds to a process, and must flow into or out of a compartment
- has units that correspond to compartment units per unit time
- can be negative (flow in reverse direction)
- may be one of several flows between two or more compartments, in either direction
- ideally represents a single process, so it is preferable to have one flow for each separately-analysable process.

[1] This module is a modification of lecture material prepared for students at Southern Cross University in March 2001.

A *cloud*
- is like a compartment (at start or end of a flow), but when its value is irrelevant, unknown, or unspecified
- corresponds to 'the outside world', and thus cannot receive or be the source of an influence arrow
- in Simile, clouds are created automatically when a flow starts nowhere or goes nowhere.

A *variable* can be
- a parameter that is 'constant' during a simulation (e.g. a coefficient in an equation)
- an intermediate variable, which can be both the source and recipient of influence arrows
- an output variable used only for reporting on model behaviour
- an exogenous variable (i.e. an 'external variable', just a function of time) that influences the model, but is not influenced by it (e.g. climatic factors).

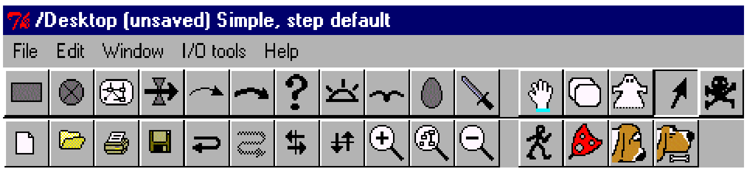
An *influence*
- usually corresponds to a link in an 'influence diagram'
- captures the idea that something affects something else
- formally, represents the fact that one term is used in the calculation of another
- can start from a compartment, flow or variable, and go to a flow or a variable (but *not* to a compartment).

Finally, it should be recognised that an *equation* may appear within a compartment, a flow or an influence. The equation
- says how a value for a variable is to be calculated
- often represents the relationship between one quantity (Y), and one or more influencing quantities (X_1, X_2, etc)
- uses standard algebraic expressions and conditional elements
- may include a sketched graph function or a tabulated function

In Simile, a single 'equation window' is used for all quantities (incl. parameters and initial compartment values).

2. GETTING STARTED

When the new version of Simile is opened, a window like the one opposite is obtained. One of the nice features of this version is that if you're not sure what a button does, hold the mouse over it, and an explanation will appear.



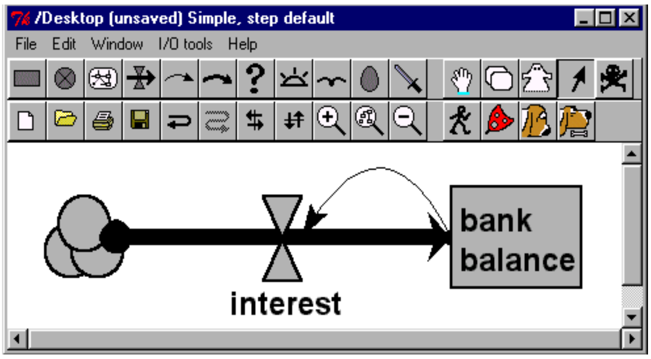
3. BUILDING A SIMPLE MODEL OF A BANK ACCOUNT

To build a simple simile model which will simulate changes in the level of savings over time in a bank account, follow the steps listed below:

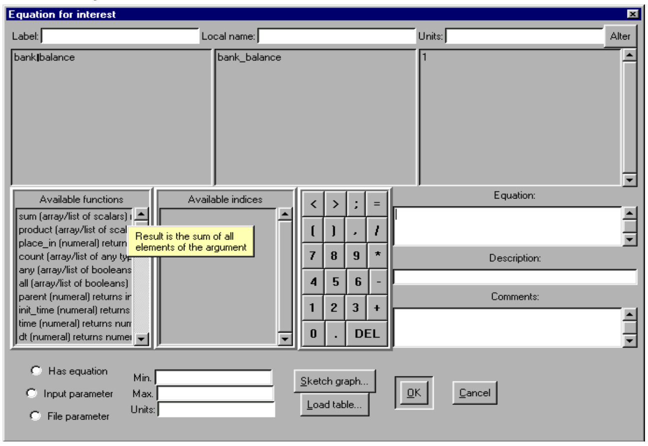
1. Click on the COMPARTMENT tool (left-most rectangle) in the top toolbar.
2. Click in the centre of the 'Desktop' window to deposit a compartment symbol there. *Note that it is automatically labelled 'comp1'.*

3. Click on the FLOW tool (straight thick arrow) in the toolbar.
4. Drag a flow into the compartment symbol in the Desktop window. *'Drag' means:move the mouse onto the Desktop window, depress the mouse button where the flow should begin, move the mouse to touch the compartment symbol (it should turn green), and release the mouse button. Note the flow is automatically labelled 'flow1'*

5. Click on the INFLUENCE ARROW tool (curved thin arrow) in the toolbar.
6. Drag an influence arrow from the compartment symbol to the flow symbol in the Desktop window. *This indicates that the interest payments depend on (are influenced by) the bank balance.*
   Your diagram should now look like the one on the previous page.



7. Click on the POINTER tool (arrow inside a rectangle) in the toolbar.
8. Click on the compartment symbol in the Desktop window.
9. Use the Delete and/or backspace keys to remove the label 'comp1'.
10. Type in a name for the compartment, e.g. balance.
11. Click on the flow symbol in the Desktop window.
12. Use the Delete and/or the backspace keys to remove the label 'flow1'.
13. Type in a name for the flow, e.g. interest.
    We've finished defining the structure of the model, but have not initialized it; that's why it is in red.
14. Double-click on the compartment symbol.
15. Enter the value 100 in the 'Equation' box.
    This states that the initial bank balance is $100.
16. Click on OK. *Notice the compartment is now black, indicating that it has been initialized.*
17. Double-click on the flow symbol (the double-triangle bit).
18. Enter the expression 0.1*bank_balance in the 'Equation' box.
    This expresses the fact that the annual interest is 10% of the balance. Notice that Simile offers help with functions if you place the mouse over them. Also note that you must type names exactly the same way as they appear in the equation box – it is often convenient to double-click on variable names rather than typing them.

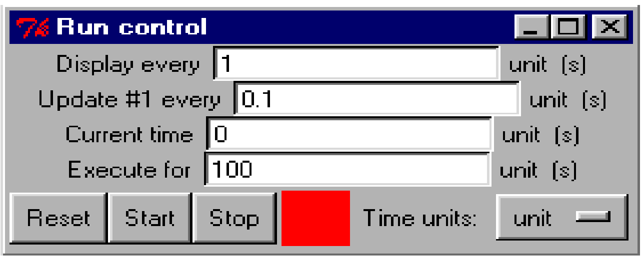


19. Click on OK.
    The model is now completely set up, and ready to run. It should look like this the window opposite.

4. RUNNING THE MODEL

This comprises the following steps:

1. Click on the word 'File' in the Desktop menu bar.
2. Select the 'Run' option in the File menu, then select the 'In tcl' sub-option. *Note that a run control dialogue window appears.*
3. Re-arrange the windows to avoid overlap (drag by top title bar).



4. Click in the 'Update every' edit field.
5. Delete the '0.1' and replace it by '1' (without the quotes).
6. Change the value 100 to 10 in the 'Execute for' edit field.
7. Open the 'I/O tools' menu in the Desktop window.

8. Select the 'Add tool' option, then the 'Plot value against time' sub-option.
   Note the little window that pops up telling you what to do now.
9. Click on the compartment in the Desktop window.
   This sequence specifies that you want to plot values for your bank balance against time.
10. Again, re-arrange the windows if necessary.
11. Click on 'Start' in the Run Control dialogue window.
    The model should now simulate according to the values in the run control window: 10 years, displaying the results every year, and calculating interest once per year. Observe the results in the window titled 'Current value for balance', and note that the actual value of balance is displayed in the box in the bottom-left corner.
12. Try re-sizing the graph window; scrolling up and down; and clicking on the 'Var' and 'Time' buttons.

5. CHANGING THE MODEL

This may be achieved in a number of ways.

By changing the equation

1. Click in the Desktop window to return to the model-design system.
2. Change the interest rate from 10% to 15%.
   You should realise that this involves selecting the pointer tool, double-clicking on the flow symbol in the Desktop window, and editing the interest equation, replacing 0.1 by 0.15.
3. Run the model again, and compare the results with the previous results.

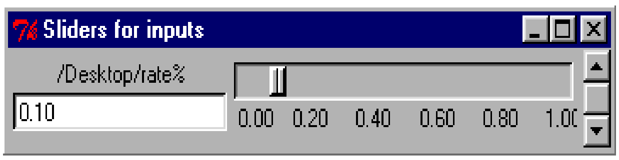
By changing a compartment value during a run

1. With the graph still on the screen, drag the slider near the y-axis down to 50.
   This simulates the effect of removing some money from the bank account, leaving just $50.
2. Click on 'Start' to continue the simulation.
   Note how the graph drops down the climbs back up again. Why do you think the line drops down at an angle rather than dropping down vertically?

With an explicit symbol on the model diagram

1. Click on the VARIABLE tool (circle with a cross inside it).
2. Click near the flow symbol in your model diagram to place a 'variable' symbol there.
3. Re-label this symbol as 'rate'.
4. Double-click on this symbol, enter the value 0.15 in the 'Equation' box, and click on OK.
5. Draw an influence arrow from the variable symbol to the flow symbol.
6. Replace the expression for interest (0.15*interest) by rate*balance
7. Run the model again.
   The results should be the same as before. You have now just changed the diagrammatic appearance of the model, not its underlying mathematical structure. In general, it's up to you whether you show model parameters explicitly or hide them in expressions

By adding a slider to the model



1. Click on the POINTER tool.
2. Double-click on the VARIABLE symbol 'rate' to open the equation window.
3. Click the button 'Input parameter' and supply the Min and Max.
4. Click on the 'Run' tool to re-run your model.
   Notice the appearance of a slider that enables you to vary the interest rate at any time during simulation.

By adding more variables to the model

As an exercise, add features to your model to represent regular deposits (e.g. your salary) and withdrawals (e.g. loan repayments) to your model.

6. SAVING THE MODEL

To save the model, take the following steps:

1. Click on the SAVE tool or select Save under File in the Desktop menu bar.
2. Find the required drive (normally A, C or your home drive) and select the required directory, as you would do in any Windows program.
3. Type the name for the file (e.g. model01). *Simile will automatically add the extension '.sml' to the file name*. *Older versions may add '.ame' or '.sim'.*
4. Click on 'Save'.

So far, the standard systems dynamics notation has been demonstrated. Now some more advanced features unique to Simile will be introduced.

7. SOME UNIQUE FEATURES OF SIMILE

Simile offers some special features that make it uniquely amenable for modelling and teaching. One of these is that the diagram is the model and the documentation, and offers a comprehensive and clear overview of a model, even to the uninitiated. Another unique aspect is the concept of multiple-instance submodels, including several special instances of submodels, both of which are now discussed.

A forestry example of diagram-is-model-is-documentation

Take a quick look at the following four illustrations of the same model implemented in different ways (BASIC, Excel, System dynamics, Simile). You don't need to understand all four approaches; these are presented to emphasise some of the strengths of Simile.

1) Denis Alder's size class model in Basic

```
SUB stproj(st())
'Updates a stand table of tree diameters by species using simple stand projection
'Get array bounds. Lower bound should be 1.
nsp% = UBOUND(st, 1)
ndc% = UBOUND(st, 2)
DIM st0(ndc%)
FOR j% = 1 TO nsp%
        'Keep old stand table as st0 to avoid overlap while updating
        FOR k% = 1 TO ndc%
             sto(k%) = st(j%, k%)
        NEXT k%
        'Get first class update using ingrowth function
        st(j%, k%) = ingrowth(j%) + st0(k%) * (1 – growth(j%, k%) – death(j%, k%))
        'Get updates to other classes
        FOR k% = 2 TO ndc%
             st(j%, k%) = st0(K%-1)*growth(j%, f%-1) + st0(k%)*(1 – growth(j%, k%) – death(k%, j%))
        NEXT k%
NEXT j%
END SUB
```

From Alder, D. (1995), *Growth Modelling for Mixed Tropical Forests*, p. 80.

2) The same model in Excel

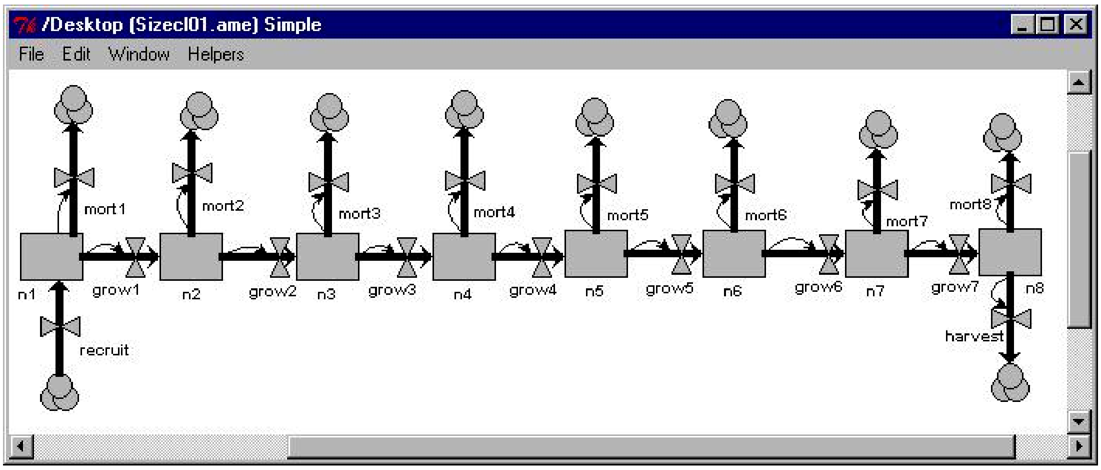
Diameter Class

From (cm)	5	10	20	40	60	80	100	120
To (cm)	=C3	=D3	=E3	=F3	=G3	=H3	=I3	+
Increment cm/yr	0.686	0.77	0.8	0.74	0.625	0.485	0.327	0.24
Outgrowth %/5yr	=5*B5/(B4-B3)	=5*C5/(C4-C3)	=5*D5/(D4-D3)	=5*E5/(E4-E3)	=5*F5/(F4-F3)	=5*G5/(G4-G3)	=5*H5/(H4-H3)	0
Mortality %/5yr	0.14	0.049	0.049	0.049	0.049	0.049	0.049	0.049
Harvest %								0.5
Year	←--- Trees/km^2 ---→							
0	500	321	232	146	164	136	64	14
=A12+5	=B12*(1-B$6-B$7)+87	=B12*B$6+C12* (1-C$6-C$7)	=C12*C$6+D12* (1-D$6-D$7)	=D12*D$6+E12* (1-E$6-E$7)	=E12*E$6+F12* (1-F$6-F$7)	=F12*F$6+G12* (1-G$6-G$7)	=G12*G$6+H12* (1-H$6-H$7)	=H12*H$6+I12* (1-I$6-I$7)*I$8
=A13+5	=B13*(1-B$6-B$7)+87	=B13*B$6+C13* (1-C$6-C$7)	=C13*C$6+D13* (1-D$6-D$7)	=D13*D$6+E13* (1-E$6-E$7)	=E13*E$6+F13* (1-F$6-F$7)	=F13*F$6+G13* (1-G$6-G$7)	=G13*G$6+H13* (1-H$6-H$7)	=H13*H$6+I13* (1-I$6-I$7)*I$8
=A14+5	=B14*(1-B$6-B$7)+87	=B14*B$6+C14* (1-C$6-C$7)	=C14*C$6+D14* (1-D$6-D$7)	=D14*D$6+E14* (1-E$6-E$7)	=E14*E$6+F14* (1-F$6-F$7)	=F14*F$6+G14* (1-G$6-G$7)	=G14*G$6+H14* (1-H$6-H$7)	=H14*H$6+I14* (1-I$6-I$7)*I$8

From Alder, D. (1995) *Growth Modelling for Mixed Tropical Forests*, p. 78.

3) The same model in a systems dynamic representation

This is an exact re-implementation of the spreadsheet model given above. It is implemented in Simile, but uses only standard system dynamic notation used in industry-standard packages like Stella (from Muetzelfeldt's submodel library).

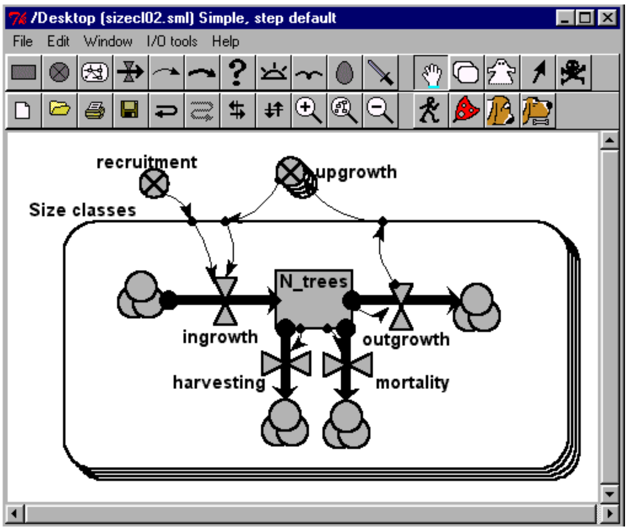


Initial values:
n1...n8 = 500, 321, 232, 146, 164, 136, 64, 14
Equations:
recruit = 87
mort1...mort8 = 0.14*n1, 0.049*n2...0.049*n8
grow1...grow8 = 0.686*n1, 0.385*n2, 0.2*n3, 0.185*n4, 0.156*n5, 0.121*n6, 0.082*n7, 0*n8
harvest = 0.5*n8

4) The same model in Simile, using a multiple-instance submodel

This is the same model, but using a multiple-instance submodel to represent a generic size-class. It has eight instances, since the published model has eight size-classes.

Initial values:
recruitment = 87
N_trees = element([500,321,232,146,164,136,64,14],index(1))
Equations:
mortality = if index(1)==1 then 0.14*N_trees else 0.049*N_trees
outgrowth = element([0.686,0.385,0.2,0.185,0.156,0.121,0.082,0],index(1))*N_trees
harvesting = if index(1)==8 then 0.5*N_trees else 0
upgrowth = [outgrowth]
ingrowth = if index(1)>1 then element([upgrowth],index(1)-1)else 0



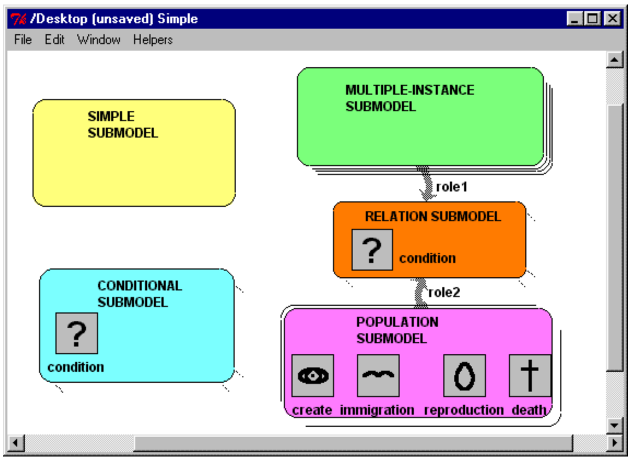
Which representation is clearer and easier to follow? Why?

Submodels

One of the powerful features you've just seen is the ability to create a submodel.

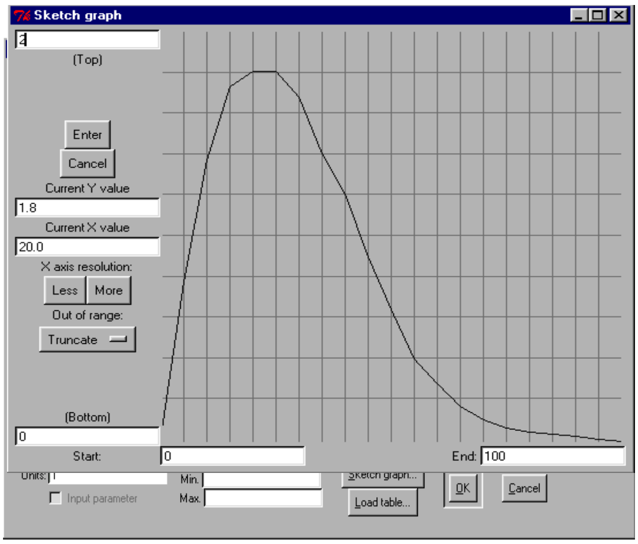
Submodels can be used to

- visually separate parts of the model
- control the appearance of a complex model
- enable separate saving and loading
- move parts of a model around
- create multiple instances
- specify relations between objects
- allow parts of the model to exist conditionally
- specify dynamically-varying populations.

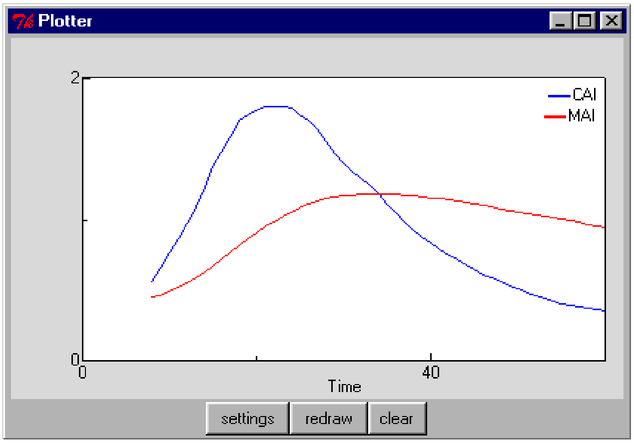


Using submodels to model individual trees

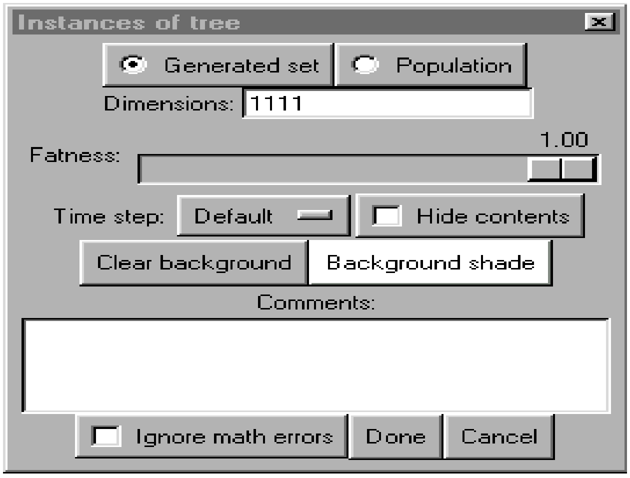
Let's build an individual tree growth model with Simile, starting with a simple model of the growth of a single tree. It's much like the bank-balance model illustrated in Section 3, but here the compartment is called diameter, and the flow is called grow. Suppose there is no established equation for tree growth, and that the model is to be constructed using a graph to represent growth. Double-click on the flow and click on the 'Sketch graph' button. The Sketch graph dialogue window appears, with its default straight line and 0,0 to 100,100 coordinates. Since growth is to be predicted from diameter (in



cm dbh), 0–100 is probably OK on the horizontal axis, but a more appropriate maximum is needed on the vertical axis (let's chose 2, a reasonable maximum when growth is expressed in cm/year). Next, 'draw' the response curve with the mouse, either by dragging, or by clicking on the grids. When this is finished, press Enter to return to the equation dialogue. Notice that "graph()" has appeared in the equation box; this must be completed by double-clicking on the local name diameter to complete the equation, or Simile won't know how the graph should be used.

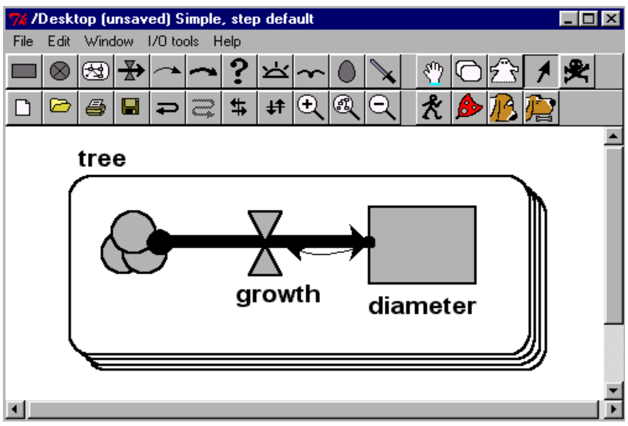


Let's also compute the mean annual increment (MAI) and compare the current (CAI or growth) and MAI curves. Create a variable called MAI, and drag an influence arrow from diameter to MAI. Double-click on the variable MAI and enter the formula for MAI: "if time(1)==0 then 0 else diameter/time(1)". The function time(1) is a system function that returns the current simulated time. Notice that Simile uses the notation "==" to indicate a test of equality to discriminate it from the common usage of "=" indicating "takes the value of".



Remember how to run the model? (If not, see Section 4 above). An "I/O tool" is needed to illustrate the CAI and MAI curves. The screen opposite illustrates the plotter. To use this, select "I/O tools", ""Add tool" and "Plotter".

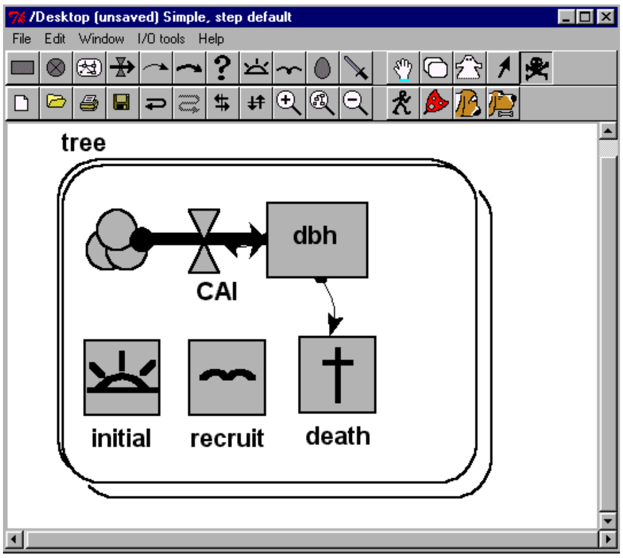
Now use the submodel tool to include all these details in a submodel. Suppose it is desired to model a plantation with 1111 trees (3 x 3 metre spacing). Click on the pointer tool and double click on the white space inside the model. A submodel dialogue box should appear, and you can specify the number of instances (1111), choose a colour if you wish, and add some comments. When this is done, the submodel diagram will have changed, now appearing like a pack of cards, to represent multiple instances.



If the screen does not appear like this, use the move tool and the zoomfit tool to tidy it up a bit.

Now try to compute the basal area. Use the variable tool to create a variable called BA outside the submodel, and drag an influence arrow from the diameter compartment to the BA variable. Double-click to open the equation dialogue. Notice that the influence arrow makes an *array* of diameters available to this variable. Double-click the *sum* function, double-click the local name [diameter], and type ^2 to square it. Next, move the cursor to the end and divide by 12732, which is equivalent to 40000/π. π is not yet available as a function in Simile. The figure 40000 converts diameter squared into radius squared, and cm^2 into m^2. Try running the model, and explore some of the I/O tools.

Let's adapt this model for uneven-aged forests, and regeneration. Double-click on the submodel again, and chose the Population button. When this is done, you'll notice that the 'pack of cards' looks 'slightly untidy', signifying that the number of instances can vary. A population submodel also activates several new symbols:

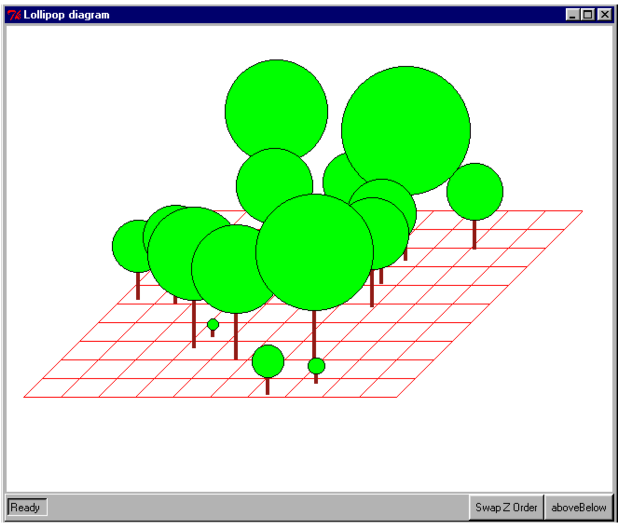


- *creation* tells Simile how many instances exist at the start of a simulation.

- *reproduction* creates 'offspring' for each existing instance; good for modelling regeneration in species with short-lived seed, and for modelling animal populations (the symbol is an egg).

- *immigration* creates new instances without regard to the existing instances; good for modelling regeneration from seed banks, of species with widely-dispersed seed, and animal migration.

- *loss* provides a way to model mortality, emigration, etc. (the symbol is a sword or a cross).



Alder's model will now be represented as a population model, so it can be compared with the illustrations above. A few numbers will need to be changed, to adopt an area of one hectare (not a square kilometre), and to use annual time steps instead of 5-year steps. Here's the model. Does it make sense to you? One thing that it portrays clearly is that growth and death depend on tree size (and tree size only), and everything else remains constant throughout the simulation.

One further embellishment will be made, namely adding x- and y- coordinates for each tree so that a 'lollipop' diagram can be created. Module 10 demonstrates how to use these Simile functions, developing a forestry example model. Also, some tutorials on Simile can be accessed on the Web site: http://www.ierm.ed.ac.uk/simile.

8. CONCLUDING COMMENTS

Simile is a powerful modeling and programming medium, which has proven valuable in modeling forestry systems. As with any computer programming language, or software package, there are a number of overheads involved in gaining familiarity with simile. The package is available as freeware, so anyone can obtain a copy and start using it. It is obviously necessary to learn the systems dynamics notation, and to become familiar with the menu items of the simile screen. After that, working through example applications such as the bank account model presented here provides initial confidence for the user. To become a competent user, it is essential to experiment with the Simile package, including the menu and sub-menus, and find out by trial and error what operations can be performed. Handy help facilities are available to assist this experimentation.

10. Simile Revisited: The One-Minute Modeller

Jerry Vanclay

The Simile programming language provides a powerful and relatively easy to use medium for developing models and simulating the behaviour of forestry systems. This is a highly visual approach to modelling, in that the flow diagram is in effect the computer program. This module provides a simple introduction to use of the Simile programming language for potential users, which has been developed to provide an initial understanding of the programming features and steps in classes and for workshops.

1. INTRODUCTION

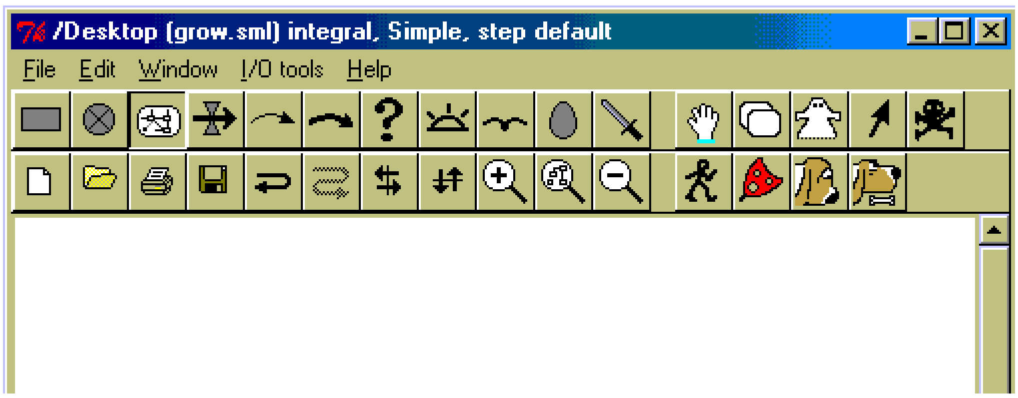
A variety of computer programming languages have been used for developing models to simulate forestry biological processes, stand growth, processing, marketing, human involvement and other 'systems' in forestry. Over time, these model development media are becoming more powerful and easy to use. The Simile programming language, originally known as AME (Agroforestry Modelling Environment), has been developed by researchers at the University of Edinburgh and elsewhere, during the last five years, with a specific focus on forestry modelling. This software can be downloaded free of charge from the web site http://www.simile.co.uk. Other languages such as Vensim and Stella offer similar capabilities.

A best-selling book was recently written called 'The One-minute Manager'. In the same spirit, this exposition of Simile can be thought of as 'The One-minute Modeller'. That is, it is a tutorial presentation designed to 'break the ice' for researchers interested in using Simile, by demonstrating some of the main features of the language in an easy-to-understand manner.

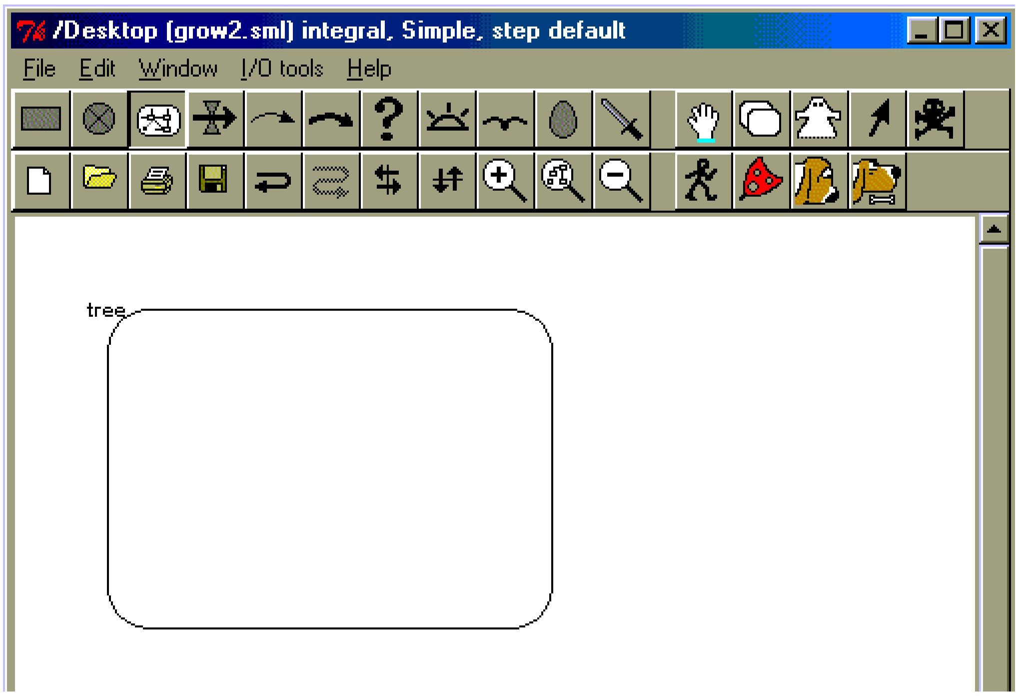
The example used for this exposition concerns growth of a forestry stand. However, the principles apply equally well to other application areas. A number of computer screen images will be presented, to indicate menu options and appearance of the visual model throughout the development steps. While there is much talk about models that are big and expensive, this paper demonstrates that modelling need not be slow, expensive, or excessively demanding of data.

2. DEVELOPING THE MODEL

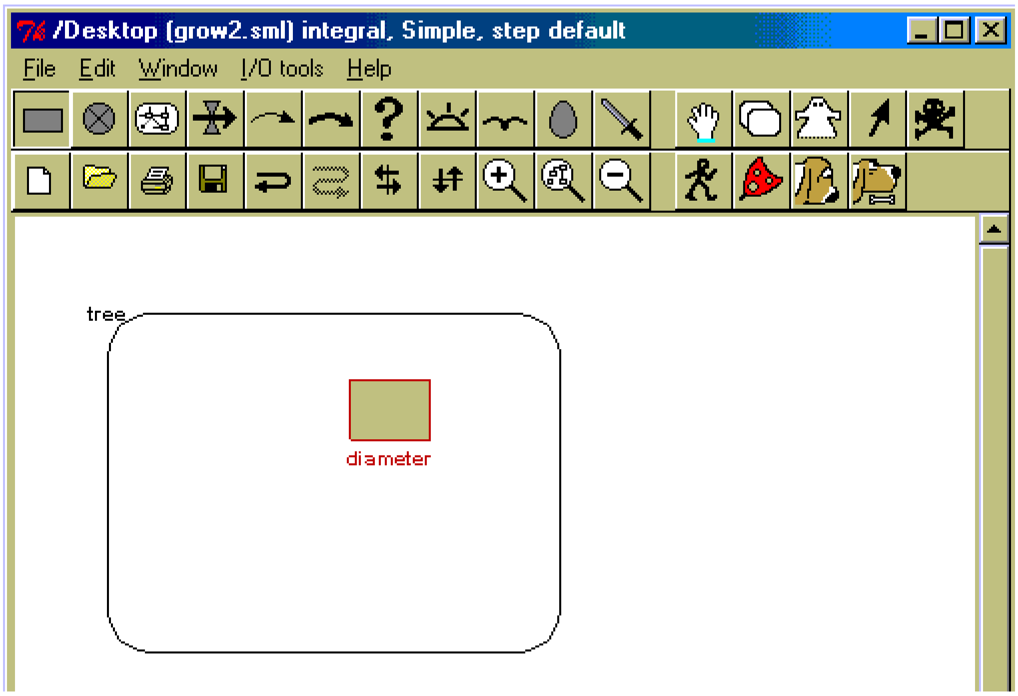
When Simile is first opened, the screen appears as below. To save space, all screen images have been reduced slightly in size in this paper, and unfortunately only black and white images can be provided here. The purposes of the various icons at the top of the screen will be explained during presentation of this application.



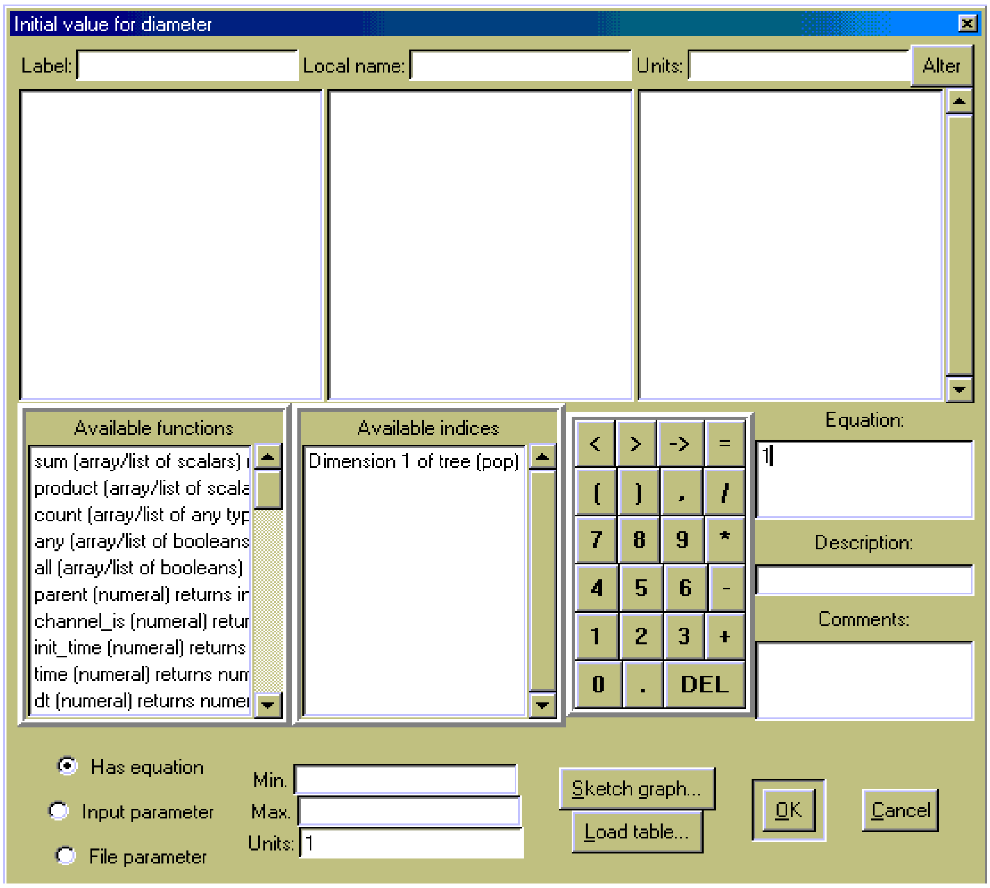
A starting point is to introduce an individual tree, and to set out what is known about trees as a submodel. Select the submodel tool by clicking on the icon showing a box with rounded corners, and then click within the modelling space at the desired location for this submodel. The submodel has been renamed 'tree', using the pointer tool in the toolbar.



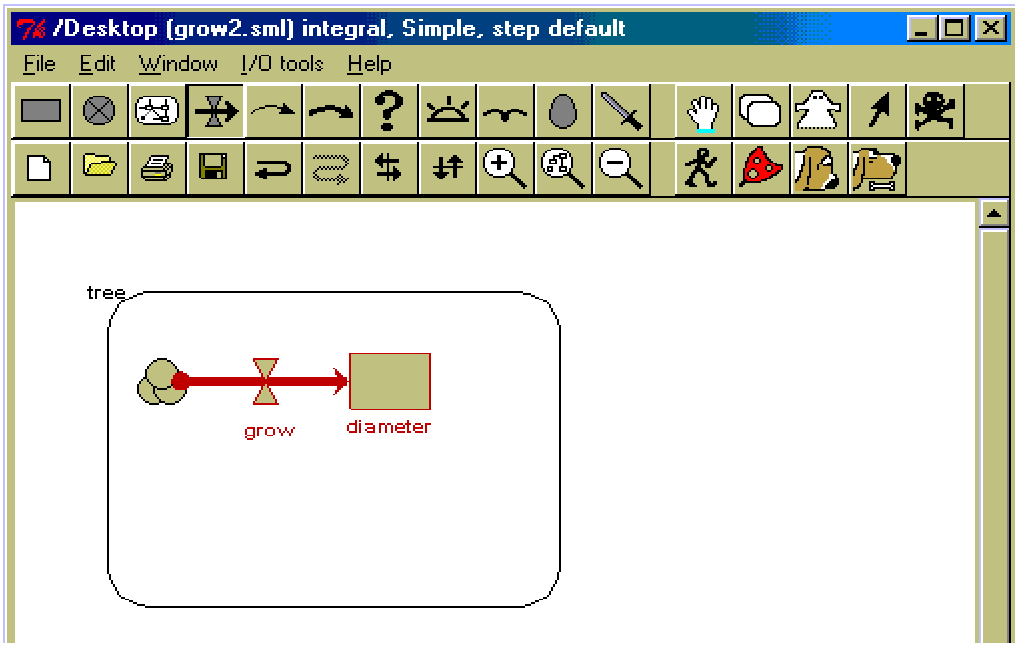
For this example, the only property of trees of interest is their diameter. Thus a compartment is added to the model diagram to represent tree diameter. Click on the compartment tool (the rectangle icon at top left) and move the mouse to the middle of the modeling space. This compartment has been renamed 'diameter', again using the pointer tool.



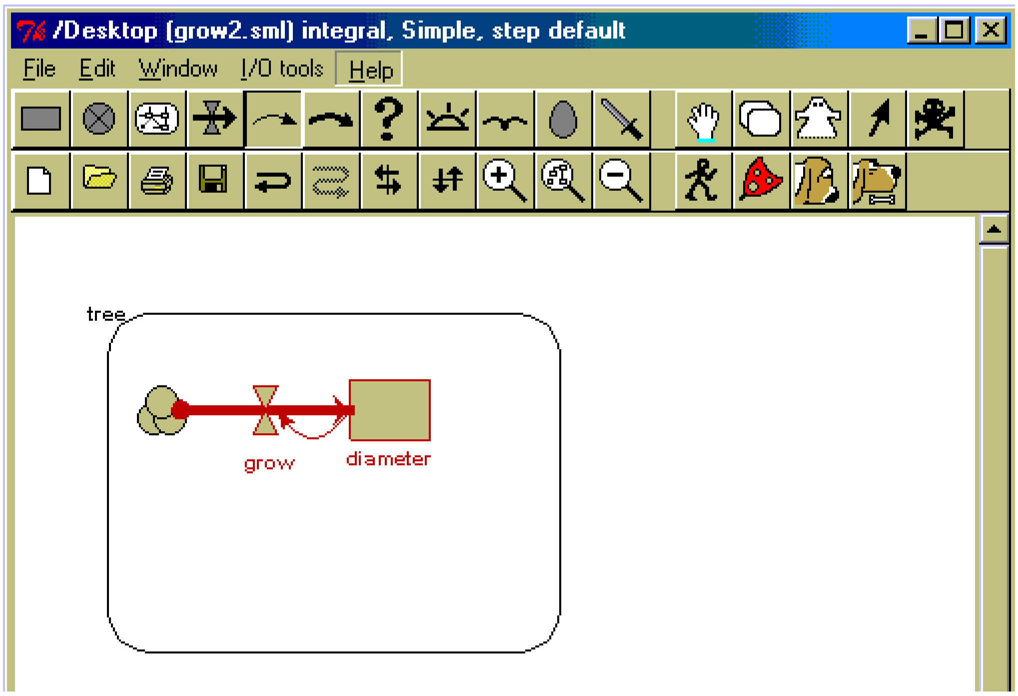
Double-click on the compartment symbol to obtain the following dialogue box. Enter an expression of say 1 in the Equation box. Here the diameter is specified in centimeters.



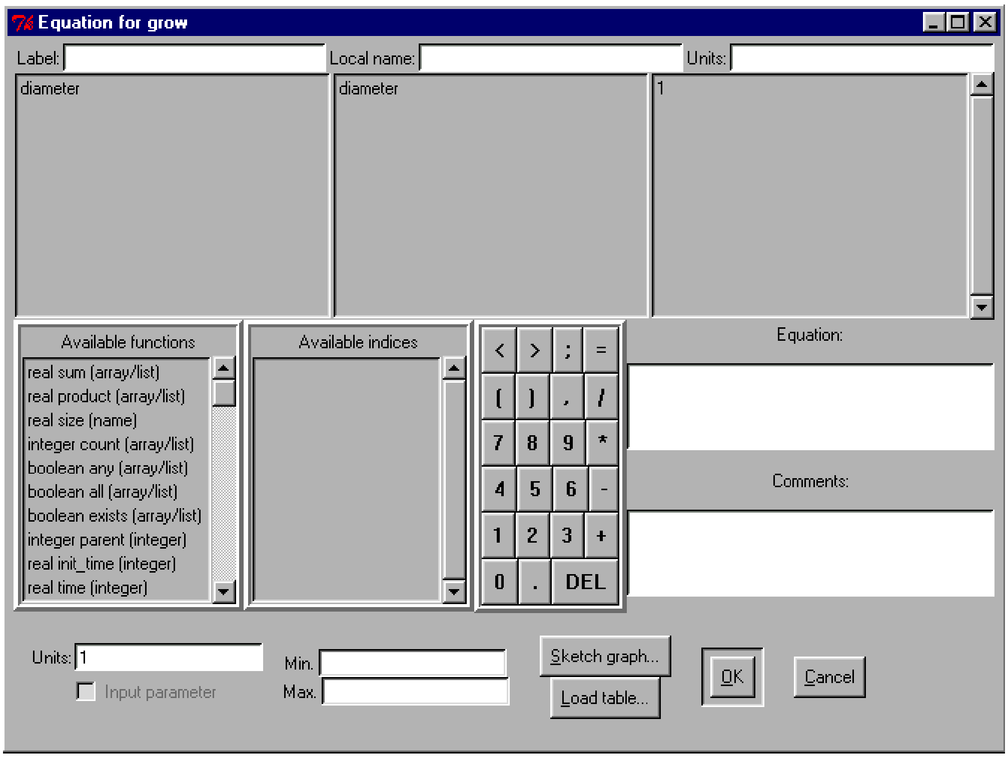
Trees grow, so it is necessary to allow the diameter to increase. This is achieved by creating a flow into the compartment called diameter. The flow tool is the icon showing an arrow with an engineering symbol for a tap or regulator. The flow has been renamed as 'grow'.



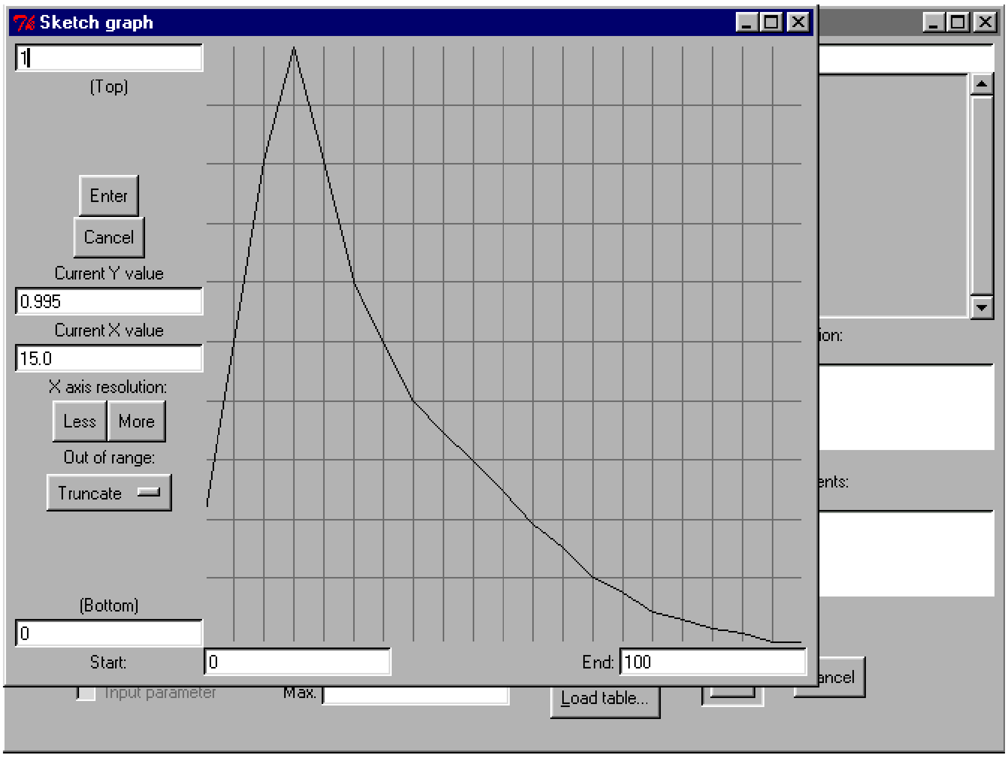
Tree growth may depend on many things, but here only size dependent growth will be modelled. This calls for drawing an influence arrow from 'diameter' to 'grow' in the model diagram. Click the influence tool with a curved thin arrow in the toolbar and draw an influence arrow from the compartment symbol to the flow symbol.



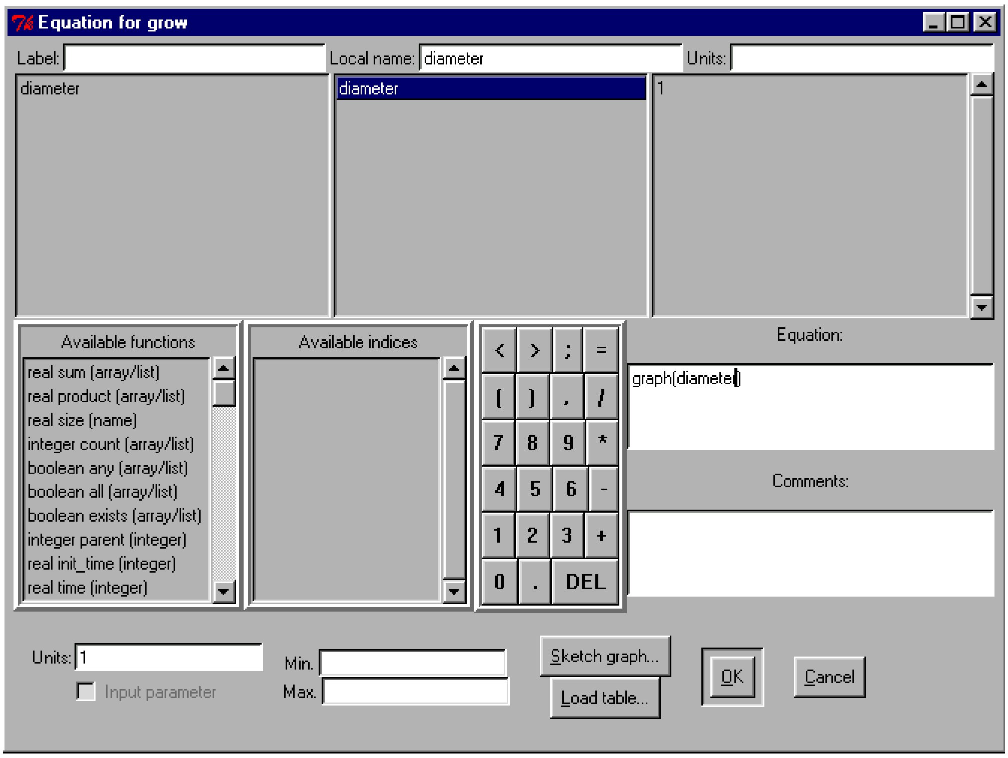
Next, a growth relationship must be supplied. By double-clicking on the flow icon called 'grow', the following dialogue box is obtained:



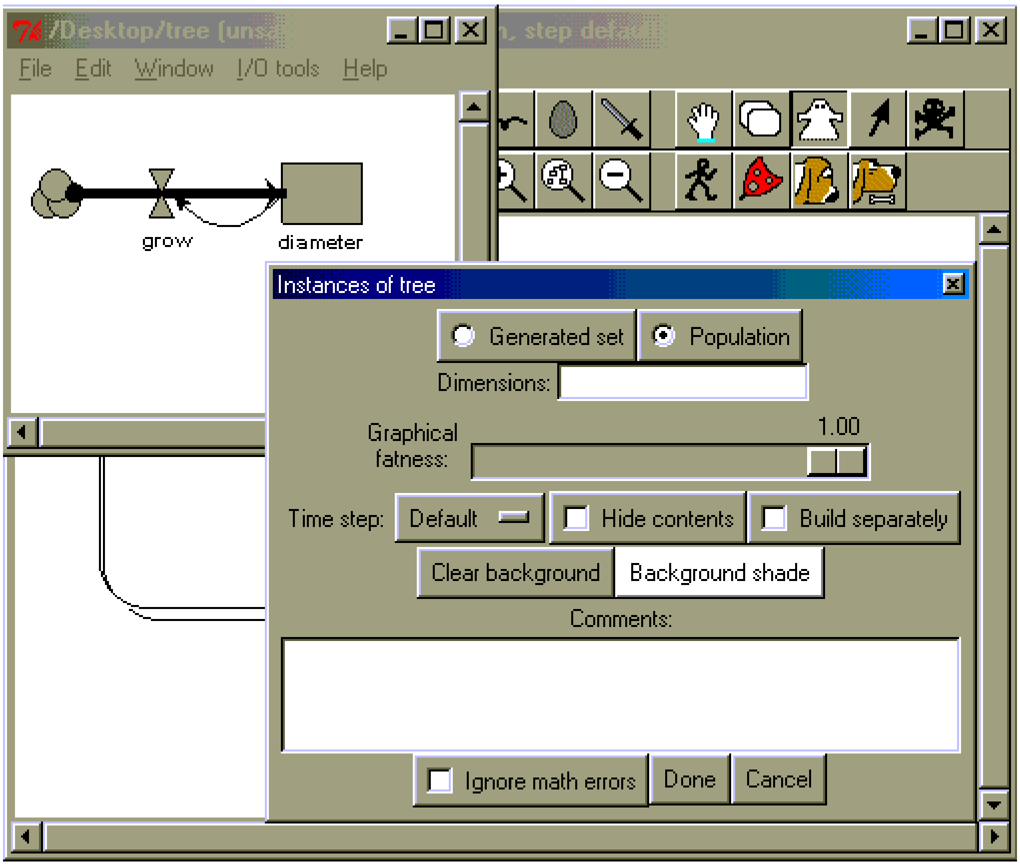
Not everyone is comfortable with equations. To avoid the maths, the relationship may be supplied as a hand-drawn sketch. By clicking on the 'Sketch graph' button, a graph window is obtained. A graph of any shape can be sketched with the mouse on the graph window. The default is to interpolate on 20 line segments, but this can be varied with buttons for "Less" and "More" X-axis resolution.



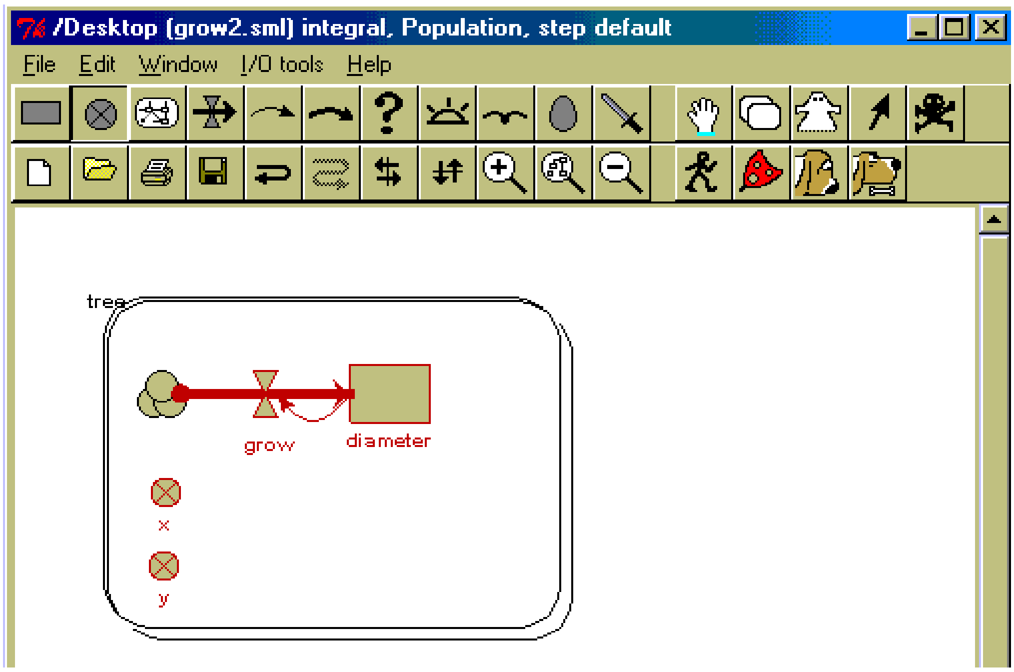
Notice that start and end values for both the x and y axes can be provided. The relationship can now be accepted by pressing the 'Enter' key. Simile knows that the y-axis represents the growth rate, but it is necessary to specify in the 'Equation' box that the x-axis represents the diameter of a tree. Having done that, press the 'OK' button. A model has been built for the first tree for the moment.



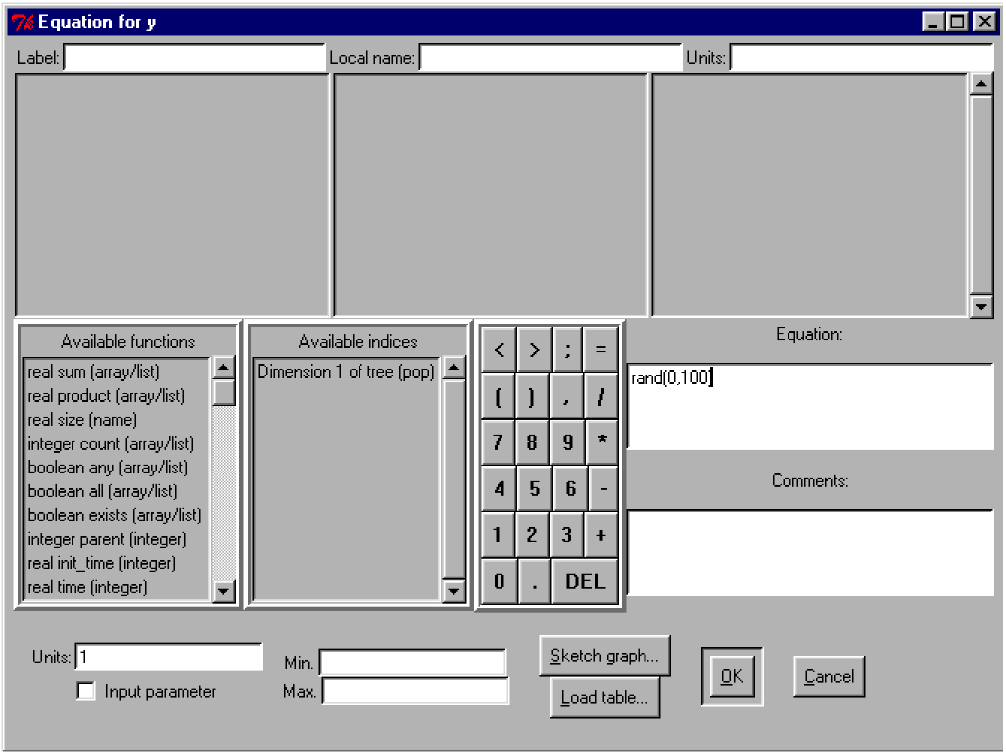
Now, the objective is to model not just one tree, but a forest. To do this, double-click on the boundary of the tree submodel. A new window for the tree submodel appears. In this window, select 'Properties' under 'Edit' on the menu bar. Next, specify the number of instances of the tree submodel in the 'Dimensions' box. If the aim was to model a plantation, a set of say 1000 trees may be generated. However, here the interest is in a natural forest, in which the number of trees varies continually. Thus, just check that the population option has been chosen, with a dot in the 'population' circle.



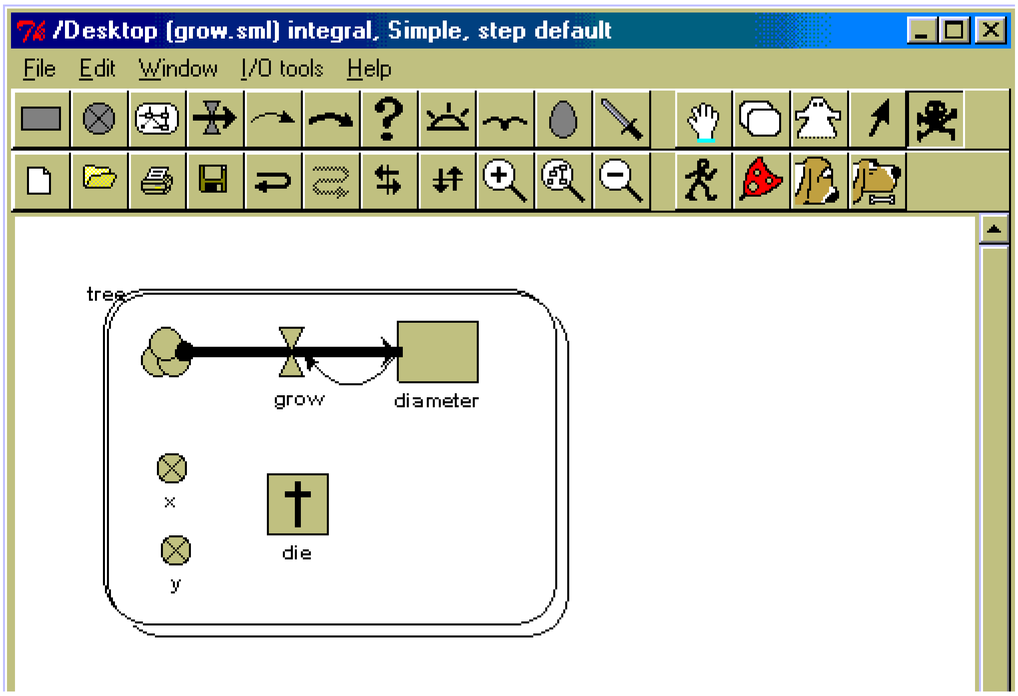
When the 'Done' button is pressed, it will be noticed that the tree submodel is no longer represented as a single instance, but has double lines around the bottom-right and top-left corners like an untidy pack of cards. This is the Simile notation for a population. Now, in order to draw some diagrams of the forest, it is necessary to add variables representing the x- and y-coordinates. Click on the variable tool (a circle icon with a cross in it), and then click in the model diagram to deposit a variable symbol. Repeat clicking in the model diagram to add additional variables. Two variables have been renamed as x and y.



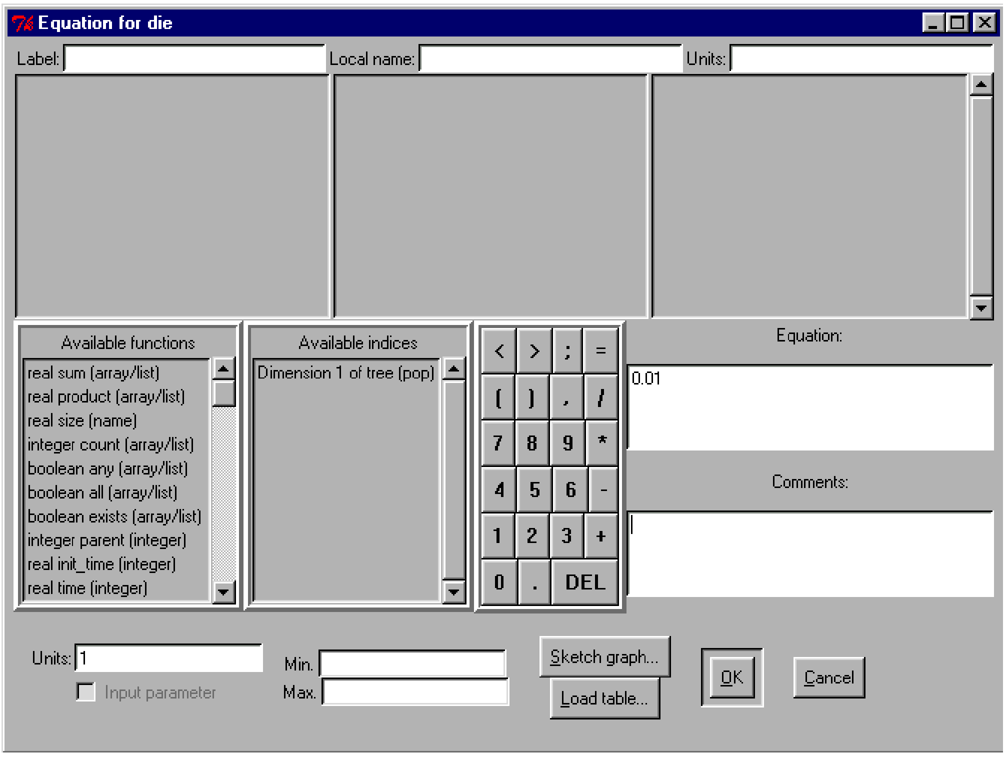
By double-clicking on each variable symbol, the following dialogue box is obtained. It is necessary to specify the initial value for each tree. For simplicity, random numbers between 0 and 100 will be chosen for each tree. Enter the expression rand(0,100) in the Equation box. The location, size and growth of trees have now been specified.



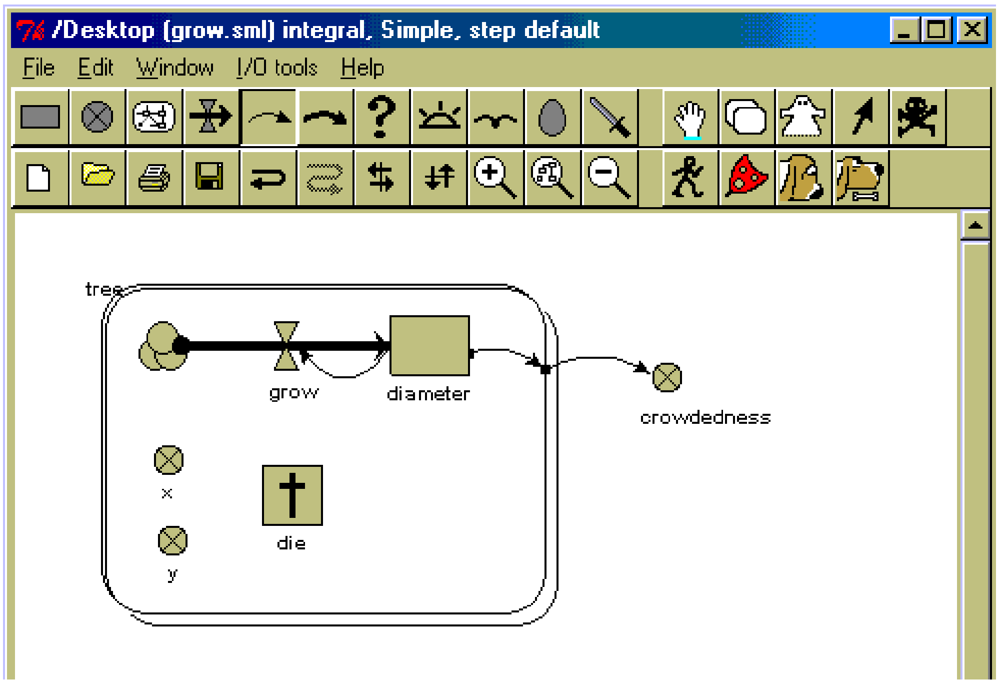
Since these trees are not immortal, it is necessary to specify a mortality function. The mortality tool is used to specify the destruction of instances of a population submodel. Click on the mortality tool (the icon with a sword in the toolbar), and then move the mouse to within the envelope of the population submodel to deposit the mortality symbol in the submodel. Notice that the icon with a cross instead of a sword turns up. The next version of Simile will use a hatchet for both the tool bar and the icon, but for the moment, Simile users just have to accept both the sword and the cross and recognise that they mean the same thing.



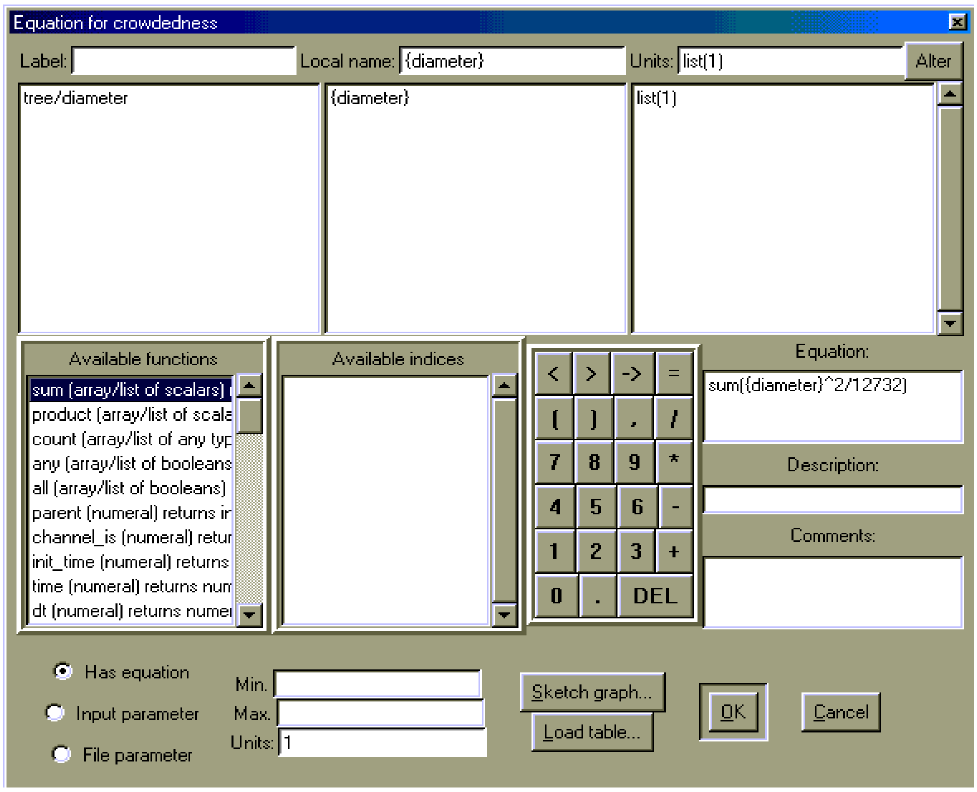
By double-clicking the mortality symbol renamed as ‘die’, the following equation dialogue window appears. Suppose a 1% mortality rate is assumed, irrespective of size, competition or any other factor. To express this, enter 0.01 in the Equation box.



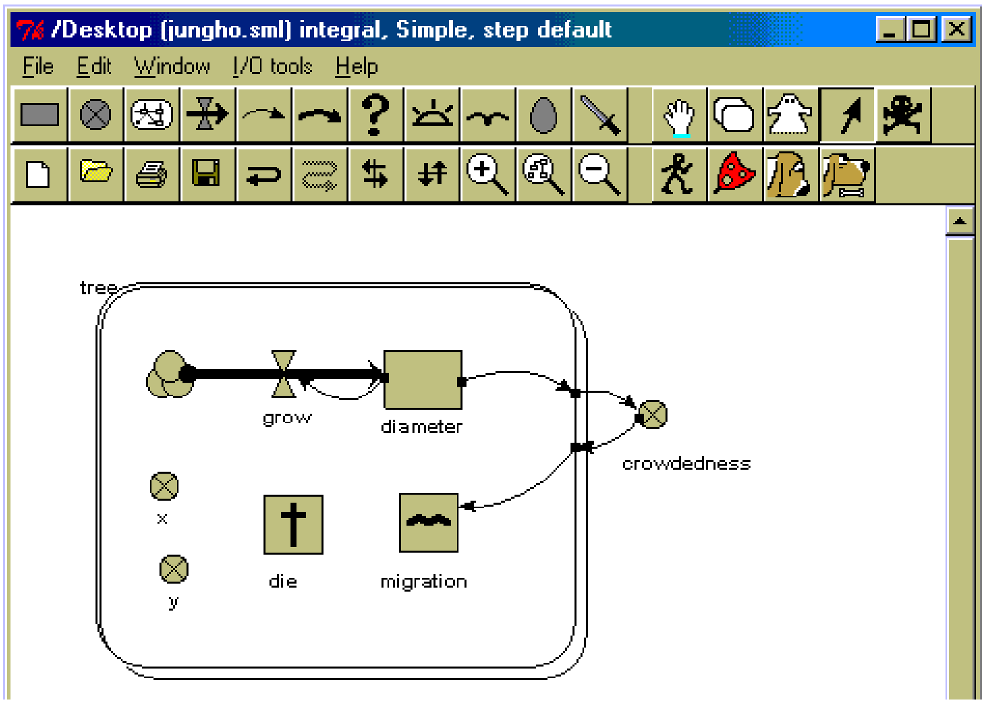
Next, to introduce new trees into our population as seedlings, while making them density-dependent, it is necessary to introduce the notion of crowdedness. Place a variable symbol outside the population submodel and rename it ‘crowdedness’. The level of crowdedness will depend on the number and sizes of the trees in our population. Thus, ‘crowdedness’ receives an influence arrow from ‘diameter’.



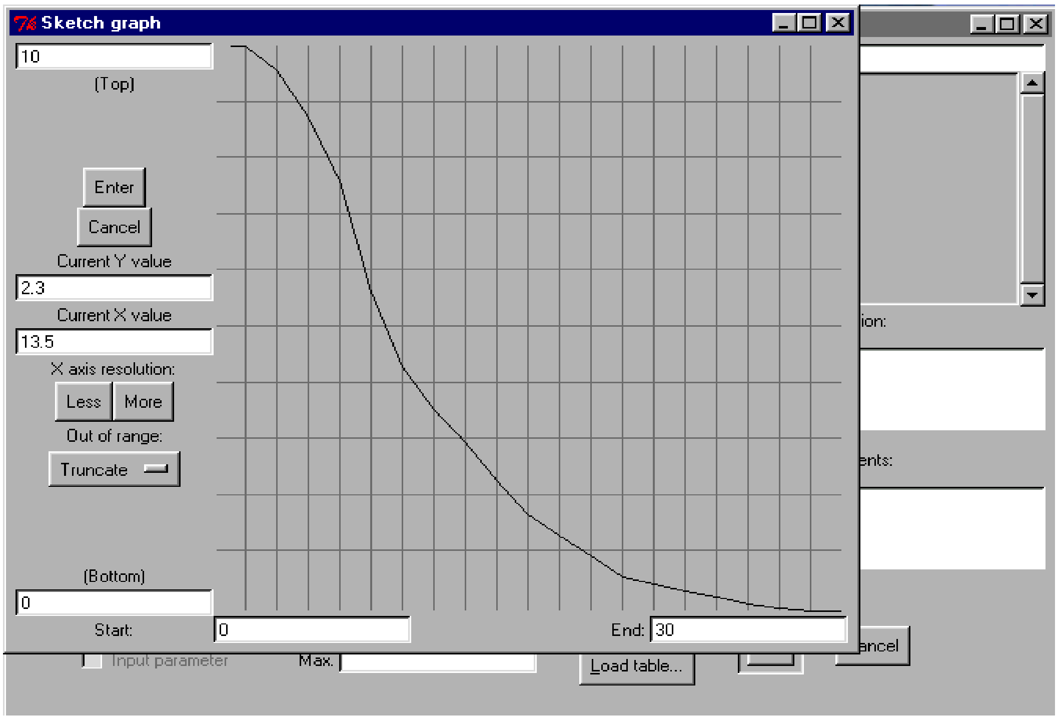
The forester's concept of stand basal area is adopted here, which is the sum of the sectional areas of the tree stems, because it is easy to compute, and because it is well correlated with many interesting properties of a forest, such as biomass, biodiversity and evapo-transpiration. Stand basal area is simply the sum of the squares of the diameters, converted to sectional area in square metres per hectare. Double-click on the variable symbol for 'crowdedness' to obtain the following dialogue box. Double-click the sum function from the 'Available functions' box at left in this window. Next, double-click the local name {diameter} and type ^2/12732 to square and divide by 12732 the diameter. The number 12732, which is equivalent to 40000/π, is used because π is not yet available as a function in Simile. The 40000 converts diameter squared into radius squared, and cm^2 into m^2.



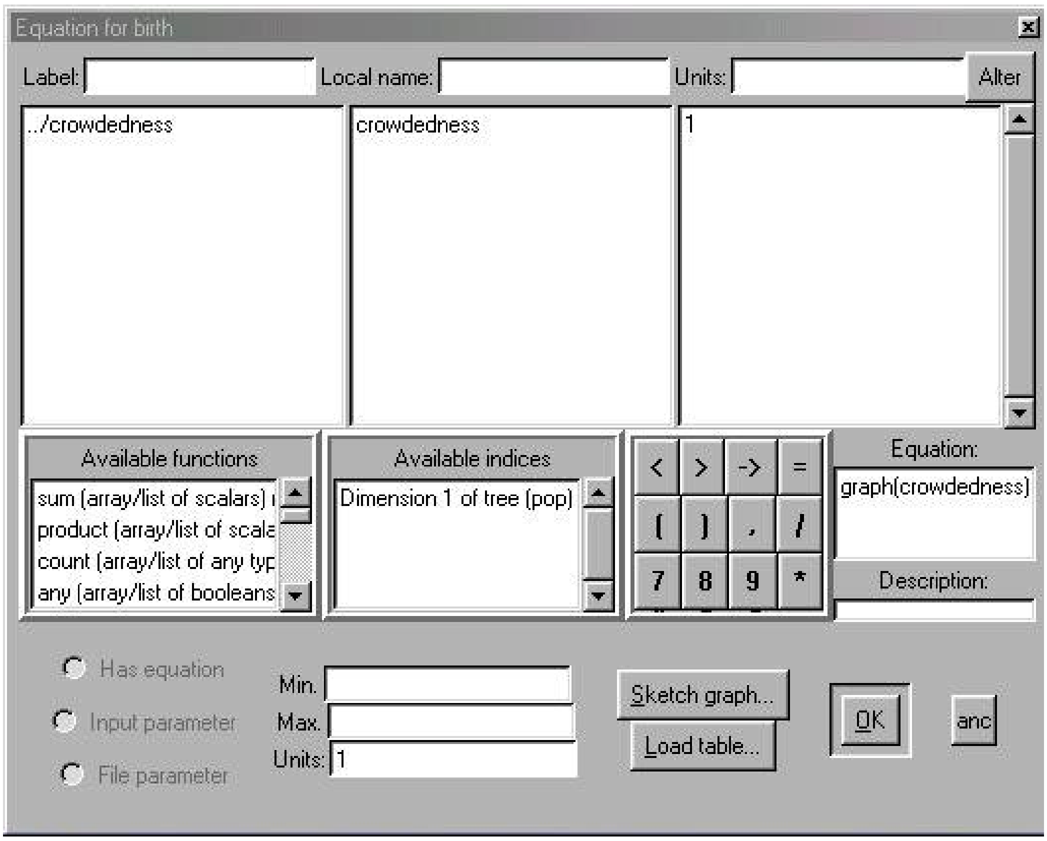
Recruitment of seedlings into the population may be represented as 'migration' (the bird symbol) rather than 'birth' (the egg symbol), to recognise that seeds may be transported by birds, wind, and other factors. This process has been renamed as 'migration'. An influence arrow goes from 'crowdedness' to 'migration'.



Double-click on the bird symbol to obtain a dialogue box. Again to avoid equations, click on the 'Sketch graph' button. A graph may be sketched illustrating the negative relationship between stand density and the number of seedlings that survive to the end of the first year. The graph predicts numbers of recruits (stems/ha/year) from stand basal area (m^2/ha). It seems reasonable to assume say 10 recruits at low stand density, with an asymptotic decline approaching zero as stand basal area reaches 30 m^2/ha. The relationship can now be accepted by pressing the 'Enter' key.

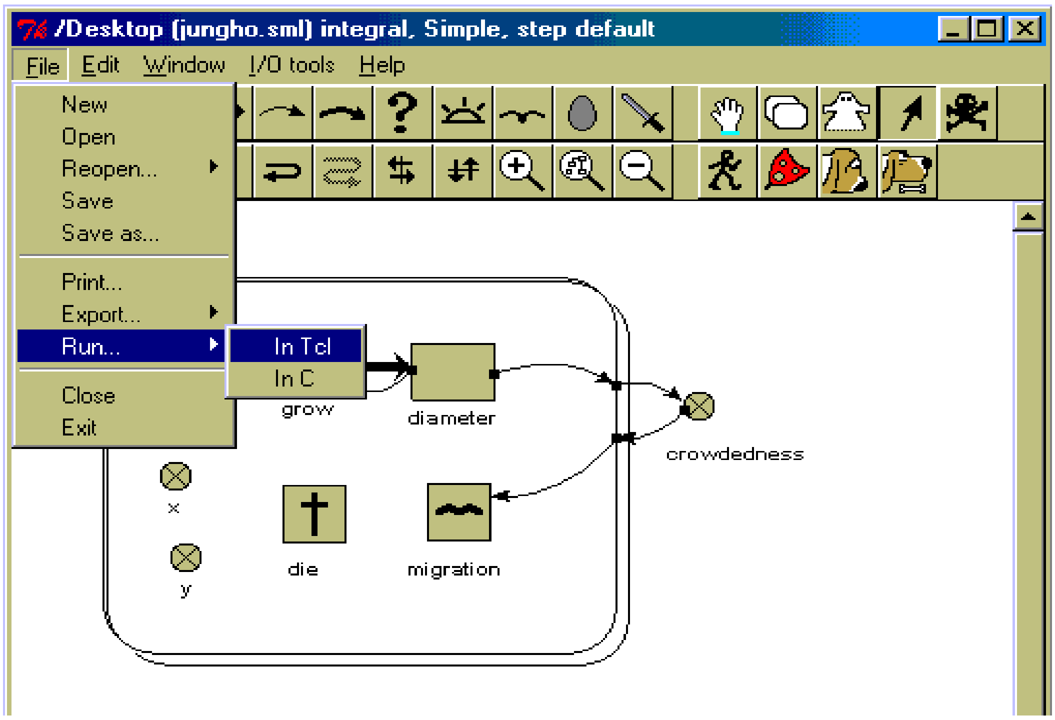


Simile knows that the y-axis represents 'migration', but it is necessary to specify in the 'Equation' box that the x-axis represents 'crowdedness'.

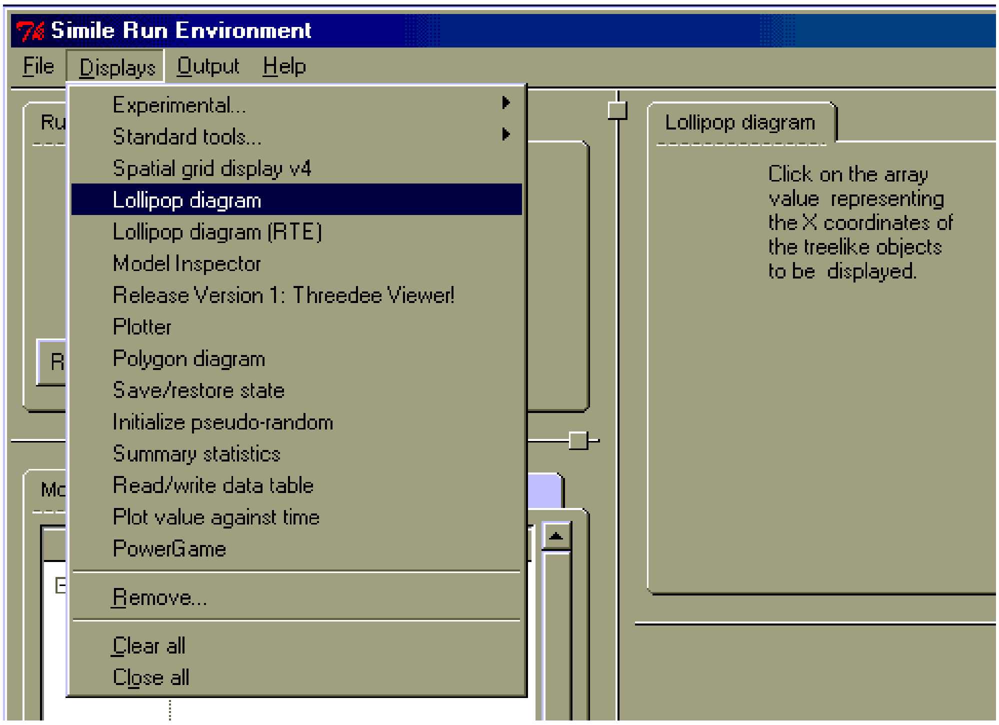


3. RUNNING THE MODEL

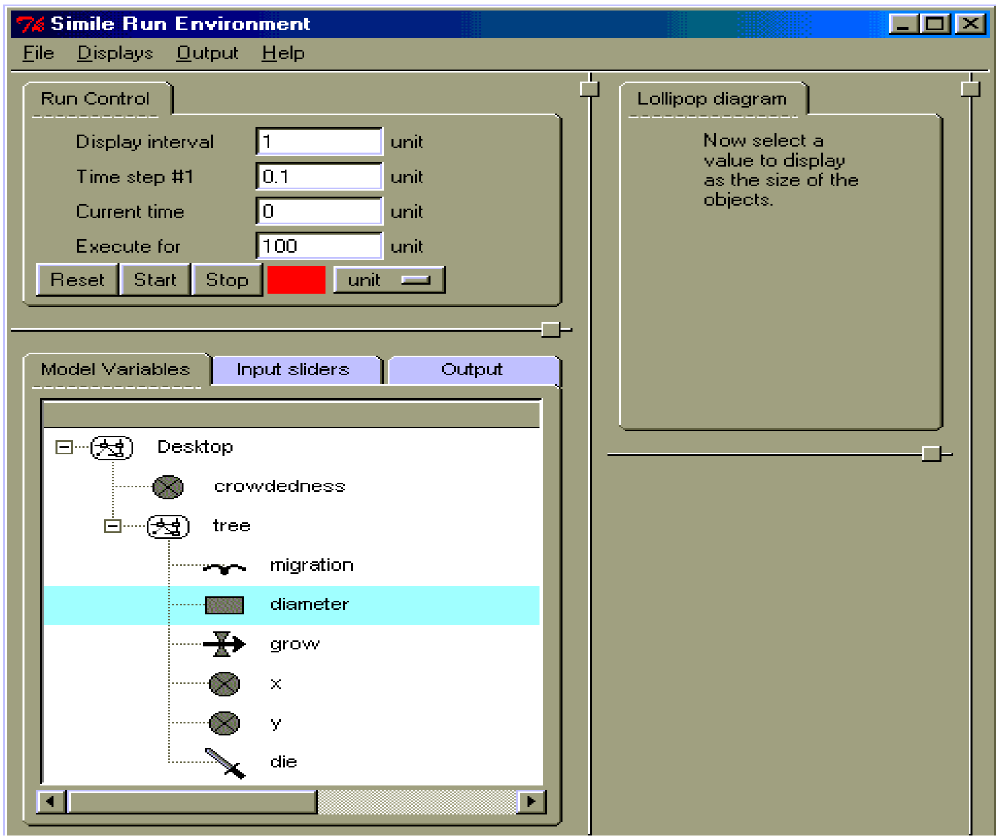
Now that the model has been developed, it is time to run it and inspect the output. In this demonstration, the model will be run with the TCL interpreter, rather than cross-compiling it as C-code allowing run time to be reduced. The two options are available in the Run command under the File menu.



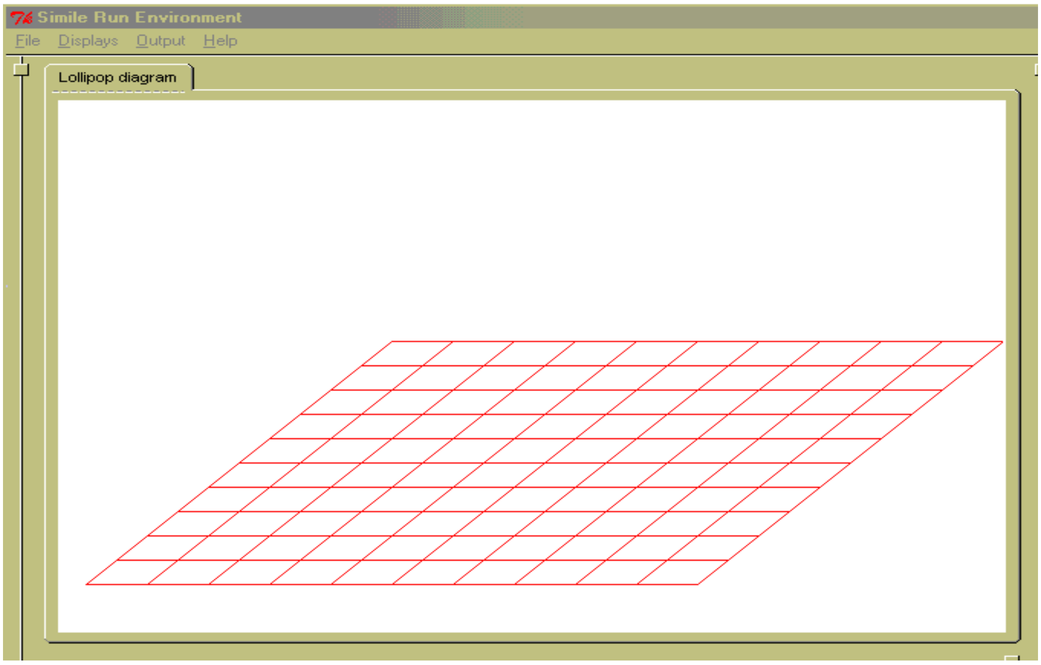
When 'In Tcl' under File/Run is selected, the Simile Run Environment widow appears. Some displays may be chosen to improve the output display. One of the standard Simile displays will be used here, namely the 'lollipop diagram'.



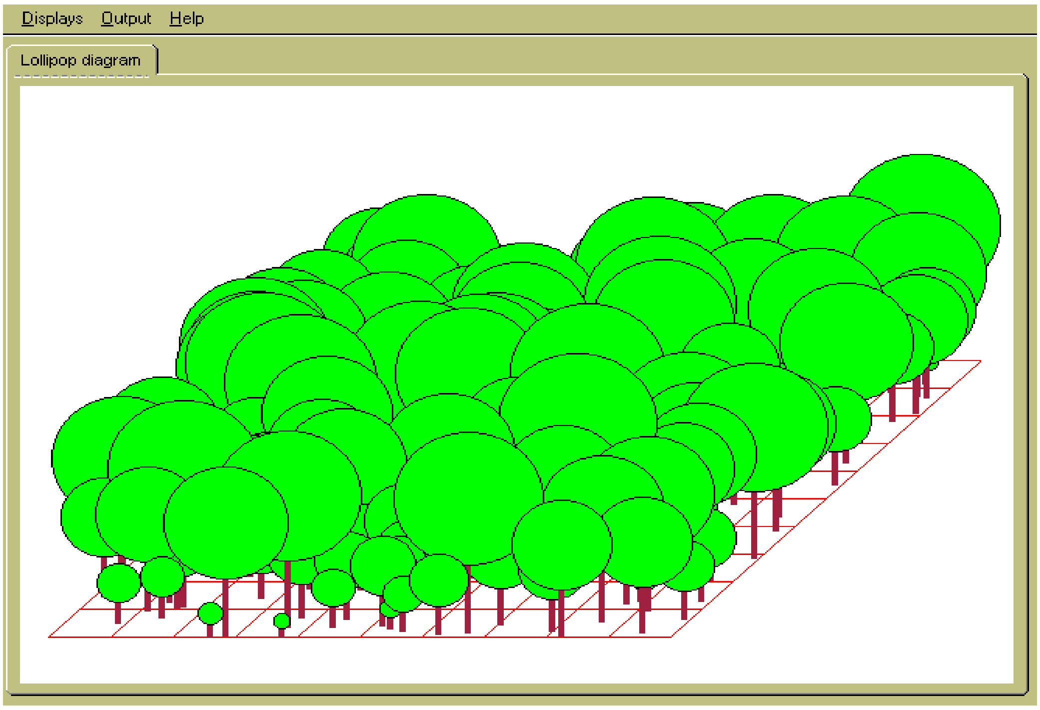
When the lollipop diagram is selected, the x- and y-coordinates must be chosen. Also, the object to be displayed, which is diameter in this forestry example, must be indicated.



Simile will draw a picture of a forest in the lollipop window, with no trees because it has not been initialized. Maximise the lollipop window.



When the 'Start' button in the Run control box within the Simile Run Environment window is pressed, trees start to grow. In the mean time, some trees die, and new instances are created. When the 'Stop' button is pressed, Simile will now display a picture of the forest with the size of the trees at a given time. This forest simulation is an impressive example of what can be achieved. It could, however, be extended to have more than one tree species, to make mortality size dependent, to make the growth rate dependent on competition as well as tree size, and so on. But this demonstration has been sufficient to formalize and communicate some ideas about tree growth, and to test whether the ideas are reasonable, in both the logical and empirical sense.



4. CONCLUDING COMMENTS

Simile is a powerful medium for model development and implementation as a computer program. The example model developed here has demonstrated various features on Simile, including submodels, flows, populations, and birth and death.

This model took just one minute to build, a little less to run, and a little more to document. Hopefully, this example has illustrates some of the capabilities of Simile, and demonstrated that, using this powerful model and program development medium, models with considerable functionality do not necessarily need to be big, complicated, expensive or excessively demanding on data.

11. Introduction to Discounted Cash Flow Analysis and Financial Functions in Excel

John Herbohn and Steve Harrison

The financial and economic analysis of investment projects is typically carried out using the technique of discounted cash flow (DCF) analysis. This module introduces concepts of discounting and DCF analysis for the derivation of project performance criteria such as net present value (NPV), internal rate of return (IRR) and benefit to cost (B/C) ratios. These concepts and criteria are introduced with respect to a simple example, for which calculations using MicroSoft Excel are demonstrated.

1. CASH FLOWS, COMPOUNDING AND DISCOUNTING

Discounted Cash Flow (DCF) analysis is the technique used to derive economic and financial performance criteria for investment projects. It is important to review some of the basic concepts of DCF analysis before proceeding to topics such as cost0benefit analysis (CBA), financial analysis (FA), linear programming and the estimation of non-market benefits.

Cash flow analysis is simply the process of identifying and categories of cash flows associated with a project or proposed course of action, and making estimates of their values. For example, when considering establishment of a forestry plantation, this would involve identifying and making estimates of the cash outflows associated with establishing the trees (e.g. the cost of buying or leasing the land, purchasing seedlings, and planting the seedlings), maintaining the plantation (such as cost of fertilizer, labour, pruning and thinning) and harvesting. As well, it would be necessary to estimate the cash inflows from the plantation through sales of thinnings and timber at final harvest.

Discounted cash flow analysis is an extension of simple cash flow analysis and takes into account the time value of money and the risks of investing in a project. A number of criteria are used in DCF to estimate project performance including Net Present Value (NPV), Internal Rate of Return (IRR) and Benefit-Cost (B/C) ratios.

Before discussing criteria to measure project performance, it is necessary to introduce some concepts and procedures with respect to compounding and discounting. Let us begin with the concepts of simple and compound interest. For the moment, consider the interest rate as the cost of capital for the project.

Suppose a person has to choose between receiving $1000 now, or a guaranteed $1000 in 12 months time. A rational person will naturally choose the former, because during the intervening period he or she could use the $1000 for profitable investment (e.g. earning interest in the bank) or desired consumption. If the $1000 were invested at an annual interest rate of 8%, then over the year it would earn $80 in interest. That is, a principal of $1000 invested for one year at an interest rate of 8% would have a future value of $1000 (1.08) or $1080.

The $1000 may be invested for a second year, in which case it will earn further interest. If the interest again accrues on the principal of $1000 only, it is known as *simple interest*. In this case the *future value* after two years will be $1160. On the other hand, if interest in the second year accrues on the whole $1080, known as *compound interest*, the future value will be $1080 (1.08) or $1166.40. Most investment and borrowing situations involve compound interest, although the timing of interest payments may be such that all interest is paid before further interest accrues.

The future value of the $1000 after two years may also be derived as

$$\$1000\ (1.08)^2 = \$1166.40$$

In general, the future value of an amount \$a, invested for n years at an interest rate of i, is $\$a\ (1+i)^n$, where it it to be noted that the interest rate i is expressed as a decimal (e.g. 0.08 and not 8 for an 8% rate).

The reverse of compounding – finding the present-day equivalent to a future sum – is known as *discounting*. Because \$1000 invested for one year at an interest rate of 8% would have a value of \$1080 in one year, the *present value* of \$1080 one year from now, when the interest rate is 8%, is \$1080/1.08 = \$1000. Similarly, the present value of \$1000 to be received one year from now, when the interest rate is 8%, is

$$\$1000/1.08 = \$925.93$$

In general, if an amount \$a is to be received after n years, and the annual interest rate is i, then the present value is

$$\$a\ /\ (1+i)^n$$

The above discussion has been in terms of amounts in a single year. Investments usually incur costs and generate income in each of a number of years. Suppose the amount of \$1000 is to be received at the end of each of the next four years. If not discounted, the sum of these amounts would be \$4000. But suppose the interest rate is 8%. What is the present value of this stream of amounts? This is obtained by discounting the amount at the end of each year by the appropriate discount factor then summing:

$$\$1000/1.08 + \$1000/1.08^2 + \$1000/1.08^3 + \$1000/1.08^4$$

$$= \$1000/1.08 + \$1000/1.1664 + \$1000/1.2597 + \$1000/1.3605$$

$$= \$925.93 + \$857.34 + \$793.83 + \$735.03$$

$$= \$3312.13$$

The *discount factors* – $1/1.08^t$ for t = 1 to 4 – may be calculated for each year or read from published tables. It is to be noted that the present value of the annual amounts is progressively reduced for each year further into the future (from \$925.93 after one year to \$735.03 after four years), and the sum is approximately \$700 less than if no discounting (a zero discount rate) had been applied.

2. DEFINITION OF ANNUAL NET CASH FLOWS

DCF analysis is applied to the evaluation of investment *projects*. Such a project may involve creation of a terminating asset (such as a forestry plantation), infrastructure (such as a road or plywood plant) or research (including scientific and socio-economic research). Any project may be regarded as generating cash flows. The term *cash flow* refers to any movement of money to or away from an investor (an individual, firm, industry or government). Projects require payments in the form of capital outlays and annual operating costs, referred to as cash outflows. They give rise to receipts or revenues, referred to as project benefits or cash inflows. For each year, the difference between project benefits and capital plus operating costs is known as the net cash flow for that year. The net cash flow in any year may be defined as

$$a_t = b_t - (k_t + c_t)$$

where b_t are project benefits in year t
k_t are capital outlays in year t
c_t are operating costs in year t.

It is to be noted that when determining these net cash flows, expenditure items and income items are timed for the point at which the transactions takes place, rather than the time at which they are used. Thus for example expenditure on purchase of an item of machinery rather than annual allowances for depreciation would enter the cash flows. It is to be further noted that cash flows should not include interest payments. The discounting procedure in a sense simulates interest payments, so to include these in the operating costs would be to double-count them.

Example 1

A project involves an immediate outlay of \$25,000, with annual expenditures in each of three years of \$4000, and generates revenue in each of three years of \$15,000. These cash flows data may be set out, and annual net cash flows derived, as in Table 1.

Table 1. Annual cash flows for a hypothetical project

Year	Project benefits ($)	Capital outlays ($)	Operating costs ($)	Net cash flow ($)
0	0	25000	2000	-27000
1	15000	0	4000	11000
2	15000	0	4000	11000
3	15000	0	2000	13000

Two points may be noted about these cash flows. First, the capital outlay is timed for Year 0. By convention this is the beginning of the first year (i.e. right now). On the other hand, only half of the first year's operating costs are scheduled for the Year 0 (the beginning of the first year). The remaining half of the first year's operating costs plus the first half of the second year's operating costs are scheduled for the end of the first year (or, equivalently, the beginning of the second year). In this way, operating costs are spread equally between the beginning and the end of each year. (The final half of the third year's operating costs are scheduled for the end of Year 3.) In the case of project benefits, these are assumed to accrue at the end of each year, which would be consistent with lags in production or payments. These within-year timing issues are unlikely to make a large difference to overall project profitability, but it is useful to make these timing assumptions clear.

A second point to note about Table 1 is that net cash flows (second column less third plus fourth column) are at first negative, but then become positive and increase over time. This is a typical pattern of well-behaved cash flows, for which performance criteria can usually be derived without computational difficulties.

3. PROJECT PERFORMANCE CRITERIA

Let us now consider a number of project performance criteria which can be obtained by discounted cash flow analysis. These criteria will be defined, then derived for the cash flow data of Example 1.

Net present value

The net present value (NPV) is the sum of the discounted annual cash flows. For the example, taking an interest rate of 8%, this is

$$NPV = a_0 + a_1/(1+i) + a_2/(1+i)^2 + a_3/(1+i)^3$$

A project is regarded as economically desirable if the NPV is positive. The project can then bear the cost of capital (the interest rate) and still leave a surplus or profit. For the example,

$$NPV = -27000 + 11000/(1.08) + 11000/(1.08)^2 + 13000/(1.08)^3$$

$$= -27000 + 11000/1.1664 + 11000/1.2597 + 13000/1.3605$$

$$= \$2935.73$$

The interpretation of this figure is that the project can suppport an 8% interest rate and still generate a surplus of benefits over costs, after allowing for timing differences in these, of approximately $3000.

Net future value

An alternative to the net present value is the net future value (NFV), for which annual cash flows are compounded forward to their value at the end of the project's life. Once the NPV is known, the NFV may be obtained indirectly by compounding forward the NPV by the number of years of the project life. For Example 1, the net future value is

$$NFV = NPV\ (1.08)^3 = \$2935.73 \times 1.3605 = 3994.06$$

Internal rate of return

The internal rate of return (IRR) is the interest rate such that the discounted sum

of net cash flows is zero. If the interest rate were equal to the IRR, the net present value would be exactly zero. The IRR cannot be determined by an algebraic formula, but rather has to be approximated by trial and error methods. For the above example, we know that the IRR is somewhere above 8%. Deriving the NPV with a range of discount rates would reveal that the IRR falls between 13% and 14%, but closer to the latter. In practice, a financial function can be called up to perform the trial-and-error calculations. It would be found in this case that the IRR is about 13.8%.[1]

The IRR is the highest interest rate which the project can support and still break even. A project is judged to be worthwhile in economic terms if the internal rate of return is greater than the cost of capital. If this is the case, the project could have supported a higher rate of interest than was actually experienced, and still made a positive payoff. In the above case, the project would be profitable provided the cost of capital was less than 13.8%.

The IRR as a criterion of project profitability suffers from a number of theoretical and practical limitations. On the theoretical side, it assumes that the same rate of return is appropriate when the project is in surplus and when it is in deficit. However, the cost of borrowed funds may be quite different to the earning rate of the firm. It could be more appropriate to use two rates when determining the IRR. The actual cost of capital could be used when the project is in deficit, and the earning rate (unknown, to be determined by trial-and-error) could be applied when the project is in surplus. This would give a better indication of the earning rate of the project to the firm or government.

From a practical viewpoint, the IRR may not exist or it may not be unique. This problem may be examined in terms of the *NPV profile*, a graph of NPV versus the rate of interest. When the IRR is well behaved, this profile takes the form as in Figure 1. As the interest rate increases the NPV falls, being zero where the NPV curve crosses the interest rate axis; the IRR corresponds with this discount rate.

Consider a project for which the net cash flow in each year (including Year 0) is positive. Regardless of the interest rate, the NPV will never be zero, so it will not be possible to determine an IRR. Similarly, a project with a large initial capital outlay and for which future benefits are relatively small or negative may not have a positive NPV regardless of the interest rate, so again the curve for the NPV profile may never cross the interest rate axis.

If a project generates runs of positive and negative net cash flows, the NPV profile may take the form of a roller-coaster curve, crossing the interest rate axis in several places.[2] This indicates multiple internal rates of return, one at each interest rate where NPV is zero. It is then by no means clear which if any of the rates we should choose to call the IRR. Further, for some sections of the NPV profile (those that are upward sloping), the NPV is increasing as the interest rate increases. This implies that the greater the cost of capital the more profitable the project. Clearly, multiple internal rates of return and perverse relationships between the NPV and the discount rate are not very satisfactory.

Benefit-to-cost ratios

A number of benefit-cost ratio concepts have been developed. For simplicity, we will consider only two concepts, referred to as the gross and net B/C ratio and defined respectively as

Gross B/C ratio = PV of benefits/(PV of capital costs + PV of operating costs)

Net B/C ratio = (PV of benefits – PV of operating costs)/PV of capital costs

[1] Calculation of internal rate of return in fact involves solving a polynomial equation, and efficient solution methods such as Newton's approximation are used in computer packages.

[2] Mathematically, the polynomial equation defining the NPV can have up as many 'roots' or solutions as there are turning points in NPV values (changes from positive to negative or negative to positive). A project which has alternating runs of positive and negative cash flows is a candidate for problems with estimation of the IRR.

For the above project, the present value of capital outlays is $25,000, since outlays are made immediately and as a single amount. The present values of project benefits and operating costs are:

PV of benefits = $\$15000/1.08 + \$15000/1.08^2 + \$15000/1.08^3 = \$38656.45$

PV of operating costs = $\$2000 + \$4000/1.08 + \$4000/1.08^2 + \$2000/1.08^3 = \$9418.38$

Hence the benefit-to-cost ratios are

$$\text{gross B/C ratio} = \frac{38{,}656.45}{25{,}000 + 9418.38} = 1.12$$

$$\text{net B/C ratio} = \frac{38{,}656.45 - 9418.38}{25{,}000} = 1.45$$

A project is judged to be worthwhile in economic terms if it has a B/C ratio is greater than unity, i.e. if the present value of benefits exceeds the present value of costs (in gross or net terms). If one of the above ratios is greater than unity, then the other will be greater than unity also. In the above example, the ratios are greater than unity, indicating that the project is worthwhile on economic grounds. It is not clear on logical grounds which of the ratios is the most useful.

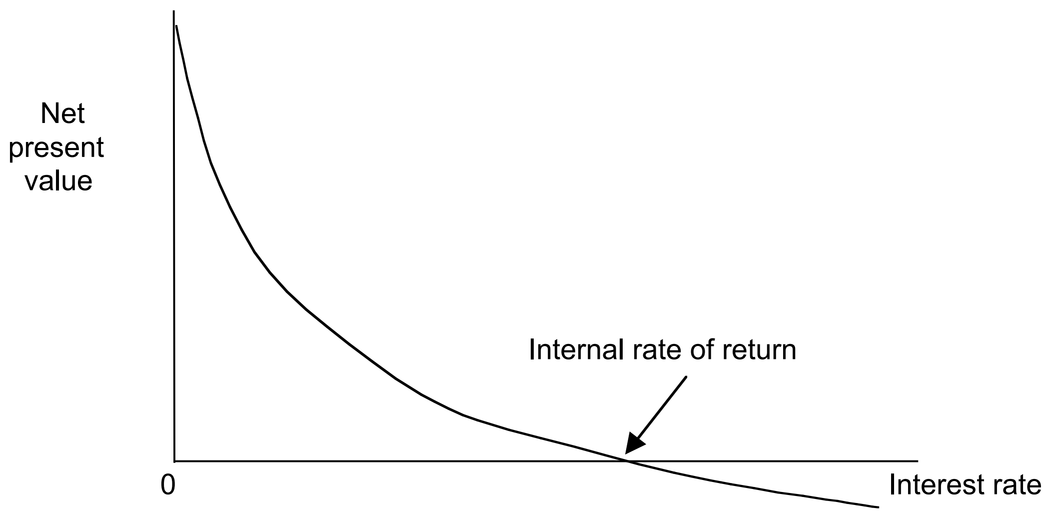


Figure 1. The NPV profile for a project

The payback period

The payback period (PP) is the number of years for the projects to break even, i.e. the number of years for which discounted annual net cash flows must be summed before the sum becomes positive (and remains positive for the remainder of the project's life). The payback period for a project with the above net cash flows can be determined as in the following table. From this table, it is apparent that the sum of discounted net cash flows does not become positive until Year 3, so the payback period is three years.

The payback period indicates the number of years until the investment in a project is recovered. It is a useful criterion for a firm with a short planning horizon, but does not take account of all the information available, i.e. the net cash flows for years beyond the payback period.

The peak deficit

This is a measure of the greatest amount that the project 'owes' the firm or government, i.e. the furtherest 'in the red' it goes. In the above table, the largest negative value is -$27,000, so this is the peak deficit. Peak deficit is a useful measure in terms of financing a project, since it indicates the total amount of finance that will be required.

Table 2. Derivation of 'project balances' and payback period

Year	Net cash flow ($)	PV of net cash flow ($)	Cumulative PV of net cash flow (or project balance) ($)
0	-27000	-27000.00	-27000
1	11000	10185.19	-27000 + 10185.19 = -16814.81
2	11000	9430.73	-16814.81 + 9340.73 = -7384.09
3	13000	10319.82	-7384.09 + 10319.82 = 2935.73

Review of DCF performance criteria

The most commonly used discounted cash flow performance criteria – NPV, IRR, B/C ratios and payback period – may be summarised as in Table 3. The various criteria are closely related, but measure slightly different things. In this respect, they tend to complement one another, so that it is common to estimate and report more than one of the measures.

Perhaps the most useful measure, and the one most often reported, is the net present value. This tells the total payoff from a project. A limitation of the NPV is that it is not related to the size of the project. If one project has a slightly lower NPV than another, but the capital outlays required are much lower, then the second project will probably be the preferred one. In this sense, a rate of return measure such as the IRR is also useful.

The payback period and peak deficit are useful supplementary project information for decision-makers. They have greater relevance for private sector investments, for firms which cannot afford long delays in recouping expenditure, and where careful attention must be paid to the total amount of funds that will need to be committed to the project to remain solvent.

Table 3. Summary of definitions of main DFC performance criteria

Net present value (NPV)	$\sum_{t=1}^{p} a_t / (1+i)^t$, where t is time, c_t is the annual net cash flow, *i* is the discount rate p is the planning horizon
Internal rate of return (IRR)	The value of r such that $\sum_{t=1}^{p} a_t / (1+r)^t = 0$
Benefit to cost ratio (B/C)	Present value of project benefits / present value of project costs
Payback period	Number of periods until NPV becomes (and remains) positive

4. USING EXCEL FINANCIAL FUNCTIONS

It is common to use a spreadsheet program such as Excel to undertake financial and cost benefit analyses. Excel has a range of financial functions that can be useful (as listed in Apppendix 1). Two of the most useful and commonly used financial functions in Excel are those to calculate Net Present Value ('NPV') and Internal Rate of Return ('IRR'). This section illustrates the use of two of the NPV and IRR functions, applied to the data in Example 1, in an Excel worksheet:

Using the 'NPV' and 'IRR' functions

The 'NPV' function calculates the net present value of an investment (i.e. series of cash flows) by using a discount rate and a series of future payments (negative

values) and income (positive values). The syntax for this function is:

NPV(discount rate, value1, value2 ….)

It is important to remember that the NPV calculation is for cashflows starting at the end of period 1. If the first cash flow occurs at the beginning of period 1 (i.e. period 0), the first value must be added to the NPV result, not included in the value 'arguments' of the function.

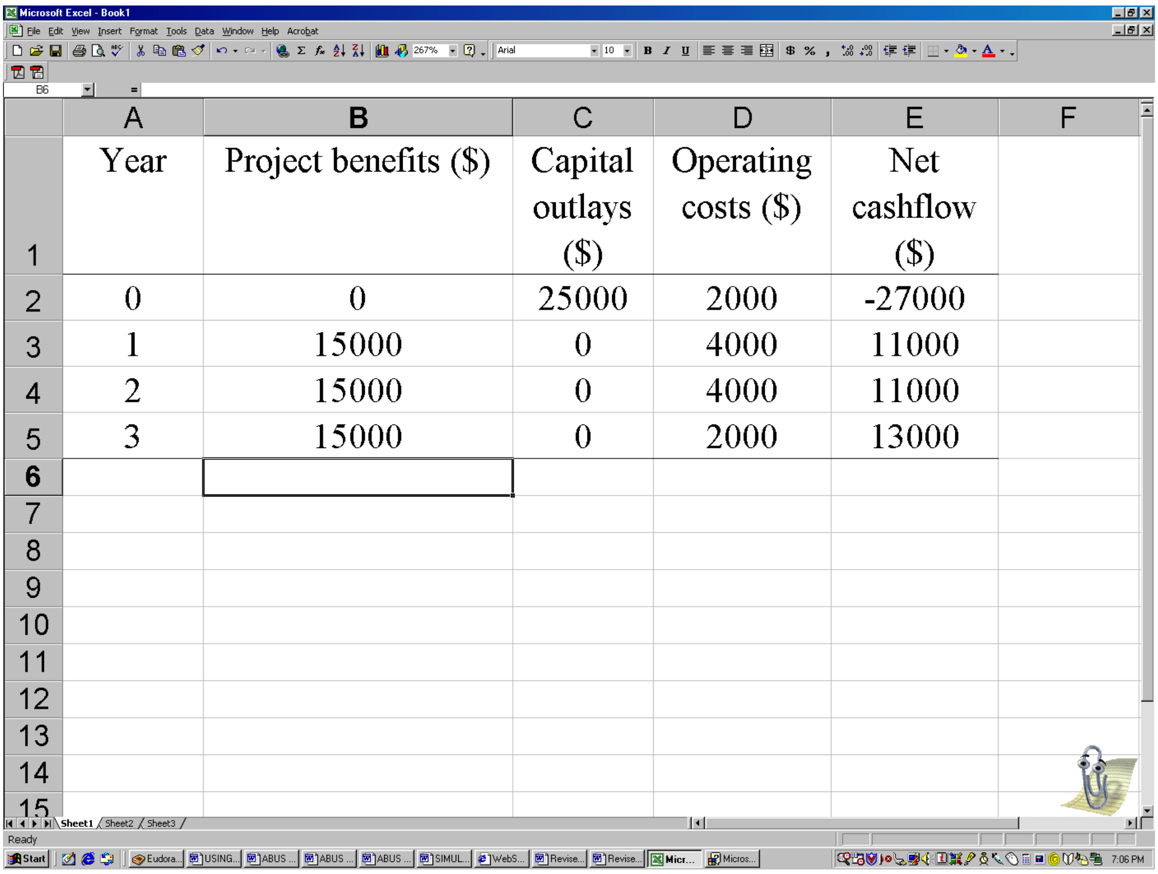
Thus to value the NPV for example 1, the following formula would be inserted into a spreadsheet cell:

=NPV(8%, 11000, 11000, 13000) - 27000

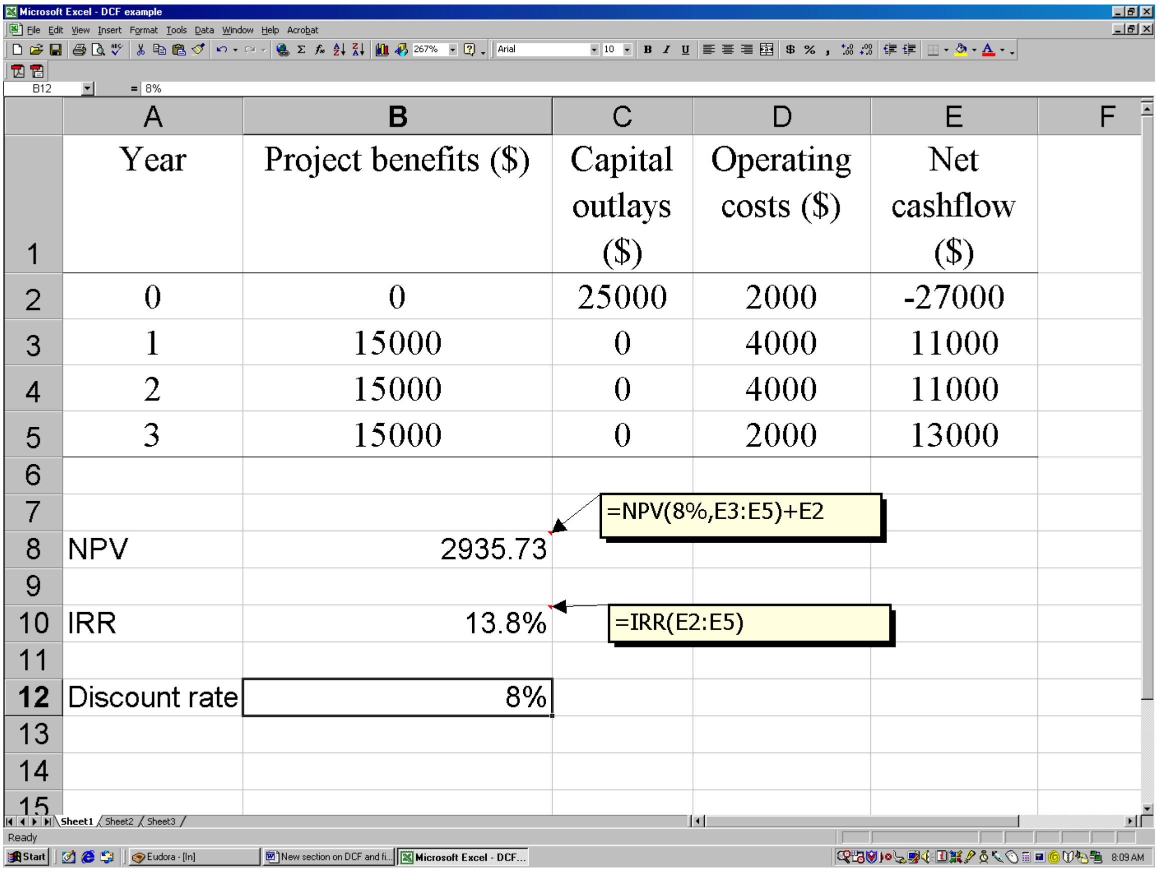
Note that the cashflow sequence used in the NPV calculation is for years 1 to 3. The value for year 0 (-$27000) is added to the NPV result.

An alternative means of calculating NPV and IRR in the above example would be to replace the string of values with cell references to where the values are located within the workbook. This is illustrated in the worksheet below. This approach has the advantage of allowing changes to the made to the cashflows (in Column E) which then automatically flow through to the calculation of the NPV. This is particularly useful when it is desired to investigate how changes in cash flows affect NPV. Similarly, the discount rate given in the formula (in this case 8%) could be replaced with a cell reference. In this case the effects of a change in discount rate on NPV could be observed simply changing the cell value rather than the formula. In the worksheet below, the '8%' in the formula for NPV in cell B8 would be replaced by a reference to cell B12. The new formula would be '=NPV(B12, E3:E5)+E2'.

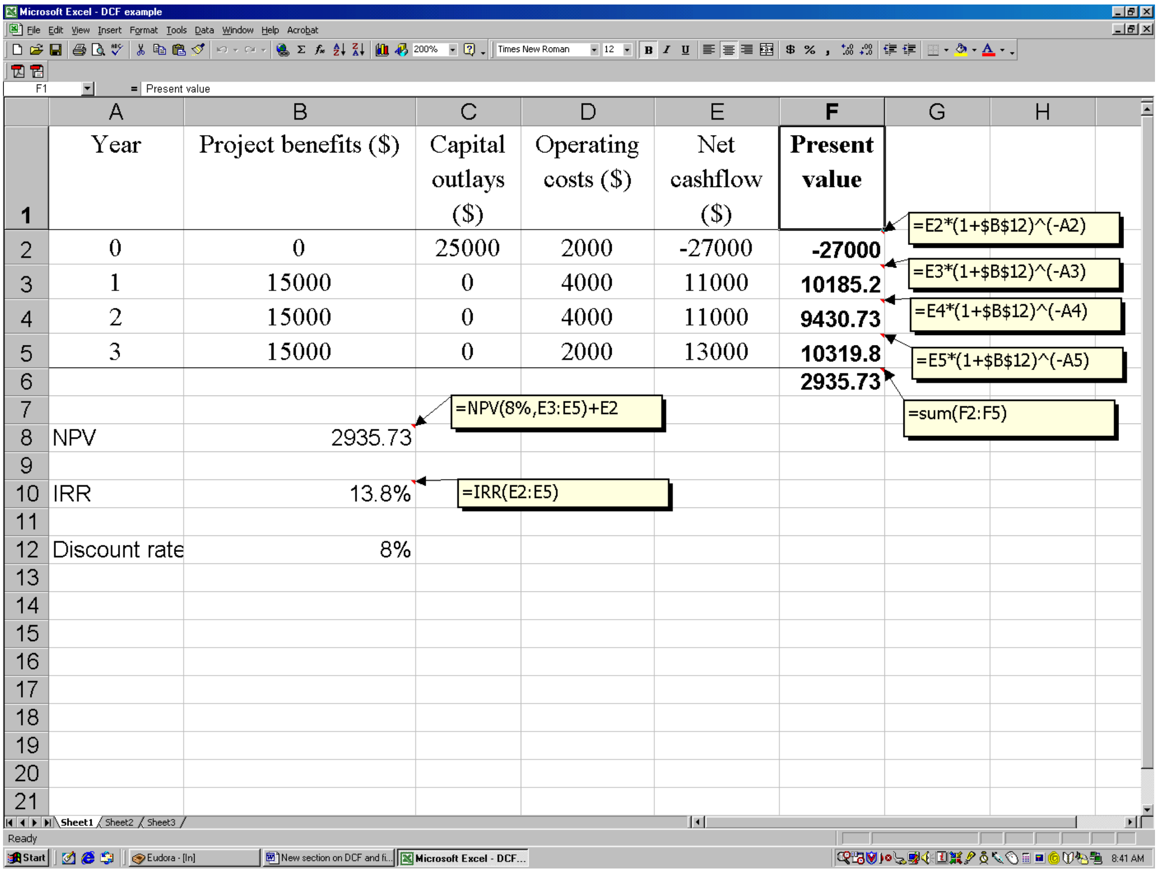
Another approach to calculating NPV is to simply calculate the present value of cash flows for each year using the discounting formula given in Equation 1 and then simply add these individual figures. This approach has a number of advantages in more complex applications – particularly those involved complex financial models with variable periods of cash flows and where discount rates change over the life of a project. This is illustrated in the workbook below:



	A	B	C	D	E	F
1	Year	Project benefits ($)	Capital outlays ($)	Operating costs ($)	Net cashflow ($)	
2	0	0	25000	2000	-27000	
3	1	15000	0	4000	11000	
4	2	15000	0	4000	11000	
5	3	15000	0	2000	13000	



	A	B	C	D	E
1	Year	Project benefits ($)	Capital outlays ($)	Operating costs ($)	Net cashflow ($)
2	0	0	25000	2000	-27000
3	1	15000	0	4000	11000
4	2	15000	0	4000	11000
5	3	15000	0	2000	13000
8	NPV	2935.73	=NPV(8%,E3:E5)+E2		
10	IRR	13.8%	=IRR(E2:E5)		
12	Discount rate	8%			



	A	B	C	D	E	F	G
1	Year	Project benefits ($)	Capital outlays ($)	Operating costs ($)	Net cashflow ($)	**Present value**	
2	0	0	25000	2000	-27000	**-27000**	=E2*(1+$B$12)^(-A2)
3	1	15000	0	4000	11000	**10185.2**	=E3*(1+$B$12)^(-A3)
4	2	15000	0	4000	11000	**9430.73**	=E4*(1+$B$12)^(-A4)
5	3	15000	0	2000	13000	**10319.8**	=E5*(1+$B$12)^(-A5)
6						**2935.73**	=sum(F2:F5)
8	NPV	2935.73	=NPV(8%,E3:E5)+E2				
10	IRR	13.8%	=IRR(E2:E5)				
12	Discount rate	8%					

The formula in Cells F2, F3, F4 and F5 calculate the present value of the net cash flow in each of the years 0 through to 3 using the formula given in Equation 1. These present values for each of the years are then summed in cell F6 to give the Net Present Value of the series of cash flows associated with the project. The resulting NPV in cell F6 is identical to the NPV calculated in cell B8. Note that an absolute cell address for the discount rate has been used ($B$12) using '$'. This allows the formula in cell F2 to be simply copied into cells F3, F4 and F5.

APPENDIX 1: FINANCIAL FUNCTIONS IN EXCEL

ACCRINT Returns the accrued interest for a security that pays periodic interest
ACCRINTM Returns the accrued interest for a security that pays interest at maturity
AMORDEGRC Returns the depreciation for each accounting period
AMORLINC Returns the depreciation for each accounting period
COUPDAYBS Returns the number of days from the beginning of the coupon period to the settlement date
COUPDAYS Returns the number of days in the coupon period that contains the settlement date
COUPDAYSNC Returns the number of days from the settlement date to the next coupon date
COUPNCD Returns the next coupon date after the settlement date
COUPNUM Returns the number of coupons payable between the settlement date and maturity date
COUPPCD Returns the previous coupon date before the settlement date
CUMIPMT Returns the cumulative interest paid between two periods
CUMPRINC Returns the cumulative principal paid on a loan between two periods
DB Returns the depreciation of an asset for a specified period using the fixed-declining balance method
DDB Returns the depreciation of an asset for a specified period using the double-declining balance method or some other depreciation method
DISC Returns the discount rate for a security
DOLLARDE Converts a dollar price, expressed as a fraction, into a dollar price, expressed as a decimal number
DOLLARFR Converts a dollar price, expressed as a decimal number, into a dollar price, expressed as a fraction
DURATION Returns the annual duration of a security with periodic interest payments
EFFECT Returns the effective annual interest rate
FV Returns the future value of an investment
FVSCHEDULE Returns the future value of an initial principal after applying a series of compound interest rates
INTRATE Returns the interest rate for a fully invested security
IPMT Returns the interest payment for an investment for a given period
IRR Returns the internal rate of return for a series of cash flows
ISPMT Calculates the interest paid during a specific period of an investment.
MDURATION Returns the Macauley modified duration for a security with an assumed par value of $100
MIRR Returns the internal rate of return where positive and negative cash flows are financed at different rates
NOMINAL Returns the annual nominal interest rate
NPER Returns the number of periods for an investment
NPV Returns the net present value of an investment based on a series of periodic cash flows and a discount rate
ODDFPRICE Returns the price per $100 face value of a security with an odd first period
ODDFYIELD Returns the yield of a security with an odd first period
ODDLPRICE Returns the price per $100 face value of a security with an odd last period
ODDLYIELD Returns the yield of a security with an odd last period
PMT Returns the periodic payment for an annuity
PPMT Returns the payment on the principal for an investment for a given period
PRICE Returns the price per $100 face value of a security that pays periodic interest
PRICEDISC Returns the price per $100 face value of a discounted security

PRICEMAT Returns the price per $100 face value of a security that pays interest at maturity

PV Returns the present value of an investment

RATE Returns the interest rate per period of an annuity

RECEIVED Returns the amount received at maturity for a fully invested security

SLN Returns the straight-line depreciation of an asset for one period

SYD Returns the sum-of-years' digits depreciation of an asset for a specified period

TBILLEQ Returns the bond-equivalent yield for a Treasury bill

TBILLPRICE Returns the price per $100 face value for a Treasury bill

TBILLYIELD Returns the yield for a Treasury bill

VDB Returns the depreciation of an asset for a specified or partial period using a declining balance method

XIRR Returns the internal rate of return for a schedule of cash flows that is not necessarily periodic

XNPV Returns the net present value for a schedule of cash flows that is not necessarily periodic

YIELD Returns the yield on a security that pays periodic interest

YIELDDISC Returns the annual yield for a discounted security. For example, a Treasury bill

YIELDMAT Returns the annual yield of a security that pays interest at maturity

12. Financial Models of Farm and Community Forestry Production Systems

John Herbohn, Nick Emtage and Steve Harrison

An integral component of investment in forestry is the need for financial information about the likely cash flows associated with the establishment, management and final harvest of a plantation. This paper uses the development of a financial model for forestry investment as a case study of financial modeling. Financial evaluation of forestry projects poses many challenges and this chapter includes discussion of the methods employed in model development. The techniques of model development are illustrated by case studies of the development of the Australian Cabinet Timber Financial Model (ACTFM) and the Australian Farm Forestry Financial Model (AFFFM).

1. FORESTRY EVALUATION MODELS: USES AND USER GROUPS

The traditional use of financial models is to assess simply the financial viability of a proposed investment. As well as deriving single point estimates of financial performance criteria, sensitivity analysis and risk simulation may be performed to explore the impact of variations from the core assumptions. Models can also be used to explore investment alternatives. For instance forestry models applied to joint ventures can be used to explore the impact on cash flows of different equity participation arrangements. Large pastoral companies can also use financial models to explore the impact that forestry may have on their overall business risk. The way in which a model is to be used is a crucial consideration in design of the model.

Forestry financial models can potentially be used by a number of different groups and for a number of different purposes. For instance, a company which frequently promotes forestry schemes to investors would require a model that could be easily adapted to new areas and different tree species. Similarly, a government department promoting joint venture arrangements involving only one or two species but which involve different equity sharing arrangements with private landholders would need a model that was flexible in different way. In both cases, it would be inefficient to develop a new spreadsheet-based model for each new situation.

Globally, forestry is a major economic activity at the farm level and can be a significant contributor to gross domestic product (GDP), particularly in many European countries. Farm operations are however highly diverse, in terms of area planted (1000s of hectares down to very small plots), species chosen, management regime, harvest age, and other variables. Forestry is typically only part of the overall business undertaking of the landowner, and must be considered in this context.

Farmers often lack the skills to develop financial models capable of modelling their activities. The provision of models to be used as a tool to undertake this type of analysis is often done as a form of assistance or extension service by governments. In such cases, models need to be highly flexible and take into account the impact that forestry has on the overall operations of the farm, particularly cash flows. The impact that forestry has on the farm financial structure and cash flows can be critical. In such cases, a project with a positive NPV may be rejected because it cannot be accommodated within the financial structure of the business (i.e. the question of how, if at all, the project can be financed becomes critical).

Forestry financial models are also useful for land valuers, real estate agents and lending agencies, all of which often have a need to place a value on immature forests when estimating market values or evaluating loan proposals for rural properties. They further have a role in training foresters and as a

teaching device.

2. KEY PARAMETERS FOR FORESTRY MODELS

The most common cash outflows associated with forestry operations include: establishment costs (e.g. land and site preparation such as clearing, fencing and ripping; purchase of planting stock; planting, mulching and watering); weed control; fertilization; pruning and thinning to waste, and commercial harvesting. The major cash inflows come from sale of commercial thinnings and final harvest, and perhaps tree residues such as firewood. In some situations revenues are also derived from non-wood forest products such as wild berries, mushrooms, honey and hunting, all of which are made possible by the environment created by the forest.

Cash flow estimates of growing native timbers as a business enterprise will vary depending on a variety of factors, some of the more important of which are listed in Table 1. In principle, a spreadsheet could be devised which incorporates each of these factors. One factor that can have a major impact on eventual cash flows is timber yield from plantations. Yield is commonly expressed as 'mean annual increment' (MAI) which is the aggregate volume of harvestable timber produced in a year from the growth of trees in a plantation. It is usually expressed as cubic metres of timber produced per ha per year. Other factors, such as site productivity, species mixture, harvest ages and timber prices (usually expressed as $ per cubic metre) may also strongly affect eventual cashflows.

Table 1. Key factors in financial performance

Site characteristics (e.g. climate, soil, aspect)
Species mixture
Silvicultural system (e.g. planting density, weed control and pruning)
Harvest age and harvest scheduling
Final yield or MAI of individual species
Interactions between species in mixed-species plantations
Stumpage price (affected by a number of supply and demand factors)
Costs (land preparation, planting and establishment, maintenance, harvesting)
Amount of government assistance
Harvest rights (buffer zones, harvest on steep land, roading permission)
Taxation regime (deductions allowable, treatment of harvest revenue)
Allowance for non-wood benefits
Discount rate

3. SOURCES OF VARIABILITY IN FORESTRY INVESTMENT PERFORMANCE

Plantation forestry is not a risk-free investment. A variety of factors can contribute to the uncertainly in revenue generated. Most cash inflows from plantations occur with the final harvest of trees when they reach maturity, although some additional cash inflows may be generated from the sale of trees that are thinned as part of the normal management practices. Typically it takes 30 or more years until plantations are ready for final harvest trees, hence long-term predictions of physical and financial performance have to be made. The main sources of variability in financial performance of forestry are summarised in Table 2.

4. DEVELOPING A FINANCIAL MODEL – A STEP BY STEP APPROACH

As with any project appraisal, undertaking a financial analysis of a forestry project can be broken down into a number of individual and relatively simple steps:

1. identify the forestry system to be adopted – for example, the type of trees to be planted and at what density, when they will be pruned, the types of product intended (e.g. sawlogs, veneer logs, pulpwood, poles, or some combination

of these) and when harvesting is likely to occur.
2. estimate the likely cash outflows.
3. estimate the cash inflows from harvest.
4. develop the financial model and estimate financial performance measures.
5. evaluate investment risk.

Table 2. Sources of risk in farm forestry

Risk category	Major sources
Risk of poor establishment	Dry weather, poor weed control
Production (timber yield) risk	Storm or cyclone Fire Pest (insect, disease) Unsuitable species or mixture for the site Collateral damage at harvest
Timber quality and product type risk	Inappropriate pruning and thinning regime Insect damage Product type and fashion changes in demand
Sovereign risk	Regulatory changes re machinery use and roading Changes in taxation arrangements Uncertain harvest rights and compensation
Market risk	Uncertain future timber prices

5. REVIEW OF MODEL DEVELOPMENT AND DESIGN OPTIONS

Models may be developed on a one off basis – that is a unique model is developed for each investment being appraised. Alternatively a generic model can be developed which is flexible enough to evaluate a different investment scenarios. This is basically a question of cost-effectiveness. If a model is to be used repeatedly, then greater effort in model development is warranted. For these generic models to be used by a variety of stakeholders, most of whom have little understanding of the mechanics of the model, the user interface or series of screens for data input and output has proved to be a critical design factor. Also, ease of navigation around the model is critical, and a convenient way to provide this is through button bars which are 'clickable' with the mouse.

What are the options?

It may be that a generic model is already available, which has been developed elsewhere. Regardless of whether this is 'freeware' or moderately expensive to acquire, it may be in the best interests of a firm or individual to gain access to such a package rather than commit resources to developing their own model. Some of the important modeling options for forestry are outlined in Table 3.

Well-designed generic models can cope with a wide range of investment circumstances. However, it is not possible to predict all of the situations in which people may wish to use a model. In the design of a generic model, it is critical to give consideration to the degree of flexibility that is required for changes to be made, and the availability of skilled modelers to make changes in the model to accommodate novel or unforeseen applications, and choice of computing platform to suit user needs. This can also be a major influence on the software used to develop the model.

Table 3. Modelling options for forestry investments

Characteristic of model	Options
Model complexity and resource demands	One-off vs generic
Model property rights	Developed by investor, off-the-shelf
Level of business activity	Forestry enterprise or whole-of-business
Level of flexibility	Narrow to wide range of applications
Modeling platform	Spreadsheet programs (e.g.Excel), programming languages (e.g. C^{++}, Visual Basic) or simulation languages (e.g. Stella, Simile)
Type of yield modeling	Growth curves, growth simulations using packages such as Plantgro, discrete harvest age and MAI
Source of input data	Default (trial-based, subjectively estimated, simulated) vs user provided
User interface	Menu system, button bars
Choice of project benefit categories	Wood only vs wider benefits (financial vs economic)
Representation of investment risk	RADR, sensitivity analysis, risk analysis
Choice of project performance variables	NPV vs LEV, IRR, not payback period
User support	None, access to technical assistance, package maintenance

What type of platform to use?

For a large and complex generic model, a spreadsheet platform may not be suitable. In such cases, it may become necessary to resort to programming languages such as Visual Basic or C^{++}. In other cases modeling environments such as Simile, which have been specifically developed for use in developing models for complex resource management decisions (as is often involved with forestry projects) may be appropriate.

Spreadsheets are a very convenient modeling environment in which to develop financial models. Excel for instance has a range of features that are extremely useful in model building including button bars, financial functions, the ability to easily record macros, and input forms. Many people are now reasonably proficient at using spreadsheets. As such, one of the major benefits of using spreadsheet models is that these models need far less technical support provided to users and can be easily modified. In contrast, models developed using a programming language such as Visual Basic or C++ can require significant ongoing technical support and require expert programmers for modifications to be made.

6. THE AUSTRALIAN CABINET TIMBERS FINANCIAL MODEL (ACTFM)

The Australian Cabinet Timber Financial Model has been developed as a flexible and user-friendly computer-based model suitable for assessing the financial viability of timber plantations using Australian native eucalypt and cabinet species in North Queensland. A schematic representation of the structure of the model is provided in Figure 1. Little published information exists on growth rates of many of these native tree species. To provide information for the model about the growth rates and optimal harvest ages for the 31 species included in the model, it was necessary to undertake a Delphi survey of experts on North Queensland forestry. Details of this study are provided in later sections.

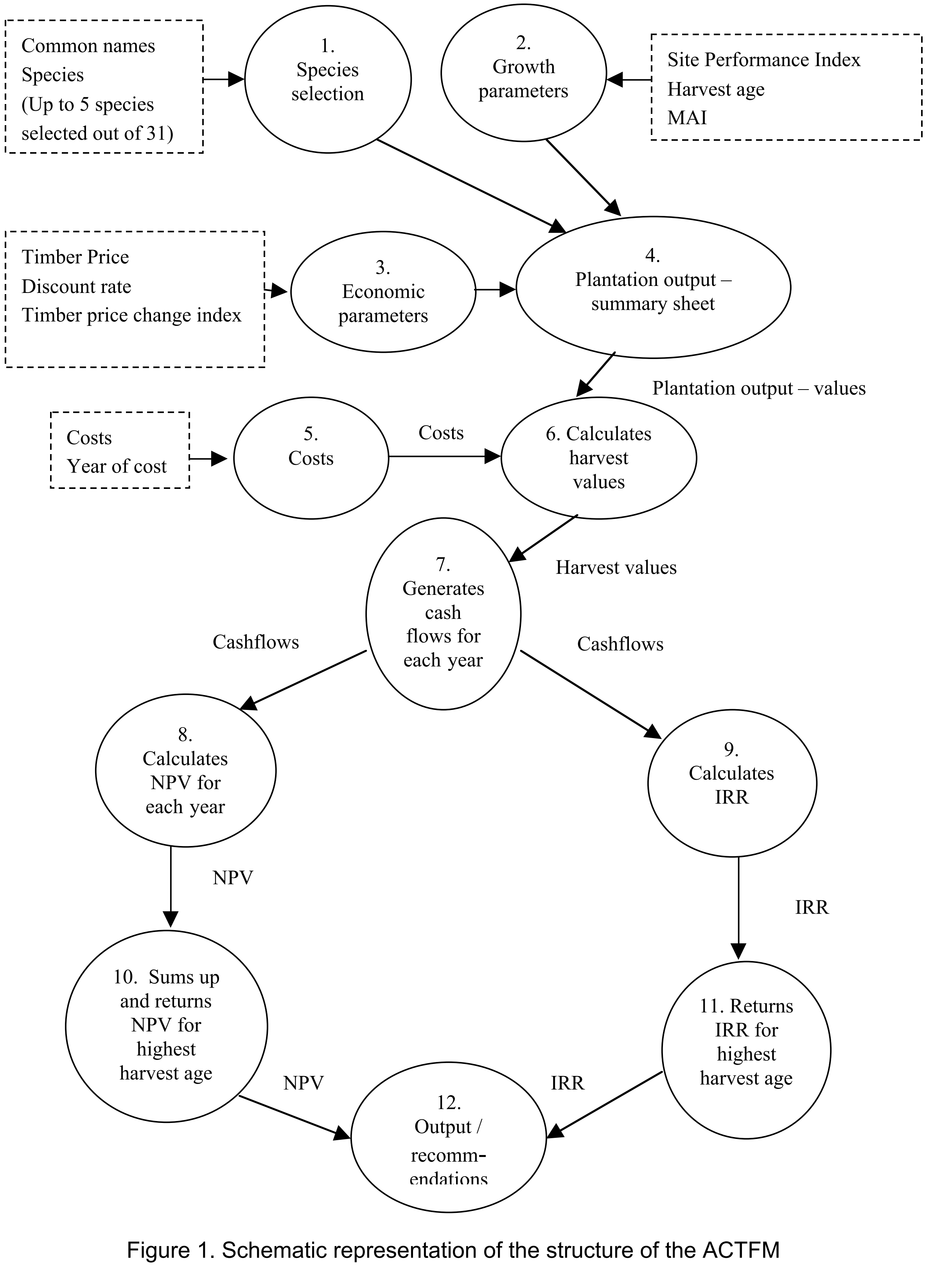


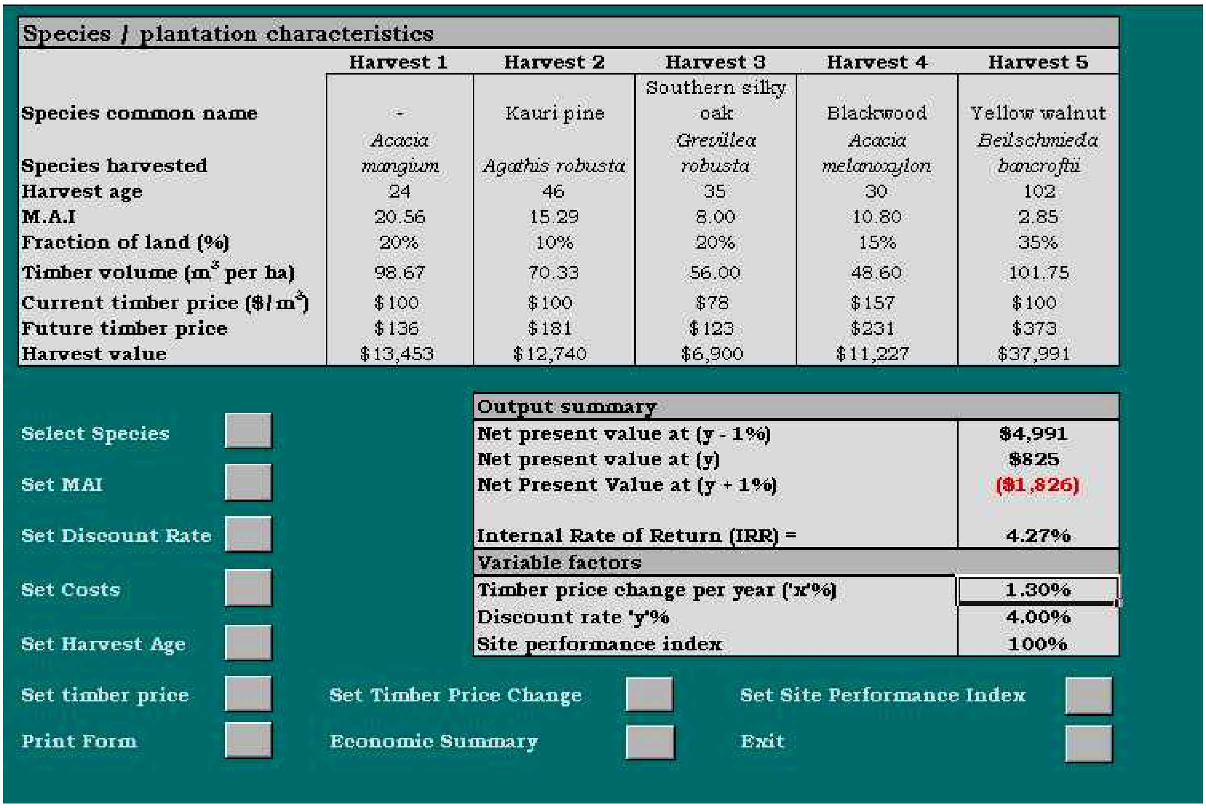
Figure 1. Schematic representation of the structure of the ACTFM

The model contains a number of linked sheets in an Excel workbook format. Sheets are used to store or display data or information about the program, and to perform calculations of the value of the plantation scenario established by the user. The model was developed on the Microsoft Excel (version 7.0) software and contains a collection of Visual Basic macros. It requires the Excel program to operate. To navigate between various data input and model output screens, a system of ‘button bars’ is provided. Notably, the model was

developed by researchers and not professional computer programmers.

The model has been set up with an attractive opening sheet, which indicates the model name, names of the developers and version of the model, and acknowledges the agencies providing financial support for model development. From the title page access is gained to the 'Plantation Output' summary sheet (Figure 2) by clicking on a 'Start' button. This in turn is linked to the various other sheets in the workbook which provide default data and instructions and other information to the user, to allow calculation of net present value and the internal rate of return. The menu system to allow the user to move freely between sheets in the workbooks consists of a series of 'buttons', on the user simply clicks the computer mouse. The various sheets provide:

- information about the program;
- data relating to MAI's, stumpage prices, and harvest ages of 30 different cabinet timber species;
- data relating to establishment and maintenance costs;
- tables for the entry of alternative MAI, price, costs and harvest age data; and
- pages containing visual basic code.



Species / plantation characteristics	Harvest 1	Harvest 2	Harvest 3	Harvest 4	Harvest 5
Species common name	-	Kauri pine	Southern silky oak	Blackwood	Yellow walnut
Species harvested	*Acacia mangium*	*Agathis robusta*	*Grevillea robusta*	*Acacia melanoxylon*	*Beilschmieda bancroftii*
Harvest age	24	46	35	30	102
M.A.I	20.56	15.29	8.00	10.80	2.85
Fraction of land (%)	20%	10%	20%	15%	35%
Timber volume (m^3 per ha)	98.67	70.33	56.00	48.60	101.75
Current timber price ($\$/m^3$)	$100	$100	$78	$157	$100
Future timber price	$136	$181	$123	$231	$373
Harvest value	$13,453	$12,740	$6,900	$11,227	$37,991

Output summary	
Net present value at (y - 1%)	$4,991
Net present value at (y)	$825
Net Present Value at (y + 1%)	($1,826)
Internal Rate of Return (IRR) =	4.27%
Variable factors	
Timber price change per year ('x'%)	1.30%
Discount rate 'y'%	4.00%
Site performance index	100%

Figure 2. Example of Plantation Output Sheet

Users can see and access all the information about a plantation scenario with up to five species from one sheet (Figure 2). Users can also choose to select default data to run the model, or may enter their own growth, cost and price data.

The main outputs of the model are summarised in the 'Plantation Output' sheet in order to make it easy to see what is happening in terms of the volumes of timber produced, the value of the harvest, and timing of returns. This sheet also assists in understanding the effects of changing each variable on the overall financial performance of the plantation. More detailed output from the financial analysis is displayed in the 'Economic Summary' sheet which can be accessed from the 'Plantation Output' sheet by a button bar.

Default values of the model

The model contains embedded estimates for a number of parameters used in the calculation of NPV and IRR, some of which are presented in Table 4. Default values for MAI and harvest ages were derived from a Delphi survey of experts on North Queensland forestry undertaken for 31 cabinet timber species in 1996-97, and are probably the best projections currently available of future growth of many of the species. They are however for restricted growth conditions (relatively fertile basalt derived soils in areas of 1,600-2,000 mm rainfall) and so users may wish to enter site-specific values. This need is to some extent overcome by the inclusion of a site index variable, which allows the user to vary the MAI (growth) data for all species in a scenario according to the conditions for which the model is being used relative to those specified for the estimates.

Table 4. Estimated harvest ages, timber yields and timber prices for eucalypt and cabinet timber species in North Queensland

Species	Harvest age (yrs)	Yield (MAI in m^3/ha)	Timber price ($/$m^3$)
Acacia mangium	23.9	20.6	
Acacia melanoxylon	29.5	10.8	299
Agathis robusta	46.2	15.3	
Araucaria cunninghamii	43.8	19.4	
Beilschmieda bancroftii	102.0	2.9	
Blepharocarya involucrigera	47.9	8.7	
Cardwellia sublimis	58.1	7.5	317
Castanospermum australe	66.7	5.9	280
Cedrela odorata	37.5	10.3	242
Ceratopetalum apetalum	113.3	5.0	355
Elaeocarpus angustifolius	33.8	16.9	317
Endiandra palmerstonii	156.0	4.5	653
Eucalyptus camalduensis	28.1	17.0	
Eucalyptus citriodora	32.0	16.1	
Eucalyptus cloeziana	35.5	17.2	47
Eucalyptus cloeziana (poles)	25.5	15.8	
Eucalyptus drepanophylla	44.4	10.8	
Eucalyptus grandis	31.4	20.6	
Eucalyptus microcorys	30.6	15.8	
Eucalyptus pellita (resinifera)	30.6	16.3	75
Eucalyptus tetreticornis	31.7	16.0	
Flindersia australis	57.5	-	355
Flindersia. bourjotiana	52.5	10.3	373
Flindersia. brayleana	43.0	13.5	205
Flindersia iffliana	56.7	8.0	243
Flindersia pimenteliana	55.0	9.0	317
Flindersia schotiana	48.3	8.5	467
Gmelina fasciculiflora	54.2	8.6	280
Grevillia robusta	35.0	8.0	149
Melia azedarach	31.1	13.7	243
Paraserianthes toona	55.0	7.1	
Toona cilata (australis)	48.8	8.9	336

Source: Herbohn *et al*. (1999).
Notes: Assumed biophysical conditions: basalt-derived soils, annual rainfall 1,600-2,000 mm per year. Price estimates were not available for some uncommon species.

If, for example, the soils are less fertile than basalt derived soils or the site under consideration receives less rainfall, then the site productivity index can be set at less than 100%, reducing the estimated volume of timber produced in the scenario. However, if the rainfall regime in the area for which the plantation scenario is established is greater than 2,000 mm per year and on fertile soil, or irrigation of the plantation is planned, then the site productivity index may be set at greater than 100%. The maximum allowable value for the index is 150%, and the minimum 50%. The default value for the index is 100%.

The default values for the costs of establishing and maintaining a plantation that are supplied with the model are those for Community Rainforest Reafforestation Program (CRRP) plantings in the Atherton Tablelands region. Cost items include land preparation and planting, pruning and certification of pruning, weeding, fertilisers and fencing. Also, allowance is made for opportunity cost for financial returns to the land foregone with plantation establishment, and insurance and road maintenance. It is arguable whether a land rental amount should be included as part of the cost of the plantation. particularly in the case of degraded or otherwise unproductive land.

Timber prices and price changes over time

The generic stumpage price of $\$100/m^3$ has been used as a default value, being the approximate royalty received by the State forest service for plantation Hoop Pine at the time of model development. The model allows users the option to select their own generic stumpage price. Stumpage prices have a major influence on estimates of NPV and IRR, hence sensitivity analysis is reported in the ‘Economic Summary’ sheet.

Given that the most of the species being planted are premium cabinet timbers it is likely that the average stumpage received for these will be considerably higher than that Hoop Pine. The only available estimates of future stumpage price for many of the species in the model have been derived by a commercial consultant and are reported in Russell *et al.* (1993). No default price is supplied for species not examined by Russell *et al.* (1993). Users can set their own timber prices by selecting the ‘Set timber prices’ button on the front sheet, and selecting the ‘Set own timber prices’ checkbox in the dialog box that is subsequently displayed. This will result in a new table being displayed, showing the species and their prices that can be altered individually by users.

Figures 3, 4 and 5 report default costs (CRRP costs) from the Australian Cabinet Timbers Financial Model. Additional prescriptive costs can be added to the prescriptive costs of Figure 3 by entering the cost type and amount under section 2, to be added to the sub-total.

The rate at which timber prices change per year can be set in the ACTFM model by selecting the button ‘Set timber price change rate’. Russell *et al.* (1993) suggested a possible increase of up to 1.3% per annum in timber prices, which is considered conservative by the authors compared to a number of long-term forecasts. The default value for the model is 0% although the option is available to set a real price change of up to +5% or -5%.

Output values generated by the ACTFM

The ACTFM calculates the financial returns that may be expected from a hypothetical eucalypt or cabinet timber plantation, expressed in the form of net present value (NPV) and internal rate of return (IRR) per hectare of the plantation. The NPV of a plantation in the model is the financial returns that can be expected from harvesting the trees for timber in today’s values. The model adopts a 7% discount rate, but also provides NPV estimates with the discount rate varied by +1% and -1%. Outputs from the model are displayed on screen, but a ‘print button’ is available to generate a hard-copy version.

1. Prescriptive costs	Cost/ha
Planning and design	74
Incidental clearing	158
Site preperation and cultivation	265
Cover crop establishment	88
Pre-plant weed control	92
Cost of plants	450
Planting and refilling	645
Post plant weed control	540
Fertilizer	83
Fencing	560
2. Additional Prescriptive Costs	
Sub-total	2955.00

Figure 3. Prescriptive costs sheet

Costs During Plantation	Year of Cost	Cost/ha
Post plant weed control	1	1310
Post plant weed control	2	812
Post plant weed control	3	213
First prune (plus certification)	4	880
Second prune (plus certification)	8	648.6
Third prune (plus certification)	12	864.4
Thinning	8	501
First harvest marking and inventory	49	57
Second harvest marking and inventory	58	57
Third harvest marking and inventory	55	57
Fourth harvest marking and inventory	48	57
Fifth harvest marking and inventory	102	80
Additional costs during plantation		

Figure 4. Costs during plantation sheet

Annual costs	
Land value (per ha)	$2,000
Land rental (capitalised @ 4% of land value)	80
Protection and management	40
Insurence	
Additional Annual costs	
Total annual costs	120

Figure 5. Annual cost sheet

ACTFM takes into account only financial benefits, and does not include non-wood benefits to the firm or agency establishing the plantations. If for example carbon credits were available from growing the trees, then this additional income could be included in the model.
The model assumes fixed harvest ages (default values or entered by the user) and does not directly allow for a trade-off between harvest age and timber revenue. However, the user could enter their own estimates of harvest age and timber price, and could re-run the model with different combinations. The model does not allow directly for any interaction between species in mixed species plantings, although the user could adjust timber yields if they were confident about their estimates of interactions.

Evolutionary development of the ACTFM

A stochastic version of this model has been developed (Harrison *et al.* 2001), while other versions have had other features added such as the calculation of LEV and the ability to add new species for use in specific financial analyses. It has been used by a variety of forestry stakeholders, including forestry investors, consultants, farm forestry advisers, researchers and students; corporate users; valuers and real estate agents.

Development of ACTFM commenced with a very simple financial model that had been produced in a national forestry inquiry. The AFFFM model utilized components of two previous models (ACTFM and the CARE Pty. Ltd. model). Even though the ACTFM represents the culmination of several years of model development, it is still recognized to have limitations in terms of its suitability to evaluate a wide range of forestry investment projects. Experience in developing and trialling the ACTFM led to recognition of a need of some users to model the impact of forestry on the wider business activities along with a more robust modelling framework not dependant on Excel. In response to this, the whole-farm planning components of the CARE Pty. Ltd. financial model and the ACTFM financial model have been integrated in a user-friendly stand-alone Visual Basic programmed package, known as the Australian Farm Forestry Financial Model (AFFFM). This model is designed to evaluate forestry investment proposals on mixed farming properties in eastern Australia.

7. THE AUSTRALIAN FARM FORESTRY FINANCIAL MODEL

The Australian Farm Forestry Financial Model (AFFFM) was developed to assess the potential effects of farm forestry activities on an individual farms' financial position. The key user groups for the AFFFM are anticipated to be farm financial advisors, farm forestry extension officers, companies and corporations involved in plantation establishment and management, and individual landholders. The model was designed to account for the financial effects of livestock and cropping activities, timber plantation development, forestry activities in native forests, and other basic financing costs and revenues for farm enterprises (Figure 6). By adding the details of the financial performance of farm enterprises and financing details the AFFFM allows for not only the assessment of the financial viability of forestry enterprises, but also assessment of whether a farm enterprise can finance new forestry enterprises.

The model was developed as a series of 'modules' relevant to the different activities on a landholding (e.g. grazing, cropping, farm financing and forestry) (Figure 6). Users decide which activities are relevant for each financial analysis. Users of the model specify 'scenarios' for the different activities on a hypothetical landholding. The cashflows included in each scenario are determined by the activities selected to be included in the scenario on the 'Farm Structure' form (Figure 7). The data entered by users to create scenarios can be saved as text files for later retrieval. The settings for the main three activities covered by the AFFFM (i.e. plantations, native forestry and agriculture) can be saved as separate scenarios for each activity. Users can also save and retrieve 'whole-farm' scenarios that include the settings of the main activities and those for farm finances.

Data availability is a key limitation for modeling any forestry system using native tree species in Australia. There is little

published data about the potential growth rates, harvest ages and products for most Australian tree species that can potentially be grown in plantations. The model contains and is linked to data about tree growth and potential products from harvests in several Eastern Australian regions that can be accessed by users. Data relating to cabinet timber species suitable for plantations in far north Queensland is the same as that supplied with the ACTFM (Herbohn *et al*. 1999). Data relating to tree species suitable for plantations in the New England Tablelands and Darling Downs regions was compiled using the Plantgro software package. The Plantgro software was developed by researchers from CSIRO to estimate plant growth in circumstances where there is limited field trial data (Hackett 1991). Data about individual species is stored as text files that can be linked to the AFFFM. When resources become available to expand the regions covered by the model, the technique of using linked text files allows data about new tree species and regions to be added with only small modifications of the code controlling the AFFFM.

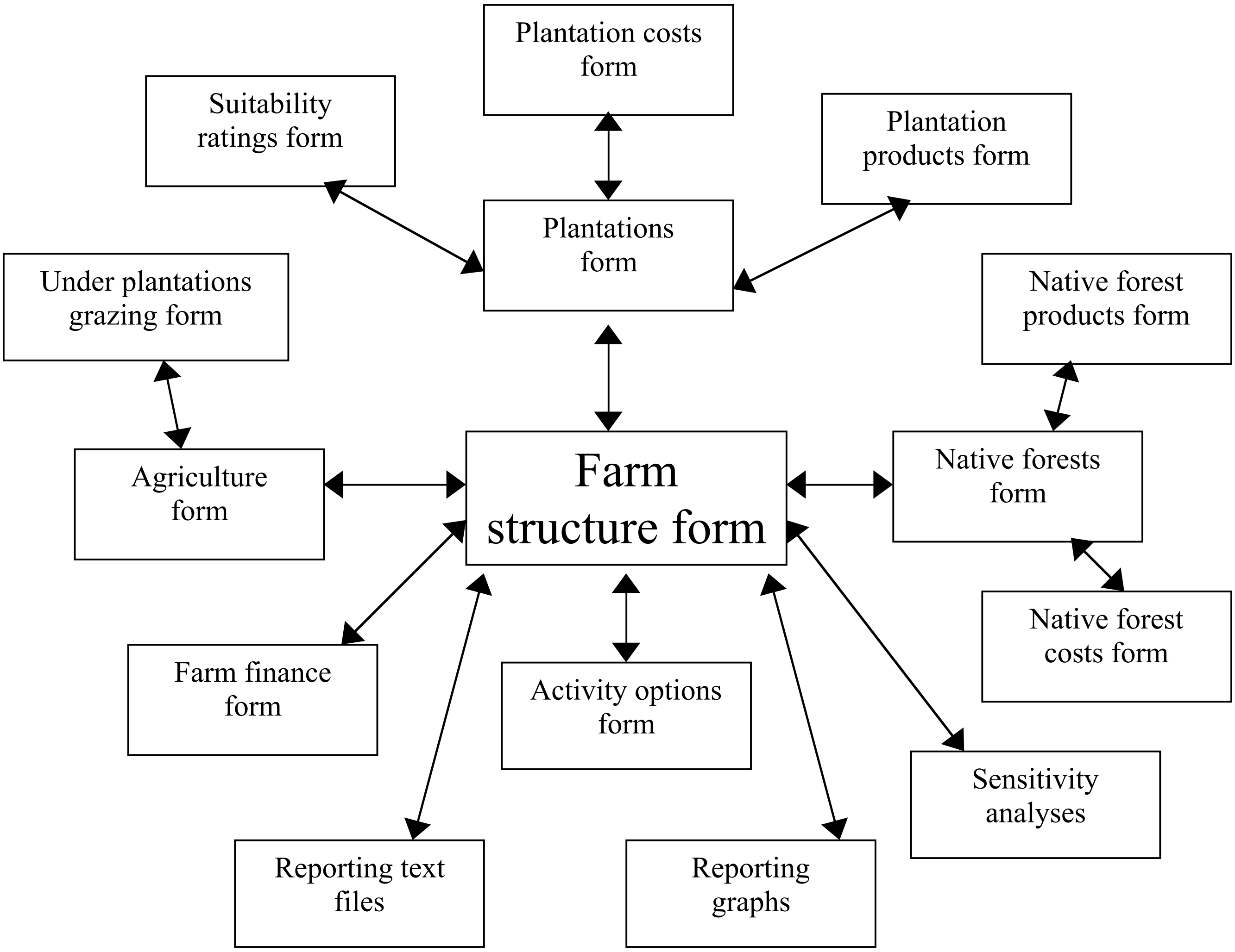


Figure 6. Schematic diagram of the structure of the AFFFM

Notes: Forms in black font are reporting forms whilst those in green font are data entry forms. Lines between forms indicate that users can navigate directly between them.

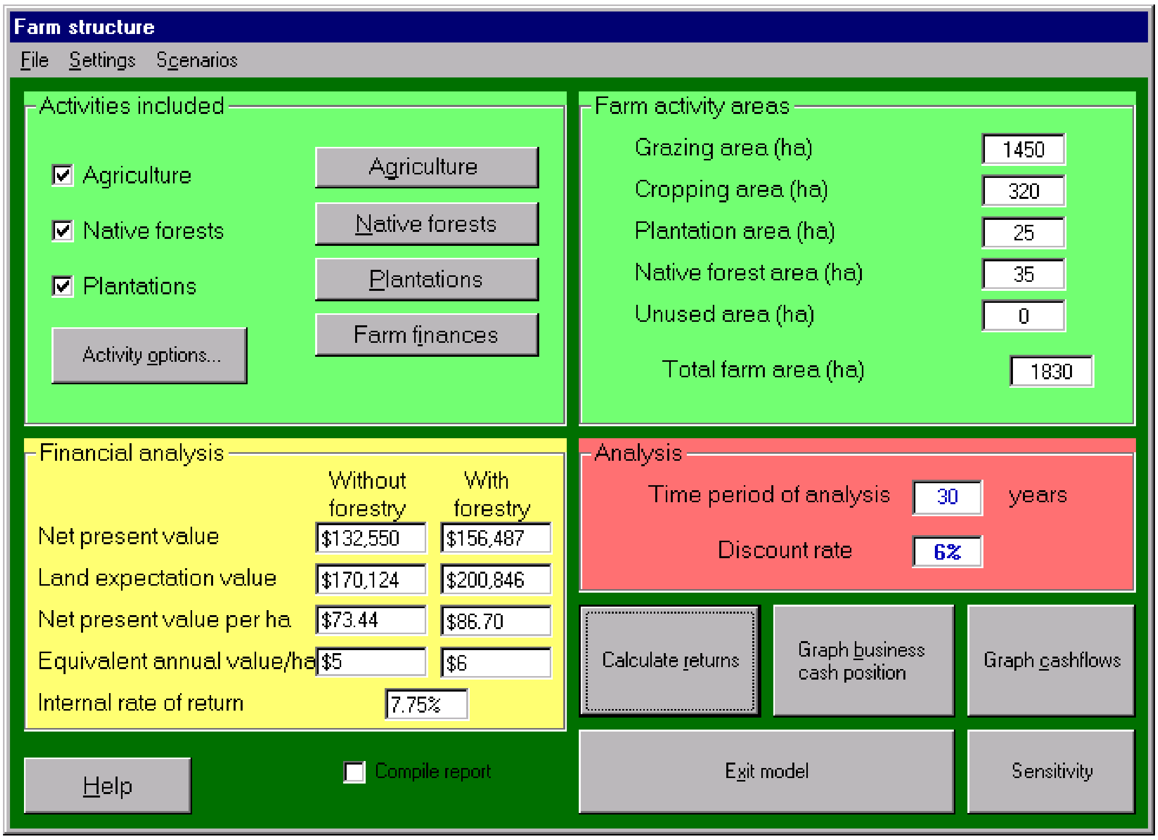


Figure 7. The main reporting form from the Australian Farm Forestry Financial Model

Functions of the model

The financial viability of each activity can be assessed separately or in conjunction with any or all of the other activities. The model calculates the cashflows for each relevant activity in each year of a scenario, and then calculates the overall cashflow for the landholding in each year. The costs and returns for each year in the future are then discounted back to their present value at the discount rate set by the user. The AFFFM lets users assess the relative influence of the various costs and revenues on the farms' financial position by allowing adjustment of these figures for the different activities. Users can also assess the effect of changes in timber prices and plantation growth rates on the NPV of a scenario.

The potential financial impacts of forestry developments are reported using a number of commonly used measures of long term investments including the net present value (NPV), internal rate of return (IRR), land expectation value (LEV) and business cash position of a scenario. Most measures of financial performance are reported to the main form of the program, the 'Farm Structure' form (Figure 7). The model can be used to assess the cashflows from different activities and overall farm cashflows from a combination of activities. When forestry activities are included in conjunction with agriculture the model reports the IRR as the difference between the 'with' and 'without' forestry situations. Other reporting features of the AFFFM include graphic display of yearly cashflows and the business cash position of the farm enterprise; reports of the settings and financial measures of scenarios; and writing of text files with yearly values of key variables.

Data requirements of the model

The data required to operate various modules of the model are listed in Table 5. Those parameters listed as 'basic' are required to operate the various activities of the model while those listed as 'advanced' may be used when the user has sufficient information.

Table 5. Parameters used in the Australian farm forestry financial model

Farm activity	Basic parameters	Advanced parameters
Plantations	Species used – name	Harvest cycles (three harvest ages per species)
	Plantation size(s) – hectares	Plantation product mix per harvest cycle (as % of standing volume) and product prices
	Stems – per hectare OR	Series of plantations
	Percent of hectare per species	Time between plantation establishments
	Mean annual increment(s) – m^3/ha/yr$^-$	Number of plantations to be established
	Harvest age(s) – years	Plantation design (woodlots/shelterbelts)
	Stumpage price(s) – dollars	
	Establishment costs – dollars	
	Maintenance costs – dollars	
	Harvest costs – dollars	
	Annual costs – dollars	
Native forestry	Forest area(s) – hectares	Non-commercial thinning costs and timing
	Mean annual increment – m^3/ha/yr$^-$	Timber products and recovery rates
	Maintenance costs – dollars	
	Stumpage price(s) – dollars	
Livestock	Carrying capacity – DSE/ha	% DSE improvement with shelter
	Stock numbers – No. of head	
	DSE ratings – per head	
	Gross margins – $/DSE	
Crops	Crop type(s) – name	
	Crop area(s) – hectares	
	Gross margin(s) – $ per ha	
Farm finances	Overhead costs – $	Loan(s) – $
	Capital costs – $	Interest rates – loans, savings, overdrafts
	Off-farm income – $	
	Other farm income – $	

The model was developed to cater for different levels of complexity of data within the modules. For example, in the plantation scenario the prices received for timber can be average prices for a single harvest or different prices can be set for different mixtures of products for up to three harvest cycles. In the case of native forestry activities users can again set an average price for all the forest products for various areas or else they can set the proportion of the harvests that will be allocated to various products with individual prices for each product.

8. DISCUSSION

The use of the Visual Basic (VB) programming language has a number of advantages over the use of Excel linked to Visual Basic macros. In editions of Excel since 1997 have used the Visual Basic language to construct macros. Macros are computer code written in Visual Basic language. Macros can be used to automate data processing and reporting, navigate through spreadsheets, carryout data analysis, and provide input and reporting forms. Macros used in Excel require Excel software to operate. Using the Visual Studio software package allows users to utilise elements of the Windows operating system and functions from Microsoft software such as Word, Access and Excel to create their own software programs. Programmers can access virtually all the functions of Excel, Access and other popular Microsoft programs with computer code plus create specific data input, analysis and reporting forms, as well as help systems, all similar to those in standard commercial software. Once the code for the data input, analysis and reporting has been written it can be compiled into software that can be copied, distributed and installed onto most computers with Windows operating systems.

Using computer code to develop the AFFFM has other advantages over the use of macros linked to Excel workbooks. The use of a programming language allows the development of more complex models that better represent real-life situations. One example is the plantation module of the AFFFM. Whereas the ACTFM allowed only one harvest age and timber price per species in a plantation, the AFFFM allows users to use up to three harvest ages per species and up to 8 product categories per harvest. The processing speed of programs written in Visual Basic relative to that of Excel has allowed this improvement in the sophistication of the model. It has also allowed the incorporation of the financial details of other farm enterprises, and for more options for users in each part of the model. These functions were not possible in the Excel based ACTFM because of limits to the file size that could be handled by most computers. Other features of the AFFFM such as the links to text files and the incorporation of an on-line help system are not possible in Excel.

The expansion of the ACTFM to a whole-farm model allows it to be used to answer a question basic to financial appraisals that could not be handled by the ACTFM, can the new enterprise be financed? The AFFFM still allows for financial appraisals of individual enterprises like the ACTFM as well. By looking at the cashflows and financing costs and revenues of forestry enterprises in conjunction with those specific to the individuals other farm and non-farm enterprises it is possible to assess the effects of different rates and scales of plantation establishment, plus the impacts of finance costs and revenues.

Like all models, the ACTFM and AFFFM cannot perfectly simulate what will occur in terms of the risk factors listed in Table 2. While the AFFFM does have some features to calculate the interaction between forestry and agricultural activities, these are limited and no default data is supplied for these parameters. Agricultural returns remain static in terms of the gross margins used to calculate returns ignoring the likelihood of climatic and market factors changing over the course of a scenario. Likewise timber prices are assumed to remain static. Some form of sensitivity or risk analysis should be used for key variables to determine the risk profile of the scenario.

REFERENCES

Hackett, C. (1991), *Plantgro: A Software Package for Coarse Prediction of plant Growth*, CSIRO, Melbourne.

Harrison, S.R. and Herbohn, J.L. eds. (in press), *A Whole-Farm and Regional Agroforestry Decision-Support System: A report for the RIRDC/Land and Water Australia/FWPRDC Joint Venture Agroforestry Program*, Canberra.

Herbohn, J.L, Harrison, S.R. and Emtage, N.F. (1999), 'Mixed species plantations of rainforest and eucalyptus cabinet species in North Queensland', *Australian Forestry*, 62(1): 79-87.

13. Estimating and Integrating Non-market Benefits of Forests in Project Appraisal

Jungho Suh

Forests generate a variety of environmental effects, including watershed protection, air pollution reduction, flooding mitigation, wildlife habitat and public recreation. These environmental benefits are not marketed and not site specific. Thus, they need to be considered from a social perspective. In recent years, a number of applications have been made to estimate non-market environmental values in monetary terms and incorporate these estimates into cost-benefit analyses in forestry research projects. This module provides an introduction to the concepts and methods of non-market valuation by which to estimate the aggregate benefits to society of environmental improvement. In the next section, the categorical framework of economic value is introduced. Non-market valuation methods are then discussed, including the hedonic price method, the travel cost method, the contingent valuation method and conjoint analysis. A contingent valuation survey of the Kakadu Conservation Zone in Australia is examined in particular. A few alternative valuation techniques are next briefly reviewed. Finally, ways of integrating non-market values in project appraisal are discussed.

1. THE ECONOMIC VALUE OF A FORESTLAND

As long as attempts are made to quantify the value of non-market environmental goods and services, the term ‘value’ is confined to *economic value*. In the resource economics literature, the economic value of a forestland is usually broken into use value and non-use value, the latter also being referred to as passive use value. The use class of economic value consists of direct and indirect use value, and option value. The non-use class of benefits falls into two sub-categories, namely bequest value and existence value. Figure 1 illustrates the use and non-use values that a forestland provides.

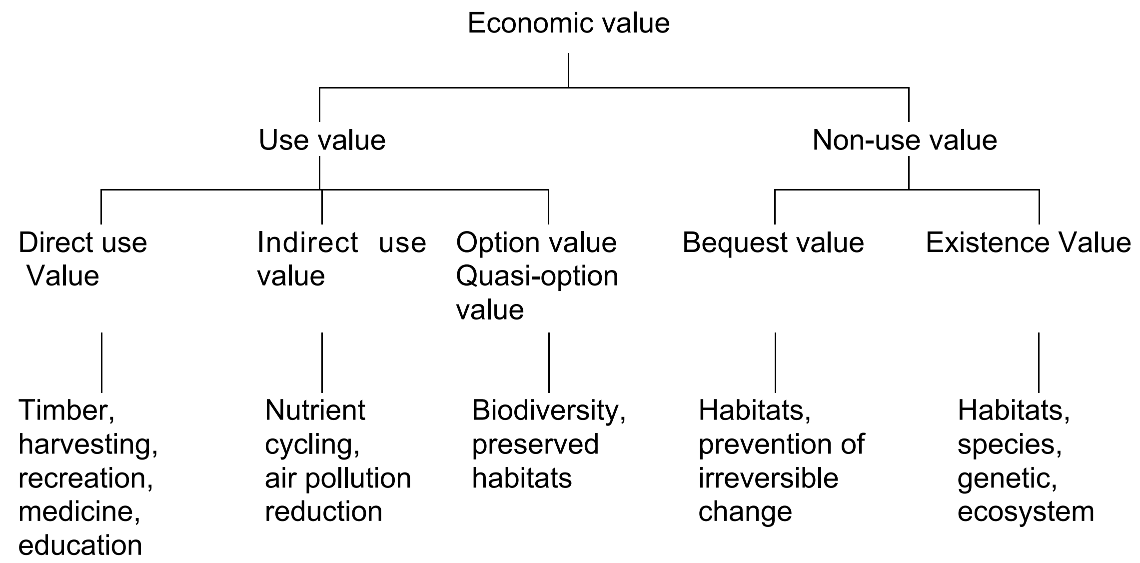


Figure 1. The economic value of a woodland

Source: Adapted from Bateman and Turner (1993) and Barbier (1994).

Direct use value is made up of consumptive and productive use value. Indirect use value refers to values conferred by forests such as carbon sequestration and water purification. Option value is defined as the potential use benefit, opposed to present use value, of an environmental good. The value is viewed as the willingness-to-pay (WTP) for preservation of a natural resource that will be made use of at a later date by the present generation. Bishop (1982) provided an excellent review of the evolution along with an extension of the concept of option value. It is known that Weisbrod (1964) originated this concept by proposing that many individuals expect they may possibly visit a national park for example and are willing to pay for an option that would guarantee their future access.

Bishop (1982) extended the concept of option value with supply side uncertainty. If a risk averse consumer is certain of demanding the services of an environmental asset in the future and uncertain about its future availability, there exists a positive option value. That is, the maximum WTP to avoid the risk to the supply of the environmental resource is larger than the expected loss. This concept is grounded on the fact that an individual in the real world will be willing to pay more than the expected consumer surplus in order to ensure that he or she can make use of the environmental resource later on. Edwards (1988) reported empirical evidence of positive option value from a study of households' WTP to prevent uncertain future nitrate contamination of groundwater in Cape Cod, Massachusetts. Bishop noted that option value ceases to exist in the case of supply side certainty.

Quasi-option value is present when there is uncertainty about future availability of a natural resource given some expectation of the growth of knowledge. For example, there are uncertain benefits for scientific or commercial purposes from the preservation of a tropical forest, which could become more certain through time as information is accumulated about the uses. Arrow and Fisher (1974) originally introduced the concept of quasi-option value in the context of an irreversible development decision. Quasi-option value is always positive if the expected growth of information is independent of a proposed development of a natural resource. In contrast, as Freeman (1984) argued, if the uncertainty is primarily about the benefits of development, this strengthens the case for development. That is, quasi-option value is negative. Whatever the case is, the presence of quasi-option value implies that the value of additional information is likely to be of greatest importance in valuing goods subject irreversible changes (Mitchell and Carson 1989).

Krutilla (1967) argued that many persons may be willing to pay for the satisfaction derived from knowledge of the bequest of unique environmental resources to future generations. Thereafter, bequest value is often defined as the benefit accruing to current generations from knowing that future generations will benefit from the resources. This concept takes a strong stance for intergenerational moral duty so as to prevent future sufferings from the shortage or degradation of the environment.

Pearce and Turner (1990) defined existence value as a value placed on an environmental good that is unrelated to any actual or potential use of the good. The concept of existence value becomes complicated because it is sometimes mixed with that of intrinsic value of a resource. Existence value is often recognised on the basis of ecocentric value orientation that nature has the right to exist for its own sake, and destruction of species and wilderness is intrinsically wrong. From this viewpoint, Aldred (1994) insisted that existence value captures a WTP to preserve the environment for the continued existence of particular species or whole ecosystems. He argued that existence value may include intrinsic value, but they are not the same thing. Solow (1993) supported the view that particular landscapes or species have to be preserved for their own sake because they are intrinsically important to preserve. Pearce and Turner (1990) stated that intrinsic value equals existence value as long as the former is defined as something unrelated to human use but captured by people through their preferences in the form of non-use value. On the other hand, Mitchell and Carson (1989) saw the term 'intrinsic' being contradictory to the term

'economic' so that the intrinsic value cannot be part of economic value. A confusing use of the term 'existence value' can be avoided by supposing that existence value embodies the welfare of particular species and whole ecosystems, but ultimately appeals to human welfare. In other words, existence value may stem from intrinsic motives such as sympathy, responsibility and a concern about the state of the world that some people may feel towards non-human beings, but the value is still anthropocentric and does not reveal the intrinsic value to non-human beings (Bateman and Turner 1993). In this context, Bateman and Langford (1997) clarified that existence value is a human value whereas intrinsic value is a non-human value, which cannot be estimated.

2. CLASSIFICATION OF THE BEHAVIOURAL LINKAGE APPROACH TO NON-MARKET VALUATION

Smith and Krutilla (1982) divided measurement techniques of environmental benefits into the physical linkage approach and the behavioural linkage approach. Under the category of physical linkage approach, a researcher can specify a model of the relationship between levels of an environmental contaminant and some type of observed damage such as reduced agricultural crop yields or impaired human health. Linked with physical data, the benefit of the reduction in the contaminant can be estimated in dollar values. When there is no such physical link to be observed, an alternative is the behavioural linkage approach. This approach uses a conceptual linkage between the services provided by environmental resources and some directly or indirectly observable consumer responses. The behavioural linkage approach for valuing environmental amenities is traditionally divided into four categories of economic valuation methods as presented in Table 1.

With the indirect and revealed preference approach, the information from revealed markets – that is, actual choices made by consumers – is used to develop models of choice. One problem is that data on market transactions and product characteristics are often unavailable or incomplete for environmental goods.

When revealed market data are not reliable or are unavailable, economists have used the stated preference approach that relies on hypothetical market situations. As a typical example, the economic value with respect to a change in the level of environmental service flows of an unpriced natural resource can be estimated by eliciting consumers WTP or willingness-to-accept compensation (WTA) amounts for the proposed change. WTP measures give welfare estimates for quality-improving changes in resource use, whereas WTA measures provide information about welfare decreases resulted from quality-decreasing environmental moves.

Table 1. Behaviour-based valuation methods of environmental public goods

Type of valuation approach	Direct	Indirect
Revealed (observed) market behaviour	DIRECT and REVEALED Referenda Simulated markets	INDIRECT and REVEALED Travel cost method Hedonic price method
Stated (hypothetical) markets	DIRECT and STATED Contingent valuation method Allocation games with tax Refund	INDIRECT and STATED Contingent ranking Contingent rating Choice modelling

Source: Adapted from Mitchell and Carson (1989, p. 75).

3. THE HEDONIC PRICE METHOD

The hedonic price method (HPM) has been widely used to estimate the externalities of environmental characteristics, particularly air quality and visual amenity, in the housing market. A botanical garden that gives pleasure to passers-by is an example

of a *positive externality*. By contrast, a negative externality occurs in the situation in which air pollutants such as sulfur dioxide can make life miserable for people breathing the air and cause health problems. Since the effects associated with production of botanical gardens or consumption of air pollutants extends outside the market to some third party, they are termed 'externalities'. In HPM applications, it is assumed that externalities are capitalised into the value of real properties. Differentials in property prices are then examined to find the marginal value of an environmental good under investigation. One problem with this technique is that it is often difficult to obtain an adequate sample of property transaction records.

Garrod and Willis (1992) investigated whether the amenity benefits of Forestry Commission estate in the UK are reflected in the values of nearby properties. A number of independent variables in the hedonic price model were tested, including the forest characteristics such as the proportions of Forestry Commission estate areas covered by broadleaved trees. The authors found that selling price increased by about A$111 from 1% increase in the relative proportion of forested area in a given 1km^2 to broadleaf woodland, with all other independent variables held at their mean values. The aggregate amenity benefit, which Forestry Commission estate provides to those households that live in close proximity, was found to be about A$913,000.

Tyrvainen and Miettinen (2000) examined the relationship between property prices and urban forest amenities of Salo, which is located 110km to the northwest of Helsinki. The town of Salo had approximately 1,100ha of forest, 74% of which was found in the urban fringe. Detached houses or terraced houses predominated in most housing areas.

A range of explanatory variables was tested including housing attributes, locational attributes and four different variables measuring forest amenities. Forest amenity variables included distance to nearest wooded recreation area, direct distance to the nearest forested area, relative amount of forested areas in the housing district and the view from the dwelling window. This study found that a 1km increase in distance from a forest park reduced terrace house price by 5.9%, and a forest view increased terrace house price by 4.9%.

4. THE TRAVEL COST METHOD

TCM was the first technique ever used to measure the demand for natural resources for recreation purposes in terms of what people spend in travel costs to visit them. As outlined by Bateman (1993), Hanley and Spash (1993) and Bennett *et al*. (1996), the TCM was conceptually first suggested by Hotelling in a letter to the director of the US Park Service in 1947. The letter suggested that "the price for visiting a park or other non-marketable recreational area (even one for which entry is free) would vary according to the travel costs of visitors coming from different places" (reported by Portney, 1994, p. 4). This suggestion was formally introduced to the literature by Wood and Trice (1958). Clawson (1959) modified the idea and first developed empirical models. The TCM was further developed by Knetsch (1963) and Clawson and Knetsch (1966).

In practice, two types of TCM are used. The zonal method (ZTCM) divides the entire area from which visitors originate into a set of approximately concentric zones emanating from the recreation site, representing increasing levels of visit cost, and then defines the dependent variable as the 'visitor rate' – the number of visits made from a particular zone in a period divided by the population of that zone. The individual method (ITCM), in contrast, simply defines the dependent variable as the individual's annual number of visits to a site, and the independent variable as the cost of travelling to the site. Then, the recreational demand curve can be produced. With the ITCM, it is possible to involve socio-economic variables under the hypothesis that they are separate factors influencing travel behaviour. The ITCM does not require data about zonal visitor rates. Cooper and Loomis (1990) mentioned that the ZTCM model has several drawbacks relative to the ITCM model, including statistical inefficiencies from grouping data and a less direct link to consumer demand theory. Hence, in these aspects, the ITCM seems to be preferable. On the other hand,

Fletcher *et al*. (1990) came to the view that the zonal approach more adequately accounts for declining participation rate proportioned to increasing travel distances. Uncorrected estimates of the ITCM may lead to biased results in the circumstance where visitation rate to a site declines as distance to the site increases. For these reasons, many travel cost practitioners have adopted Clawson and Knetsch's (1966) ZTCM.

Suppose there is a recreation area attracting people from three zones having populations of 1,000, 4,000 and 10,000 people, respectively as illustrated in Table 2. Visitors from Zone 1 who have a cost of $1 participate at the rate of 300 per 1,000. If the cost to the visitors from Zone 1 increased to $2, the new visitation rate would be the same as that of Zone 2, i.e. 200 per 1,000.

Table 2. Demand schedule for recreation experience for a hypothetical area

Zone	Population	Cost per visit	Number of visits	Visits per 1,000
1	1,000	$1	300	300
2	4,000	$2	800	200
3	10,000	$3	1,000	100

Using data presented in Table 2, one can map the recreation demand curve for the hypothetical area by plotting visit rates across all zones against visit costs. The total consumer surplus for recreation is then obtained by integration of the area under this demand curve. In addition, using this demand curve, one can estimate the resulting change in consumer surplus with a change in environmental quality of the site. If the demand for the recreational use of an environmental resource such as a national park increases as its quality improves, the demand curve shifts outwards under the assumption that the quality of a natural resource is positively correlated with its recreational use value. Then, the monetised incremental benefits to visitors of improving the quality of the park can be estimated.

It should be noted that the demand curve relating visitation rate and cost per visit indicates demand for the 'whole recreation experience' rather than the recreational use value of the resource alone. The whole experience includes travel to and experience on the site, travel back and recollection. This distinction is important when the researcher intends to estimate the recreation value attached to a particular to $3. It can be seen from the table that a $3 rise would result in choking off all visits. The regression of the estimated number of total visits on added travel cost will eventually generate the new demand curve, from which on-site recreation benefits can natural resource. In order to isolate the on-site recreation experience from the whole experience, Clawson and Knetsch (1996) suggested constructing the new demand schedule for the resource by plotting added entry fee to the site as a proxy for the price variable and total visits as the quantity variable.

Suppose that the entry fee on visits to the site has increased. One can next determine the effect of a rise in the entry fee of say $1. The visitors from Zone 1, for example, who used to spend $1 per visit, are now faced with the situation of spending $2 due to the rise in overall cost as much as $1. People in this region would participate in the same recreation activities on the site at the rate of 200 per thousand, as indicated in the Table 2, when faced with costs of $2 per visit. One can then estimate the number of visits under varying added travel cost using this result. The number of recreatists attending from Zone 1 would be 200 – i.e. 200 per thousand multiplied by 1,000, the base population of Zone 1. The numbers of visits to the site from other zones can be found in the same manner. Table 3 reveals the estimated total number of visits to the site under varying added entrance fee from $1 be estimated.

A number of TCM applications (e.g. Everitt 1983; Willis and Garrod 1991; Bennett 1995) have demonstrated that forestlands can have substantial recreational use value

and TCM is a powerful tool to estimate recreational use value of a particular forestland. For example, Bennett (1995) estimated economic value associated with recreational visits to two National Parks in the north-east of New South Wales, Australia. Dorrigo National Park forms two amphitheatres of 3,600ha from the tablelands to the coastal plain, and is the habitat for a wide variety of flora and fauna. Gibraltar Range National Park of 17,300ha is noted for its high number of rare and threatened plants. The park is essentially a high plateau with granite occurring over much of its area. Bennett (1995) estimated recreational value of A$34 per visit to Dorrigo National Park and A$19 per visit to Gibraltar Range National Park.

Table 3. Effect of increases in travel cost on visits to a hypothetical area – deriving a demand schedule

Zone	Number of visitors at added entrance fee			
	$0	$1	$2	$3
1	300	200	100	0
2	800	400	0	0
3	1,000	0	0	0
Total visitors	2,100	600	100	0

There are many continuing controversies over TCM. The practices of time valuation in TCM are the most controversial. Fletcher *et al*. (1990, p. 125) pointed out that “many of the most vexing problems encountered in TCM literature, both theoretical and empirical, relate to the appropriate inclusion of time constraints and time values”. Likewise, Randall (1994) argued that the level of money-valued welfare measures in TCM is inherently subjective because it depends on arbitrary and simplistic specifications of household production technology and particular accounting or analytical conventions for the household’s implicit cost of time. Another fundamental problem with TCM is that the valuation technique cannot be applied to some forestlands that people do not visit for recreation purposes. Harrison (2001) drew attention to a number of other complexities arising in application of TCM. First, when people visit multiple sites during a single recreation trip, it can be difficult to allocate a proportion of their travel cost to a specific site, which is the target for valuation. Second, since recreation demand typically is highly seasonal, and has peak visitation during school and public holidays, it is normally necessary to carry out surveys on a year basis. Third, where there is a group visit, with members of varying ages, the issue arises of which members to include as recreationists and how to allocate costs between party members.

5. THE CONTINGENT VALUATION METHOD

The CVM question involves asking individuals how much they would be willing to pay or whether they would be willing to pay a given amount to prevent a specified decrease in the quality or quantity of a particular good, or how much they would be willing to pay to obtain improvements. WTP values elicited from CVM surveys are ‘contingent’ upon a hypothetical market described to the respondents, hence this approach came to be called the contingent valuation method. Mitchell and Carson (1989) provided the full history of early development of CVM, which first came into use in the early 1960s by Davis (1963).

CVM has been widely applied in the last three decades. Mitchell and Carson (1989) listed more than 120 CVM studies, most of which were undertaken in the USA and Europe. Carson *et al*. (1995 cited in Bennett 1996, p. 190f) listed 2,131 CVM studies and papers applied to health, education and transportation as well as the environment.

Major bidding methods are continuous and discrete ones. The continuous method usually refers to as the open-ended elicitation method, where respondents are simply asked to state their maximum willingness to pay for the good being

valued. The discrete method refers to as the dichotomous choice question, where respondents determine whether their WTP is larger or smaller than a set dollar value. The dichotomous choice format is also known as a take-it-or-leave-it or closed-ended format, or a referendum format if the question is structured so that it is similar to a referendum.

Some examples of open-ended CVM applications include Mattsson and Li (1994) and Hadker *et al*. (1997). Mattsson and Li (1994) estimated non-timber use values of forests in the county of Vasterbotten in northern Sweden. Forests in the county are dominated by pine and spruce. This CVM study examined a variety of on-site human activities such as camping, berry and mushroom picking, hiking or simply taking walks, as well as the off-site visual experience of the forests. Respondents were asked the amount that they would be willing to pay annually for using or experiencing the non-timber commodities to the level that they currently use or experience them.

Hadker *et al*. (1997) estimated the preservation value of the Borivli National Park in Bombay. Notably, this study is one of the first reported uses of CVM in India (cited from www.epa.nsw.gov.au/envalue). The Borivli National Park of 104km^2 accounts for one-fifth of the Bombay Metropolitan Region. The national park includes the water bodies that supply drinking water to Bombay and is home for a large number of endangered mammals, birds and reptiles. The park being valued is the most popular park in India. According to the authors, about 2.5 M people visit the national park each year. The park has been facing financial constraints and its potential as a recreational spot and habitat for wildlife has been deteriorating. In this study, Bombay residents were asked how much they would willing to pay to maintain and preserve the Borivli National Park by funding an autonomous management body. The mean household WTP of Bombay residents was found to be 7.5 INR per month for 5 years. The authors noted that despite India being a developing country with medium to low income levels, the empirical evidence suggested people are willing to pay for preserving environmental amenities.

The open-ended CVM question is theoretically vulnerable to strategic bias. This bias occurs in either one of two different ways. Respondents may understate their WTP for a welfare-improving change if they believe that others will bid sufficiently high to provide the desired quality of an environmental good. By contrast, respondents may overstate WTP for an environmental benefit, when they believe that the possibility of the improvement going ahead will increase if they bid high. For this reason, open-ended formats had largely been superceded by the dichotomous formats in CVM applications (Blamey 1996).

The dichotomous choice format is the most commonly employed format in CVM studies nowadays. Using the dichotomous CVM, researchers may ask people as to whether they would pay a specific amount of dollars (\$*X*) if conditions of a natural resource were changed from $\{z_1^0, z_2^0, \cdots, z_k^0\}$ to $\{z_1^1, z_2^1, \cdots, z_k^1\}$, where z_k^0 represents the k^{th} attribute under the current status and z_k^1 represents the k^{th} attribute under an alternative. It is not necessary to have all attribute levels in the new option distinct from the current status.

Respondents are expected to answer 'yes' – i.e. to willing to pay \$*X* – if the utility with the new attribute levels and \$*X* decrease in income is greater or at least as much as the *status quo* utility. From the response, researchers can elicit directly each respondent's WTP for the specified environmental changes. The basic model for an individual who is willing to pay \$*X* for some changes in environmental quality is thus:

$$V(E_1, M_0 - X) \geq V(E_0, M_0) \quad (1)$$

where X is the individual bidding price and E denotes the quality of the natural resource. V represents utility or satisfaction, which is a function of the individual's income M and the quality of the resource. With an amount of money M_0, the individual can purchase environmental or non-environmental goods and services. The maximum that a person

is willing to pay for promotion of improvements in the quality of the particular public good from E_0 to E_1, will be such that both sides of the above expression become equal.

Imber *et al*. (1991), as a typical example of the close-ended CVM application, used the CVM technique to estimate the dollar value Australians place on the Kakadu Conservation Zone. The zone is located in the boundaries of the 20,000km^2 Kakadu National Park, but was not part of the national park. Most of the Kakadu National Park is registered on the World Heritage List. The park is annually visited by about 230,000 people, and renowned for its geographic features, wetlands, wildlife and scenic vistas. Mining was proposed on the conservation zone. The extent of environmental damage from the proposed mine was in dispute. Some people argued that the conservation zone should be made part of the Kakadu National Park, where any mining activities are not allowed. HPM is not applicable to estimate the value of the conservation zone simply because an adequate sample of property transaction records cannot be obtained. Nor is TCM an appropriate method due to the occurrence of part-whole bias. This bias occurs where the good being valued by the respondent differs from the commodity that is intended to be valued. In other words, respondents can experience difficulty in distinguishing between the environmental value of the conservation zone and the whole national park. This study found that Australians are willing to pay about A$124 per person per year for 10 years to avoid the effects of the major impact scenario, and A$53 to avoid the effects of the minor impact scenario. The mining proposal was finally withdrawn due to a variety of factors, including the high conservation values.

CVM mimics real market transactions. Suppose that a person comes to a shopping mall to buy a T-shirt. Before purchasing a shirt, the person needs to choose the desirable size, colour and design of the shirt, and most importantly his income level. Likewise, the CVM questionnaire must be framed in such a way that respondents are made aware there are an array of substitute and complementary goods for the environmental good under consideration, and also budget constraints. Otherwise, framing bias may occur. For this reason, Imber *et al*. (1991) asked several preliminary questions for the framing purpose, gradually bringing respondents to the WTP question. The preliminary questions included: “Do you think Australia needs to concentrate more on protecting the environment, or more on developing the economy, or would you say we currently have a reasonable balance?”; and “What do you think the two or three environmental matters most important to Australia right now?” To give sufficient information to respondents, Imber *et al*. (1991) also showed respondents a set of colour photographs of the site and its surroundings after respondents completed the preliminary section. If there is insufficient information about the commodities being valued, respondents’ WTP may not be equivalent to their actual WTP. In giving information at this stage, CVM interviewers must be in the neutral position so as to prevent potential interviewer bias.

Respondents were then asked whether they would be willing to have their income reduced by A$2 a week for the next 10 years to add the Kakadu Conservation Zone to the Kakadu National Park rather than use it for mining. If a respondent answered yes about whether he or she would pay A$2 a week, another WTP question was asked using a higher price of A$5. If the answer was no, lower price of A$1 was then proposed. Respondents were split into a few sub-samples, each sub-sample being asked if they are willing to pay the distinct combination of dollar amounts. This is necessary to ensure that there is enough variation in the WTP variable to carry out statistical analysis on it. The range of maximum WTP amounts was predetermined through focus group meetings.

It is strongly recommended to make use of follow-up questions. Socio-economic data such as age and income need to be obtained to test whether these variables influence the WTP responses. One or more follow-up questions are also required to check whether respondents understand the choice they are being asked to make and to discover the reasons for their answer. Imber

et al. (1991) included an interesting follow-up question: 'Do you recycle things such as paper or glass?' This question was designed to determine whether environmentalists predominated within the respondents and whether the attitudes had an impact on the WTP bids.

Apart from the possible occurrence of biases that are inherent to non-market valuation or generated by the questionnaire design and survey processes, CVM has come under rather philosophical criticism. Most critiques are concerned about the categorical mistake. According to Sagoff (1988), CVM has its theoretical grounding in measuring consumer preferences, but this method is inappropriate for measuring non-use value that is related to citizen preferences. In other words, if respondents are motivated to bid WTP by the role of citizen in a value survey, they are concerned with the public interest, rather than their own self-interest. Thus, when answers to CVM questions do not arise from an expression of underlying consumer preferences, CVM estimates are not a suitable source of information about values for CBA (Diamond and Hausman 1993). By the same token, Knetsch (1994) stated that CVM responses are more likely to be an indication of an attitude or good feeling of moral satisfaction than of economic value, and therefore CVM results can provide little or no guide to resource allocation policies.

6. CONJOINT ANALYSIS

Conjoint analysis is a variant of CVM in the sense that both are used to measure preservation value of non-market environmental goods based on hypothetical behaviour and they require survey respondents to trade off dollars for attributes. In conjoint analysis applications, respondents are asked to rank or rate multiple profiles, or choose only one option from the given number of options. In the environmental economics literature, these question formats are named 'contingent ranking', 'contingent rating' and 'choice modelling', respectively. Compared to CVM, a single conjoint analysis exercise can separately and simultaneously estimate the coefficients of all factors involved in choice sets.

7. SOME FURTHER VALUATION METHODS

In the economic valuation literature, alternative valuation methods that have not been discussed in this module so far include the production function approach (Barbier 1994), composite methods of TCM and CVM, the contingent activity method (Heyes and Heyes 1999) and benefit transfer.

The production function approach is unique in the sense that it views the environment as an input into production process to capture the value of ecological functions. For example, Narain and Fisher (1995) used this approach to model the contribution of the Anolis lizard to pest control. Despite provision of a useful way in which to value elements of environmental functions, the approach requires detailed knowledge of the physical effects on production of changes in an environmental resource (Barbier 1994), and often entails a number of assumptions due to the paucity of adequate data on how an environmental function is linked to the production of other goods (Acharya 1998).

Hoagland *et al*. (1995 cited in Davis and Tisdell 1996) suggested the use of multiple valuation techniques to help correct for any potential biases attached to a specific technique. For instance, the 'hedonic TCM' has been in place as a class of methods for valuing non-market public goods. The method has been applied to value the characteristics of a destination such as proportion of the forest as open space, presence of water features and diversity of species, using revealed preference data (e.g. Englin and Mendelsohn 1991; Hanley and Ruffell 1993).

Davis and Tisdell (1996) used a composite of TCM and CVM to value the consumer surplus of scuba divers at Julian Rocks Aquatic Reserve, a small marine protected area in New South Wales, Australia. Three points were estimated: (1) the cost that consumers were currently confronting in traveling to and accessing the attraction and the annual number of visits they undertook each year at that price, as in TCM; (2) the upper limit to the cost of visitation beyond which the consumer would

cease to access the site, as in CVM; and (3) the frequency with which they would willing to access the site each year if the cost was half way between their actual costs and upper limit, data of which were also obtained by a CVM question.

Heyes and Heyes (1999) developed 'hypothetical TCM'. Instead of being asked directly to place a dollar value on a hypothetical recreational site, respondents under the hypothetical TCM are asked the maximum additional distance they would have been willing to travel in order to access such a site. By converting these into monetary equivalents, alternative measures of consumer surplus can be derived. Heyes and Heyes (1999) drew attention to their finding that the estimates derived in this way lay somewhere between those taken from the TCM and CVM analyses. They stated that the principle advantage of this method is that it can be expected to reduce the extent of protest-motivated bidding. That is, respondents are in effect being asked how much they would be willing to pay for access to the site, but the payment is expressed in a non-monetary currency, i.e. travel distance.

Another possible approach is to transfer the benefit or cost assessment estimated from previous studies to a current valuation problem. The practice of benefit transfer has been developed because the high time and other resource requirements often prevent new valuation studies. Willis and Garrod (1995) suggested constructing a database of all environmental benefit estimates with details of the modelling procedures and all relevant assumptions in order to make further progress of valuation and benefit transfer studies. Morrison (2001) provided a review of three currently available databases, viz. ENVALUE (New South Wales EPA Environmental Valuation Database), EVRI (Environmental Valuation Resource Inventory) and NZDB (New Zealand Non-Market Valuation Database).[1]

[1] These are found at the wetsites: www.epa.nsw.gov.au/envalue; www.evri.ec.gc.ca/evri/; and learn.lincoln.ac.nz/markval/. The Department of the Environment, Transport and Regions of the UK provides an environmental valuation source-list for the UK at: www.detr.gov.uk/environment/evslist/.

8. INTEGRATING NON-WOOD VALUES OF FORESTRY PROJECTS APPRAISAL

A question often arises as to how to evaluate a new land-use option for a forestland. Conceptually, there are several solutions to the problem of integrating non-wood values in projects appraisal.

Cost-benefit analysis

Assume there are only two options for a particular forestland area. For example, a given habitat can be either preserved or developed. The preservation option is not available. A conventional evaluation method for dealing with discrete development-preservation options is social cost-benefit analysis (CBA). The basic CBA rule for these two options is to compare the net benefit of each option. Whether to preserve a forestland or develop it from society's standpoint is dependent upon the difference between the net present value (NPV) of these exclusive choices. The application of the traditional CBA rule as to whether a project can be accepted can be expressed as:

$$\mathrm{NPV} = \sum \frac{\mathrm{NB}_{Dt} - \mathrm{NB}_{Pt}}{(1+r)^t} > 0 \qquad (2)$$

where r is the discount rate and NB_{Dt} and NB_{Pt} are net development benefits (development benefits minus costs of the development option, i.e. input costs) and net preservation benefits (preservation benefits minus costs of the preservation option such as policing, maintenance and monitoring costs) in year t, respectively.

Safe minimum standard

Pearce and Turner (1990) and Tisdell (1991) discussed the 'safe minimum standard' (SMS) as an alternative to CBA of discrete options on an irreplaceable natural resource. The SMS approach stresses that one should avoid irreversible environmental damage unless the social cost of doing so is unacceptably high. The rule is intended for minimising the maximum losses, which

is called a *minimax* strategy. Assume there is no compromise between preservation and development, that is, there are only dichotomous choices. For example, if a forestland is developed, the maximum loss is the preservation benefit (B_P). Alternatively, if the given habitat is preserved, the maximum loss is the development benefit (B_D).

In effect, the SMS amounts to comparing the net present values of B_D and B_P. The lower B_D is, the more the *minimax* solution is likely to be the preservation option. In other words, the SMS approach makes preservation the preferred option unless the social cost of preservation (i.e. the forgone development benefit) is unacceptably large. This rule appears precise except that it requires value judgement to determine what is unacceptably 'large'. Pearce and Turner (1990) pointed out that the requirement of the SMS approach may be a deliberate attempt with the aim of not relying on a single criterion for making discrete choices. Tisdell (1991) pointed out that the SMS approach is in fact more supportive of preservation than the CBA rule, considering that the size of B_P is likely to be though it is normally believed to be greater than B_D.

Ecological approach

Ecological approaches to project evaluation stress that ecological consequences and sustainability must be taken into account seriously before making irreversible development decisions. This notion of sustainable development was an important agenda item at the 1992 Earth Summit in Rio de Janeiro, Brazil. Economists concerned about sustainable development tend to argue that the conventional CBA is a mere monetary analysis. From the viewpoint of the extremely strong sustainability conditions (Daly 1980; 1996), however, sustainable development is also dangerous, and the term is an oxymoron, and that mankind must be saved from this dangerous path.

Multi-criteria analysis

Multi-criteria analysis (MCA) – also known as multiple objective decision-support systems (MODSS) – is a structured framework for evaluation of various policy options across multiple objectives, which are often conflicting with one another. In this technique, the various criteria that reflect the multiple objectives under consideration, and the decision options, are listed in an effect table. This is an $m \times n$ matrix with m criteria and n options. The data used are either qualitative or quantitative. In order to determine the relative value of each decision option, importance weights are attached to each criteria. The qualitative data are ranked or put on a scale, and then combined with the quantitative data. The preferred option is computed by a series of mathematical techniques (Cameron 1992; Hajkowicz *et al*. 2000). Bennett (2000) noted what differentiates MCA from CBA is that in CBA the weights used are the per unit values derived for each of the impacts, whereas the weights used in MCA are driven by the analyst.

In comparison to CBA, MCA does not have to involve the conversion of all costs and benefits associated with a project into monetary terms. Cameron (1992) pointed out that MCA has the advantage in avoiding the risk of spurious quantification of non-market values: that is, the difficulties of converting to dollar value can be avoided by leaving the results of qualitative assessments of environmental values in a qualitative form. Gurocak and Whittlesey (1998) came to the parallel view that public projects often have a variety of economic, ecological, social and political objectives, many of which cannot or perhaps should not be converted to monetary terms. It is not necessary, however, to leave aside environmental value in a qualitative form. Values which can be quantified by a suited valuation method can be included in the MCA process (Cameron 1992).

The common disadvantage of MCA is that this method requires decision makers to identify preference weights for the criteria involved in the decision process (Cameron 1992). In this context, MCA tends to be much dependent on the subjectivity of preference weight inputs from decision makers (Gurocak and Whittlesey 1998). Subjectivity itself is not necessarily undesirable but may cause an inconsistent

framework in making the unavoidable hard choices, diminishing the effectiveness of MCA.

9. DISCUSSION

Several techniques have been developing to estimate non-market values. Fairly obviously, it is necessary to seek the most appropriate technique. Harrison (2000) came to the point that although choice of valuation technique becomes complex in reality, simple statements can be made as a rough guide. When the task is to value the recreation benefits at a recreation site, TCM is likely to be appropriate. Nevertheless, it is doubtful whether TCM is appropriate for capturing the value of a specific characteristic of the recreation site. CVM needs to be considered when social welfare changes associated with an environmental change within a recreation site is to be estimated.

There are various categories of non-timber values associated with forestlands, falling under the headings of indirect use value, option value and existence value. Evidence from current Philippines experiences is that extreme deforestation results in not only a timber shortage, but also environmental problems in watersheds. Other problems including loss of human life due to increased flood severity, damage to crops and fisheries, reduction in the life of hydropower generators, which are caused by flooding, are also attributed in part to deforestation. Further, some may concern for the fauna and flora issues of deforestation, for example, loss of native animals. The damage function method or production function approach can be used to estimate damage to crops caused by deforestation. CVM or choice modelling is an appropriate method to value wildlife habitat in farm forestry. Harrison (2000) provided further insights into methods of estimating various non-wood benefits of farm forestry.

While the importance of non-market values is being increasingly realized, accuracy of valuation methods reviewed in this module remains a lingering problem. Reliability of value estimates might be the top criterion, from the viewpoint of policy makers, to judge with whether to include them into project appraisal. Thus, valuation researchers must continue to strive to refine existing valuation methods, or developing new ones, in a way of enhancing the reliability. However, as Harrison (1999) stressed, the ability to estimate non-market values precisely should be treated as a separate issue to the importance of integrating them in project appraisal.

Economists have devised various project evaluation methods. The conventional cost–benefit approach originated from the utilitarian philosophy of 'the greatest happiness of the greatest number'. The philosophical notion was incorporated into the Kaldor–Hicks compensation principle. A number of decisions on environmental projects have often been made according to the principle. The justification of the compensation principle is based on a potential Pareto improvement, in which economic welfare of a society is higher if the monetary value of gains exceeds the losses. Under the principle, the identity of the gainers and losers does not matter. On the other hand, the multi-criteria approach evaluates several policy options at the same time, using the idea of weighting various criteria. Further, ecological approaches stress that researchers must also take into account scientific evidence of the adverse impact of development over native flora and fauna, water quality and other ecological considerations. Ecological approaches are applicable particularly to the discrete land-use options. Among these evaluation methodologies, each of them works well in some cases and not well in other cases.

REFERENCES

Acharya, G. (1998), 'Valuing the Environment as an Input: The Production Function Approach', in M. Acutt and P. Mason, eds., *Environmental Valuation, Economic Policy and Sustainability*: *Recent Advances in Environmental Economics*, Edward Elgar, Cheltenham, pp. 63–78.

Aldred, J. (1994), 'Existence Value, Welfare and Altruism', *Environmental Value*, 3: 381–402.

Arrow, K.J. and Fisher, A.C. (1974), 'Preservation, Uncertainty, and

Irreversibility', *Quarterly Journal of Economics*, 88(2): 312–319.

Barbier, E.B. (1994), 'Valuing Environmental Functions: Tropical Wetlands', *Land Economics*, 70(2): 155–173.

Bateman, I. (1993), 'Valuation of the Environment, Methods and Techniques: Revealed Preference Methods', in R.K. Turner, ed., *Sustainable Environmental Economics and Management: Principle and Practice*, Wiley, Chichester, pp. 192–265.

Bateman, I.J. and Langford, I.H. (1997), 'Non-users' Willingness to Pay for a National Park: An Application and Critique of the Contingent Valuation Method', *Regional Studies*, 30(6): 571–582.

Bateman, I.J. and Turner, R.K. (1993), 'Valuation of the Environment, Methods and Techniques: The Contingent Valuation Method', in R.K. Turner, ed., *Sustainable Environmental Economics and Management: Principle and Practice*, Wiley, Chichester, pp. 120–191.

Bennett, J. (1995), *Economic Value of Recreational Use: Gibraltar Range and Dorrigo National Parks*, A report prepared for the Environmental Economics Policy Unit, Environmental Policy Division, NSW National Parks and Wildlife Service.

Bennett, J. (1996), 'The Contingent Valuation Method: A Post-Kakadu Assessment', *Agenda*, 3(2): 185–194.

Bennett, J. (2000), 'On values and their Estimation', *International Journal of Social Economics*, 27(7–10): 980–993.

Bennett, J., Gillespie, R., Powel, R. and Chalmers, L. (1996), 'Economic Value and Regional Economic Impact of National Parks, *Australian Journal of Environmental Management*, 3(4): 229–239.

Bishop, R.C. (1982), 'Option Value: An Exposition and Extension', *Land Economics*, 58(1): 1–15.

Blamey, R.K. (1996), 'Citizens, Consumers and Contingent Valuation: Clarification and the Expression of Citizen Values and Issue–Opinions', in W.C. Adamowicz, W. White and W.E. Phillips, eds., *Forestry, Economics and the Environment*, CAB International, Wallingford, pp. 103–133.

Cameron, J.I. (1992), 'Valuing the Environment: A Social Ecological Perspective', in M. Lockwood and T. de Lacy, eds., *Valuing Natural Areas: Applications and Problems of the Contingent Valuation Method*, The Johnstone Centre of Parks, Recreation and Heritage, Charles Sturt University, Albury, Australia, pp. 155–174.

Carson, R., Wright, J., Carson, N., Alberini, A. and Flores, N. (1995), *The Bibliography of Contingent Valuation Studies and Papers*, Natural Resource Damage Assessment Inc., La Jolla.

Clawson, M. (1959), *Methods of Measuring Demand for and Value of Outdoor Recreation*, Reprint No. 10, Resource for the Future, Washington, D.C.

Clawson, M. and Knetsch, J.L. (1966), *Economics of Outdoor Recreation*, The Johns Hopkins University Press, Baltimore.

Cooper, J. and Loomis, J. (1990), 'Pooled Time-Series Cross-Section Travel Cost Models: Testing Whether Recreation Behavior is Stable over Time', *Leisure Science*, 12(2): 161–171.

Daly, H.E. (1980), 'The Steady-State Economy: Toward a Political Economy of Biophysical Equilibrium and Moral Growth', in H.E. Daly, ed., *Economics, Ecology, Ethics: Essays Toward a Steady-State Economy*, W.H. Freeman and Company, San Francisco, pp. 324–356.

Daly, H.E. (1996), 'Sustainable Growth? No Thank You', in J. Mander and E. Goldsmith, eds., *The Case Against the Global Economy*: *And for a Turn Toward the Local*, Sierra Club Books, San Francisco, pp. 192–196.

Davis, D. and Tisdell, C.A. (1996), *Valuing Outdoor Recreational Sites: A New Approach*, Discussion Paper No. 206, Department of Economics, The University of Queensland, Brisbane.

Davis, R.K. (1963), *The Value of Outdoor*

Recreation: An Economic Study of the Maine Woods, unpublished Ph.D. thesis, Harvard University.

Diamond , P.A. and Hausman, J.A. (1993), 'On Contingent Valuation Measurement of Nonuse Values', in J.A. Hausman, ed., *Contingent Valuation: A Critical Assessment*, North Holland, Amsterdam, pp. 3–38.

Edwards, S.F. (1988), 'Option Prices for Ground Water Protection', *Journal of Environmental Economics and Management*, 15(4): 475–487.

Englin, J. and Mendelsohn, R. (1991), 'A Hedonic Travel Cost Analysis for Valuation of Multiple Components of Site Quality: The Recreation Value of Forest Management', *Journal of Environmental Economics and Management*, 21(3): 275–290.

Everitt, A. (1983), 'A Valuation of Recreational Benefits', *New Zealand Journal of Forestry*, 28(2): 176–183.

Fletcher, J.J., Adamowicz, W.L. and Graham-Tomasi, T. (1990), 'The Travel Cost Model of Recreation Demand: Theoretical and Empirical Issues', *Leisure Science*, 12(1): 119–147.

Freeman, A.M III. (1984), 'The Quasi-Option Value of Irreversible Development', *Journal of Environmental Economics and Management*, 11(3): 292–295.

Garrod, G. and Willis, K. (1992), 'The Amenity Value of Woodland in Great Britain: A Comparison of Economic Estimates', *Environmental and Resource Economics* 2(4): 415–434.

Gurocak, E.R. and Whittlesey, N.K. (1998), 'Multiple Criteria Decision Making: A Case Study of the Columbia River Salmon Recovery Plan', *Environmental and Resource Economics*, 12(4): 479–495.

Hadker, N., Sharma, S., David, A. and Muraleedharan, T. (1997), 'Willingness to Pay for Borivli National Park: Evidence from a Contingent Valuation'. *Ecological Economics*, 21: 105–122.

Hajkowicz, S.A., McDonald, G.T. and Smith, P.N. (2000), 'An Evaluation of Multiple Objective Decision Support Weighting Techniques in Natural Resource Management', *Journal of Environmental Planning and Management*, 43(4): 505–518.

Hanley, N. and Ruffell, R. (1993), 'The Valuation of Forest Characteristics', in W.C. Adamowicz, W. White and W.E. Phillips, eds., *Forestry and the Environment: Economic Perspective*, CAB International, Wallingford, pp. 171–197.

Hanley, N. and Spash, C.L. (1993), *Cost–Benefit Analysis and the Environment*, Edward Elgar, Cheltenham.

Harrison, S.R. (1999), 'Progress in Estimation of Intractable Non-market Vlaues', in S.B. Dahiya, ed., *The Current State of Economic Science*, Spellbound Publications, Model Town, Rohtak, India, Vol. 5, pp. 2719–2737.

Harrison, S.R. (2000), 'Total Economic Value of Multiple-use Farm Forestry and Non-market Valuation Techniques', in S.R. Harrison, J.L. Herbohn and K.F. Herbohn, eds., *Sustainable Small-scale Forestry: Socio-economic Analysis and Policy*, Edward Elgar, Cheltenham, pp. 50–64.

Harrison, S.R. (2001), *Valuation of the Environment*, unpublished paper, School of Economics, The University of Queensland, Brisbane.

Heyes, C. and Heyes, A. (1999), 'Willingness to Pay Versus Willingness to Travel: Assessing the Recreational Benefits from Dartmoor National Park', *Journal of Agricultural Economics*, 50(1): 124–139.

Hoagland, P., Kaoru, Y. and Broadus, J.M. (1995), 'A Methodological Review of Net Benefit Evaluation for Marine Reserves, Paper No. 27, *Environmental Economics Series*, The World Bank, Washington, D.C.

Imber, D., Stevenson, G. and Wilks, L. (1991), *A Contingent Valuation Survey of the Kakadu Conservation Zone*, RAC Research Paper No. 3, Vol. 1, Resource Assessment Commission, Canberra.

Knetsch, J.L. (1963), 'Outdoor Recreation

Demands and Benefits', *Land Economics*, 39(4): 387–394.

Knetsch, J.L. (1994), 'Environmental Valuation: Some Problems of Wrong Questions and Misleading Answers', *Environmental Values*, 3(4): 351–368.

Krutilla, J.V. (1967), 'Conservation Reconsidered', *American Economic Review*, 57(4): 777–786.

Mattsson, L. and Li, C.Z. (1994), 'How Do Different Forest Management Practices Affect the Non-timber Value of Forests? An Economic Analysis', *Journal of Environmental Management*, 41: 79–88.

Mitchell, R.C. and Carson, R.T. (1989), *Using Surveys to Value Public Goods: The Contingent Valuation Method*, Resources for the Future, Washington, D.C.

Morrison, M. (2001), 'Non-market Valuation Databases: How Useful Are They?', *Economic Analysis and Policy*: 31(1): 33–56.

Narain, U. and Fisher, A. (1995), 'Modelling the Value of Biodiversity Using a Production Function Approach: The Case of the Anolis Lizard in the Lesser and Greater Antilles', in C.A. Perrings, K.G. Mäler, C. Folke, C.S. Holling and B.O. Jansson, eds., *Biodiversity Conservation*: *Problems and Policies*, Dordrecht, Kluwer Academic Publishers, pp. 115–125.

Pearce, D.W. and Turner, R.K. (1990), *Economics of Natural Resources and the Environment*, Harvester Wheatsheaf, London.

Portney, P.R. (1994), 'The Contingent Valuation Debate: Why Economists Should Care', *The Journal of Economic Perspectives*, 8(4): 3–17.

Randall, A. (1994), 'A Difficulty with the Travel Cost Method', *Land Economics*, 70(1): 88–96.

Sagoff, M. (1988), *The Economy of the Earth: Philosophy, Law and Environment*, Cambridge University Press, Cambridge.

Smith, V.K. and Krutilla, J.V. (1982), 'Toward Reformulating the Role of Natural Resources in Economic Model', in V.K. Smith and J.V. Krutilla, eds., *Explorations in Natural Resource Economics*, The Johns Hopkins University Press, Baltimore, pp. 3–43.

Solow, R.M. (1993), 'Sustainability: An Economist's Perspective', in R. Dorfman and N.S. Dorfman, eds., *Economics of the Environment: Selected Readings*, 3rd edn., Norton, New York, pp. 179–187.

Tisdell, C.A. (1991), *Economics of Environmental Conservation: Economics for Environmental and Ecological Management*, Elsevier, Amsterdam.

Tyrvainen, L. and Miettinen, A. (2000), 'Property Prices and Urban Forest Amenities', *Journal of Environmental Economics and Management*, 39: 205–223.

Weisbrod, B.A. (1964), 'Collective-Consumption Services of Individual-Consumption Goods', *Quarterly Journal of Economics*, 78(3): 471–477.

Willis, K.G. and Garrod, G. (1995), 'Transferability of Benefit Estimates', in K.G. Willis and J.T. Corkindale, eds., *Environmental Valuation*: *New Perspectives*, CAB International, Wallingford, pp. 191–212.

Wood, S. and Trice, A. (1958), 'Measurement of Recreation Benefits', *Land Economics*, 34: 195–207.

14. Estimation of Non-market Forest Benefits Using Choice Modelling

Jungho Suh

This module is concerned with some fundamental features and conditions of choice modelling applications to non-market valuation. Choice modelling is an advance on the contingent valuation method (CVM), and is creating strong interest among researchers, and has much potential for non-market valuation in multiple use forestry. To undertake a choice modelling application for estimating non-market values, potential practitioners need to understand the theoretical issues and practicalities involved in applying the technique. These include the statistical foundation of choice modelling, strict rules for the experimental design and ways of utilising the estimates. This module first outlines the characteristic features of choice modelling in terms of the method of evaluating resource use alternatives, compared with contingent ranking, contingent rating and CVM. Some basic assumptions and considerations needed in designing choice sets are then examined. Roles and rules of focus groups are next introduced. Some choice modelling applications made in forestry research are then briefly reviewed. Finally, ways of extrapolating welfare measures from a choice modelling application are reviewed. More complex statistical issues are explained in three appendices.

1. THE CHOICE SET FORMAT OF CHOICE MODELLING

Choice modelling originated from conjoint analysis, and is also a variation on contingent valuation. In comparison to CVM, conjoint analysis describes options by decomposing them into a number of attributes, and presents respondents with a choice between *j* available options ($j = 1, 2, \cdots, J$). This situation can be made quite realistic by mirroring actual market choice that may depend upon a number of attributes.

If the number of available options is too large, the full options are divided into several sets. Then, respondents are asked to rank, rate or choose their preferred combination from each set. Moreover, the number of sets can be increased to as many as each respondent can answer within a limited time. For this reason, it is said that one of the major advantages of conjoint analysis compared to CVM is that many options provide a large number of observations so that fewer respondents are required to yield results within acceptable confidence limits.

Conjoint techniques have been widely used in marketing studies dealing with market goods rather than non-market goods. Conjoint analysis is founded on the theory of consumer preference in an attempt to describe how consumers choose between similar products, for example, beers, coffees and soft drinks. Respondents are asked to rank or rate or choose from a set of multiple product profiles. Setting prices for the products was not necessarily the primary concern of conjoint analysis in marketing studies. In this sense, the fact that conjoint analysis was eventually developed to value non-market public goods can be dubbed a ‘paradigm change’ in the field of economic valuation.

The rationale of conjoint analysis applications for estimating environmental non-market values is that it is possible to estimate the amount that people are willing to pay to achieve a greater amount of one or more environmental attribute, given that the dollar cost is treated as one of the characteristics for a non-market good. In fact, the price factor does not represent an inherent attribute of a commodity under consideration. Rather, the price presents dollar costs that are traded off for proposed changes in attribute levels. This is why Mitchell and Carson (1989) classified conjoint analysis as a ‘hypothetical and indirect’ approach.

In contingent ranking, respondents rank three or more options from most to least preferred. In the contingent rating application, respondents are asked to rate each option separately on a given rating scale instead of ranking the options. For example, consider the case of a protected forest area as illustrated in Table 1, where z_k^j represents the *k*th attribute of the *j*th option and z_p^j is the price factor. The protected area is here defined as a combination of attributes, each of which may take various levels. If a respondent prefers the *j*th option $\{z_1^j, z_2^j, \cdots, z_k^j\}$ to the other options, a higher ranking or rating is assigned to the *j*th option. Compared to contingent ranking, contingent rating contains cardinal information. In choice-based conjoint analysis, respondents are asked only to choose their highest preference from among several options – for example the set of choices presented in Table 1. Carson *et al.* (1994) called this method 'choice modelling' to distinguish it from contingent ranking or rating. In some literature, the term 'environmental choice experiments' is used rather than choice modelling, especially by the Canadian group of practitioners.

It is notable that a dichotomous CVM question is the same as a binary choice modelling one except for the pricing format (Bennett and Carter 1993; Roe *et al*. 1996; Stevens *et al*. 2000). Consider a choice modelling question with only one alternative (two options), as in Table 2, where respondents are asked whether to accept the new option, comparing to the current status option. Note that one of z_k represents the WTP amount (z_p). It can be seen that the question is virtually identical to that of a dichotomous CVM where respondents are asked whether they would be willing to pay z_p for the same change in z_k.

Choice modelling has an advantage over contingent rating in the sense that the former is free of metric bias with which the latter is plagued. Metric bias occurs when a respondent values an amenity according to a different metric or scale than the one intended by the researcher (Mitchell and Carson 1989). This bias also relates to the problems of interpersonal comparison of cardinal measurement of utility (Morrison *et al*. 1996). Since individual rating scales in contingent rating applications reveal only relative value between the respondents, it is necessary to assume that rating scales being used are consistent across individuals (Rolfe and Bennett 1996). Similarly, contingent ranking suffers inconsistent ordinal measurement of utility across individuals.

2. ASSUMPTIONS AND CONSIDERATIONS IN THE EXPERIMENTAL DESIGN

Any type of market or non-market good can be described by a range of characteristics. In applications of choice modelling, a number of hypothetical profiles are created by combining distinct levels of attributes, which must represent a wide range of characteristics of the object being valued. The number of attributes and their levels determines the total number of distinct profiles. A full factorial design includes all combinations of the attribute levels, where every level of a given attribute is combined with all levels of every other attribute. In general, if there are *m* factors and *n* levels of each, n^m unique combinations can be made. If factor space S has *k* factors, S = $(z_1, z_2, \cdots, z_k)$, and each factor z_k has L_k possible levels, then S has $L_1 \times L_2 \times \cdots \times L_k$ possible combinations.

Question formats based on complete factorial design quickly become impractical due to the cost of administering the survey, not to speak of the respondents' confusion and fatigue, as the number of either factors or levels of the individual factors increases. Indeed, in many cases, a choice modelling researcher is simply unable to conduct a survey using a large number of profiles. Hence, the researcher is forced to adopt a fractional factorial design, where only some of the combinations of factor levels are included. In choice modelling practice, a selected fractional factorial design is again broken into a number of separate choice sets. Rolfe and Bennett (1996) noted that the number of choice sets should not be too onerous for a single respondent. They suggested that choice sets be divided into manageable blocks, with each block allocated to a sub-sample of respondents.

Table 1. A choice format with several scenarios with various levels of attributes

Option (j)	Attribute (z_k)				
	z_1	z_2	…	z_k	z_p
1	z_1^1	z_2^1	…	z_k^1	z_p^1
2	z_1^2	z_2^2	…	z_k^2	z_p^2
:	:	:	:	:	:
J	z_1^J	z_2^J	…	z_k^J	z_p^J

Table 2. A binary choice modelling question format

Option	Attribute (z_k)				
	z_1	z_2	…	z_k	z_p
Current situation	z_1^0	z_2^0	…	z_k^0	z_p^0
New option	z_1^1	z_2^1	…	z_k^1	z_p^1

Louviere (1988) warned that one must be cautious of fractional designs because a strictly additive utility function, known as the orthogonality assumption, underlies choice modelling. The orthogonality assumption means that choice modelling estimates only the main effect of each attribute on the overall utility, assuming that all interaction effects between attributes are zero. Thus, choice modelling questions should be designed to comply with the orthogonality assumption. Further explanation on the orthogonal experimental design is provided in Appendix A.

For designing a choice modelling questionnaire, a few other considerations are required. First, the number of choices available in a choice set should be manageable from the viewpoint of respondents. Hanley and Spash (1993) warned that a choice set with five or more choices for the contingent ranking method would make respondents confused and threaten the reliability of the information gained from the contingent ranking study. In choice modelling applications, the confusion would be less evident because respondents are not asked to rank the options in order of their preferences. Nevertheless, a large number of options in a choice set may cause the same problems. It is notable that the minimum number of options that should appear in each choice set is three.

Second, extreme care is called for regarding the levels and range of the payment variable. Lareau and Rae (1989, pp. 729–730) in an empirical study of the contingent ranking technique warned that “if prices are too low, respondents order options by focusing mainly on the environmental attributes, while if prices are too high, respondents order options according to the price attribute.” This warning is applicable to choice modelling studies. Respondents would choose an option by focusing mainly on the environmental attributes if prices are too low and by focusing on the financial attribute if prices are too high. By the same token, Rae (1983) called attention to the price range. Too small a range may result in underestimating the values of other attributes when a large number of respondents, who are willing to pay more than the maximum, exist. Similarly, too wide a price range with only a small number of price levels may effectively force a trade-off, but lessen the chance of obtaining precise estimates within the range.

Third, too many attributes cause respondents to state their preferences by

'indicator attributes'. The indicator attribute effect occurs when respondents face complicated combinations of attributes, and then use a single attribute to indicate what would happen to the other attributes (Morrison *et al*. 1997), rather than evaluating each option by weighing up the levels of each attribute. On the other hand, too few non-monetary attributes may generate payment vehicle bias and substantially offset the effects of de-emphasising the payment mode (Morrison *et al*. 1996). Payment vehicle bias occurs when the payment vehicle employed influences how an individual responds to questions about WTP.

Bennett (1999) stressed that the payment vehicle must be compulsory and have broad coverage so as to be relevant to all respondents, as used in Rolfe *et al*. (2000). The rationale for presenting the compulsory payment vehicle is that the use of a donation as a payment vehicle encourages strategic responses. Therefore, it is argued that respondents must be convinced that they will be called upon to actually pay the amount they agree to pay when they choose options from choice sets.

The compulsory payment vehicle is not always desirable in the sense that it may also cause strategic responses. Morrison (1999) stressed that if the payment scenario is not acceptable, people may simply refuse to be interviewed in protest against the payment vehicle, take the survey less seriously or ignore the amount of payment. That is, people may reject the notion that they should have to pay for hypothetical new improvements in the quality of natural resources because they have paid taxes for the environmental improvement, even though respondents obviously derive utility from the environmental improvement. Flatley and Bennett (1994) reported that Lockwood *et al*. (1993), in a pilot study of a CVM survey to value areas of Gippsland forest, found a strong anti-government sentiment emanating from publicity over the troubles being experienced by various state-controlled financial institutions. Consequently, the forest valuation study had to use contributions to an independent trust as the payment vehicle in order to avoid the public's sentiment against the imposition of new taxes influencing WTP bids.

3. ROLES AND RULES OF FOCUS GROUP SESSIONS

Focus group sessions are a commonly used tool in psychology and are often regarded as a crucial step in shaping the market strategy for products. Krueger (1988) defined a focus group meeting as a carefully planned discussion designed to obtain perceptions on a specific area of interest in a permissive, non-judgmental and non-threatening atmosphere. The meeting with a small group of people is initiated and guided by preferably an experienced facilitator. Leading a focus group may require the combined skills of an ethnographer, a survey researcher and a therapist. Morgan (1988) pointed out that the main advantage that focus groups offer is the opportunity to observe a large amount of interaction on a topic in a limited period of time. In comparison with directive interview that is dominated by the interviewer, members of a focus group are allowed to talk without setting boundaries or providing cues for potential response categories. The moderator is not supposed to take the lead, but is expected to allow the advantages of non-directive interviews that use open-ended questions. The facilitator's skilled control over focus groups is, therefore, a key to successful meetings.

Krueger (1988) and Bernard (1995) suggested that a focus group be characterised by homogeneity, but ideally composed of strangers who do not know each other and will likely not ever see each other again. Homogeneity is sought in terms of occupation, social class, educational level, age and family characteristics, so as to avoid some mixes of participants that do not work well because of limited understanding of other lifestyles and situations. For example, participants will be inhibited by those whom they perceive more experienced, knowledgeable or better educated. Krueger (1988), however, pointed out that sufficient variation among participants to allow for contrasting opinions is helpful to obtain the contrast and variation that spark lively discussions.

Rolfe and Bennett (1995) and Blamey

(1998) pointed out that focus group sessions are an integral part of stated preference methods, in which underlying theory of consumer behaviour is linked with psychology. Focus groups are routinely used in the developmental phase of CVM studies. Rolfe and Bennett (1995, p. 3) summarised the major roles of focus groups in an environmental valuation exercise. These are:

1. establishing the overall framework and characteristics of the good in question including the relationship to other goods and applicable institutional settings;
2. ascertaining the extent of knowledge that people have about particular goods, and the ways in which they value those goods;
3. identifying and describing the major attributes that people consider when valuing particular goods; and
4. establishing appropriate trade-offs or WTP amounts associated with changes in the particular goods.

Running focus groups is even more vital in choice modelling studies, given that a choice modelling questionnaire tends to be much more complex than a CVM questionnaire. Through focus groups, choice modelling practitioners can determine which attributes are relevant to participants and test whether the information presented is appropriate; whether the main issues are communicated effectively and whether participants understand the choice modelling process sufficiently to handle choice sets; and whether the upper and lower bounds for the levels of the financial attribute are adequate (Morrison *et al*. 1997).

Focus group participants are expected to reveal the beliefs and attitudes that they actually consider when making environmental decisions. For this reason, Rolfe and Bennett (1995) argued that focus group sessions have to be involved from the questionnaire design stage rather than from the testing stage. In practice, focus groups are often used for testing a draft questionnaire and designing the final questionnaire.

One and a half to two hours is reasonable to run a focus group meeting for stated preference studies (Morrison *et al*. 1997). Stewart and Shamdasani (1990) noted that taking part in a focus group is a time consuming activity for participants in practice. Spending two or more hours talking to a group of strangers is not likely to be viewed as an appealing prospect. This is more likely the case if one has worked all day. They noted that a variety of incentives may be used to encourage participation, and monetary incentives are commonly used to induce individuals to spend time in a focus group.

It is recommended that a focus group have five to 10 members plus a moderator (Rolfe and Bennett 1995; Morrison *et al*. 1997). The group size is determined by two factors: it must be small enough for everyone to have opportunity to share insights yet large enough to provide diversity of perceptions (Krueger 1988). Bernard (1995) pointed out that if a group is too small, the group can be dominated by one or two 'loudmouths', and if the number of group members is beyond 10, it becomes difficult to manage the group. Krueger (1988) mentioned that a focus group is typically composed of seven to 10 people, but the size can range from as few as four to as many as 12. He noted that small focus groups or mini-focus groups with four to six participants are becoming increasingly popular because the smaller groups are easier to recruit and host, and more comfortable for participants.

4. CHOICE MODELLING APPLICATIONS IN THE NATURAL ENVIRONMENT INCLUDING FORESTRY

There have been many choice modelling applications to recreation studies or environmental conservation research. Adamowicz *et al*. (1994) analysed the welfare impacts of changes in a set of attributes of recreational fishing sites. Hanley *et al*. (1998) focused on the identification and valuation of key attributes affecting forest choice by conservation-oriented recreational users and non-users. Recent choice modelling applications dealt with urban tourists' portfolio choices of destination and transportation components (Dellaert *et al*. 1997), the environmental values of water supply options of a river

(Blamey *et al*. 1999), improved wetland quality (Morrison *et al*. 1999) and choice of recreational theme parks (Kemperman *et al*. 2000). Examples of choice modelling applications to forestry studies include valuing international rainforests (Rolfe and Bennett 1996; Rolfe *et al*. 1997) and evaluating a tree clearing policy (Rolfe *et al*. 2000).

Rolfe and Bennett (1996) used choice modelling to estimate demand by Australians for rainforest conservation in overseas countries. This choice modelling study was undertaken in Brisbane in 1995. The key attributes that were used in the survey were location, rarity, effect on local people, potential for future visits, size and possession of special features. The location attribute was varied across six locations, while each of the other attributes varied across three levels. There were six options in each choice set. Each respondent was presented with nine choice sets. This study demonstrated how it is possible to derive estimates of value for overseas rainforests.

Rolfe *et al*. (2000) conducted a choice modelling experiment to estimate the values held by Brisbane residents for both environmental and social factors associated with tree clearing in the Desert Uplands region of central-western Queensland. The implication of changes to tree clearing regulations were described in terms of six attributes as presented in Table 3. Respondents were presented with a *status quo* option (Option A) and two options for increased preservation. A series of eight choice sets were presented to each respondent. This study found that Brisbane households hold substantial protection values for native vegetation in the Desert Uplands.

The choices made of interview are typically entered into a spreadsheet for subsequent processing. Suppose that a respondent chose Option 3 from the list in Table 3. With this choice observation, three records can be entered as presented in Table 4 on a typical spreadsheet such as Microsoft Excel. Choice modelling requires that the dataset be arranged with a line of data for each option. If choice modelling data are collected from 10 respondents, and each respondent is asked to tick eight different choice sets, 240 records are generated. In choice modelling applications, sample sizes are generally about 300. Staying with eight choice sets per respondent, this means that 7,200 records are obtained from a choice modelling survey.

The dataset can be analysed with a statistical package such as Limdep 7.0, a specialised program for the estimation of qualitative response models and limited dependent variable models.

Table 3. A practical example of choice set

Implications	Option A (current guidelines)	Option B	Option C
Levy on your income tax	None	$60	$20
Income lost to the region ($ M)	None	5	10
Jobs lost in the region	None	15	40
Number of endangered species lost to region	18	8	4
Reduction in population size of non-threatened species	80%	75%	45%
Loss in area of unique ecosystems	40%	15%	28%

Source: Rolfe *et al*. (2000, p. 12).

5. WELFARE MEASURES EXTRAPOLATED FROM CHOICE MODELLING ESTIMATES

Discrete choice models have historically been the main models used in choice modelling studies. What is meant by discrete choice models is that those models in which the dependent variable takes discrete values. The simplest of these models is the binary choice model in which the dependent variable *y* takes the value of 0 and 1 (Maddala 1983). An example of this is presented in Table 4: *y* is defined as 1 if the respondent chooses Option 1, *y* is defined as 0 if the respondent does not choose the option, and so forth.

Table 4. An example of choice modelling data entered on a spreadsheet

Record no.	Variable						Choice (*y*)
	Levy	Regional income	Jobs	Endangered species	Population size	Ecosystems	
1	0	0	0	18	80	40	0
2	60	5	15	8	75	15	0
3	20	10	40	4	45	28	1

For a choice set with *J* alternatives, an estimated discrete choice model can be expressed as:

$$L_j = \ln\left(\frac{P_j}{P_{j'}}\right) = V(z_k^j) = b_0 + b_1 z_1^j + \cdots + b_K z_K^j$$

$$j = 1, 2, \cdots, J \qquad (1)$$

where z_k^j refers to the *k*th attribute of the *j*th option, one of z_k represents the price, and b_k is the weight or coefficient associated with an independent variable z_k. Equation 1 indicates that the logarithm of probabilities that a particular choice will be made can be represented by the systematic component of utility of the *j*th option. This is assumed because choice modelling aims to yield the unbiased estimate of the main effects of those attributes on utility only (Adamowicz *et al.* 1994). The independence property that all the attributes must be independent of one another is the most important feature pertaining not only to choice modelling but to all types of conjoint analysis. Further details of the discrete choice model are presented in Appendix B.

The independence of irrelevant alternatives (IIA) is inherently assumed for logit models. The IIA assumption implies that the odds of choosing the *j*th option in relation to one of the other options must be constant regardless of whatever other options are present (Louviere and Woodworth 1983). Violations of the IIA property may occur in applications where options are close substitutes for one another. The property is also implausible when there exists heterogeneous tastes (Morrison *et al.* 1999). When the IIA property is invalid, parameter estimates from the relevant discrete choice model will be biased. The IIA assumption is a substantive restriction on the generality of the logit models. Thus, testing for presence of the IIA property is a standard empirical procedure in obtaining a valid empirical logit model. Appendix C provides commonly used methods of detecting the IIA violations.

The estimated parameters of a discrete choice model provide a basis to compute the trade-offs between dollars and environmental quality. Compared to CVM, a single choice modelling exercise can separately and simultaneously estimate the coefficients of all factors involved in choice sets. For example, the WTP to avoid each 1% reduction in non-threatened species can be obtained from the choice modelling analysis using the Rolfe *et al.* (2000) dataset. Consequently, compensating surpluses for numerous hypothetical options can be extrapolated. Compensating surplus estimates are put to use as paramount welfare measures associated with characteristics of an environmental resource. In contrast, CVM generally values a composite of non-monetary factors per survey. It is, however, not correct to argue

that the contribution of various attributes to the value of an environmental good cannot be estimated separately with CVM (Scarpa 2000). It can be achieved with adequately designed interviews. In particular, CVM practitioners can prepare a number of distinct profiles and present one profile to one respondent at a time, though this is costly in obtaining as many observations as in choice modelling.

Implicit prices for unit changes within attributes

The implicit price means the marginal rate of substitution between a non-monetary attribute and the monetary factor. The implicit price is also referred to as 'part-worth' or the point estimate of WTP. The implicit price is derived holding constant all other parameters except for the parameter of the attribute for which the implicit price is being computed. Mathematically, let one of z_k represent the price factor z_p. Holding $\Delta V_j = 0$ yields:

$$\Delta V_j = \Delta \sum b_k z_k^j + \Delta b_p z_p = 0 \qquad (2)$$

Assume that the level of an attribute changes from a base level to another in terms of environmental quality and levels of $k-1$ other attributes remain unchanged. The ratio of the particular attribute coefficient to the price coefficient measures the WTP for the hypothetical change of the specific attribute, because utility of the non-monetary attributes is indirectly related to the price factor. The point estimate of WTP with respect to z_k is obtained by:

$$-\frac{dz_p}{dz_k} = -\frac{b_k}{b_p} \qquad (3)$$

For an improvement in environmental quality, b_k is greater than zero. Thus, the ratio is expected to be positive because the sign on b_p, the price parameter, will be negative. Positive ratio values represent attributes that increase utility, whereas negative ratio values (i.e. $b_k < 0$) represent attributes that reduce utility. It is important to note that implicit prices are not appropriate for use as measures of overall welfare created by an environmental change (Morrison *et al*. 1999). The reason is that an implicit price is computed under the assumption that change in utility is equal to zero, as expressed in Equation 2. Compensating surpluses need to be calculated, in order to understand comprehensively the environmental attitudes of stakeholder groups towards a specific environmental option. The calculation of the compensating surplus for an environmental change involves implicit prices of model attributes multiplied by proposed units of changes within each attribute, and the alternative specific constant estimated for the specific option.

Deriving the compensating surplus for environmental improvement

Two types of compensation measures of consumer welfare may be distinguished. 'Compensating variation' is defined as the amount of income that, if given to an individual, would make the individual indifferent in utility terms between the initial and subsequent combinations of the money income and the consumption level of goods. Instead of 'compensating variation', the term 'compensating surplus' is used to describe the welfare change in terms of income, in the situation where individuals are not free to choose the consumption level of goods – for example, public goods such as air quality. In short, compensating surplus measures the difference in utility between two options — namely, the change in income that would make the utility level of an individual indifferent between the initial and subsequent options.

Using indirect utility functions, compensating surplus (CS) for beneficial environmental changes can be defined as in the following equation:

$$\begin{aligned} V_C(E_0, M_0) &= V_C(E_1, M_1) \\ &= V_C(E_1, M_0 - \text{CS}) \\ &= V_N(E_1, M_0) - \text{CS} \end{aligned} \qquad (4)$$

where E indicates a particular environmental good in consideration and M denotes an individual's total money income. The individual can purchase any combination of market goods and services including other environmental goods and services with total money income M, where $M_0 > M_1$. The person's other option is assumed to be an environmental improvement denoted by E, where $E_1 > E_0$.

V_C and V_N represent the utility levels of the current situation and alternative state, respectively. From Equation 4, $M_1 = M_0 - CS$, and therefore $CS = M_0 - M_1$. Unlike when deriving implicit prices, the utility level is not required to remain the same. Hence, compensating surplus represents marginal WTP for a change from the current bundle of commodity (E_0, M_0) to (E_1, M_0).

The welfare estimates are obtained by the difference in utility between two options, which is scaled by the marginal utility of income to determine compensating surplus:

$$\begin{aligned} \text{CS} &= -\frac{1}{b_p}(V_C - V_N) \\ &= -\frac{1}{b_p}\left(\sum b_k\, z_1^C - \sum b_k\, z_1^N\right) \end{aligned} \qquad (5)$$

where z_k^1 indicates the kth attribute (except for the price factor). Equation 5 has often been used for estimating compensating surplus for a new level of environmental quality attached to a single site (e.g. Boxall *et al*., 1996; Blamey *et al*., 1998; Morrison *et al*., 1999).

6. SUMMARY

Choice modelling is grounded on the discrete choice model, which can produce a strictly additive utility function. From the estimated utility function, one can ultimately extrapolate consumer surplus for proposed hypothetical environmental changes. For this reason, discrete choice models have recently been applied to value non-market environmental goods such as forests and wetlands around the world in a number of benefit measurement studies.

In designing a choice modelling questionnaire, keeping valid the orthogonality assumption is vital. In the modelling phase, testing for the IIA property is a standard empirical procedure. As well, given the diversity of beliefs, attitudes, interests, knowledge and other factors that might exist in populations, running focus groups sessions is clearly a beneficial step to designing and testing a non-market valuation questionnaire. Focus group sessions are essential particularly for choice modelling practices, considering that the questionnaire format with a number of choice sets is relatively new and complex to potential respondents. Using focus group sessions, a choice modelling practitioner can select design attributes, finalise choice formats, frame the hypothetical market context and identify attitudes that influence participants' responses.

In choice modelling applications to environmental resources, environmental quality as well as monetary factors are included as attributes of the options in a choice set. Thus, choice modelling allows one to obtain compensating surplus estimates so that one can account for the welfare change generated by a bundle of changes in environmental attributes. It is also possible to determine the relative importance of these attributes to people in making their choices.

In choice modelling practice, respondents are offered hypothetical combinations of attributes associated with a set of scenarios. By downplaying the payment factor, choice modelling is much more aptly able to estimate preferences with the effect of reducing the payment vehicle bias relative to CVM. The mechanical framework of a discrete choice model allows researchers to disaggregate the utility function attached to a particular good, which is a compelling advantage of choice modelling over CVM. A choice modelling exercise can produce point estimates of WTP for changes in the attribute levels varied in the choice sets as well as compensating surplus estimates for a conceptually unlimited number of underlying scenarios. Choice modelling can avoid problems posed with contingent ranking and contingent rating. These latter methods face many critics because of theoretical problems such as lack of error theory, and measurement bias that occurs because of interpersonal comparison of utility.

REFERENCES

Adamowicz, W and Louviere, J. and Williams, M. (1994), 'Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities', *Journal of Environmental Economics and Management*, 26(3): 271–292.

Aldrich, J.H. and Nelson, F.D. (1984), *Linear Probability, Logit and Probit Models*, Sage University Paper No. 45, Sage Publications, Beverly Hills, California.

Ben-Akiva, M. and Lerman, S.R. (1985), *Discrete Choice Analysis: Theory and Application to Travel Demand*, the MIT Press, London.

Bennett, J. (1999), *Some Fundamentals of Environmental Choice Modelling*, Choice Modelling Research Report No. 11, School of Economics and Management, The University of New South Wales, Canberra.

Bennett, J.W. and Carter, M. (1993), 'Prospects for Contingent Valuation: Lessons from the South-East Forests', *Australian Journal of Agricultural Economics*, 37(2): 79–93.

Bernard, H.R. (1995), *Research Methods in Anthropology: Quantitative and Quantitative Approaches*, Altamira Press, Walnut Creek, Canada.

Blamey, R. (1998), 'Contingent Valuation and the Activation of Environmental Norms', *Ecological Economics*, 24(1): 47–72.

Blamey, R.K., Bennett, J.W., Morrison, M.D., Louviere, J.J. and Rolfe, J.C. (1998), *Attribute Selection in Environmental Choice Modelling Studies: The Effect of Causally Prior Attributes*, Choice Modelling Research Report No. 7, School of Economics and Management, The University of New South Wales, Canberra.

Blamey, R., Gordon, J. and Chapman, R. (1999), 'Choice Modelling: Assessing the Environmental Values of Water Supply Options', *The Australian Journal of Agricultural and Resource Economics*, 43(3): 337–357.

Boxall, P.C., Adamowicz, W.L., Swait, J., Williams, M. and Louviere, M. (1996), 'A Comparison of Stated Preference Methods for Environmental Valuation', *Ecological Economics*, 18(3): 243–253.

Carson R.T., Louviere, J.J., Anderson, D.A., Arabie, P., Bunch, D.S., Hensher, D.A., Johnson, R.M., Kuhfield, W.F., Steinberg, D., Swait, J., Timmermans, H. and Wiley, J.B (1994), 'Experimental Analysis of Choice', *Marketing Letters*, 5(4): 351–367.

Dellaert, B.G., Borgers, A.W. and Timmermans, H.J. (1997), 'Conjoint Models of Tourist Portfolio Choice: Theory and Illustration', *Leisure Sciences*, 19(1): 31–58.

Flatley, G and Bennett, J. (1994), 'The Value of Vanuatu Forest Protection to Australian Tourists', Vanuatu Forest Conservation Research Report No. 6, Department of Economics and Management, The University of New South Wales, Canberra.

Green, P.E. (1974), 'On the Design of Choice Experiments Involving Multifactor Alternatives', *The Journal of Consumer Research*, 1(2): 61–68.

Greene, W.H. (1997), *Econometric Analysis*, 3rd edn., Prentice-Hall, London.

Hanley, N. and Spash, C.L. (1993), *Cost–Benefit Analysis and the Environment*, Edward Elgar, Cheltenham.

Hanley, N., Wright, R. E. and Adamowicz, V. (1998), 'Using Choice Experiments to Value the Environment', *Environmental and Resource Economics*, 11(3–4): 413–428.

Hausman, J. and McFadden, D. (1984), 'Specification Tests for the Multinomial Logit Model', *Econometrica*, 52(5): 1219–1240.

Kemperman, A.D., Borgers, A.W., Oppewal, H. and Timmermans, H.J. (2000), Consumer Choice of Theme Parks: A Conjoint Choice Model of Seasonality Effects and Variety Seeking Behavior', *Leisure Science*, 22(1): 1–18.

Krueger, R.A. (1988), *Focus Groups: A Practical Guide for Applied Research*, Sage Publications, Newbury Park, California.

Lareau, T. and Rae, D. A. (1989), 'Valuing WTP for Diesel Odor Reductions: An Application for Contingent Ranking Technique', *Southern Economics Journal*, 55(3): 728–742.

Lockwood, M., Loomis, J. and DeLacy, T. (1993), 'A Contingent Valuation Survey and Benefit-Cost Analysis of Forest

Preservation in East Gippsland, Australia', *Journal of Environmental Management*, 38(3): 233–243.

Louviere, J.J. (1988), *Analysing Decision Making: Metric Conjoint Analysis*, Sage Publications, Newbury Park, California.

Louviere, J.J. and Woodworth, G. (1983), 'Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data', *Journal of Marketing Research*, 20(November): 350–367.

McFadden, D. (1974), 'The Measurement of Urban Travel Demand', *Journal of Public Economics*, 3: 303–328.

McFadden, D., Tye, W. and Train, K. (1977), 'An Application of Diagnostic Tests for the Irrelevant Alternatives Properties of the Multinomial Logit Model', *Transportation Research Record*, 637: 39–46.

Maddala, G.S. (1983), *Limited-dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge.

Mitchell, R.C. and Carson, R.T. (1989), *Using Surveys to Value Public Goods: The Contingent Valuation Method*, Resources for the Future, Washington, D.C.

Morgan, D.L. (1988), *Focus Groups as Qualitative Research*, Sage Publications, Newbury Park, California.

Morrison, M. (1999), *Rethinking Contingent Valuation: Ethics and Defensibility*, unpublished paper presented at the Australian and New Zealand Society for Ecological Economics Conference, Griffith University, Brisbane.

Morrison, M.D., Bennett, J.W. and Blamey, R.K. (1997), *Designing Choice Modelling Surveys Using Focus Groups: Results From the Macquarie Marshes and Gwydir Wetlands Case Study*, Choice Modelling Research Report No. 5, School of Economics and Management, The University of New South Wales, Canberra.

Morrison, M.D., Bennett, J.W. and Blamey, R.K. (1999), 'Valuing Improved Wetland Quality Using Choice Modelling', *Water Resources Research*, 35(9): 2805–2814.

Morrison, M.D., Blamey, R.K., Bennett, J.W. and Louviere, J.J. (1996), *A Comparison of Stated Preference Techniques for Estimating Environmental Values*, Choice Modelling Research Report No. 1, School of Economics and Management, The University of New South Wales, Canberra.

Pindyck, R.S. and Rubinfeld, D.L. (1991), *Econometric Models and Economic Forecasts*, 3rd edn., McGraw-Hill, New York.

Rae, D. (1983), 'The Value to Visitors of Improving Visibility at Mesa Verde and Great Smoky National Parks', in R.D. Rowe and L.G. Chestnut, eds., *Managing Air Quality and Scenic Resources at National Parks and Wilderness Areas*, Westview Press, Boulder, Colorado, pp. 217–234.

Roe, B. and Boyle, K. and Teisl, M.F. (1996), 'Using Conjoint Analysis to Derive Estimates of Compensating Variation', *Journal of Environmental Economics and Management*, 31(2): 145–159.

Rolfe, J. and Bennett, J. (1995), *International Demand for Rainforests: Using Focus Groups to Establish Valuation Frameworks*, Vanuatu Forest Conservation Research Report No. 9, School of Economics and Management, The University of New South Wales, Canberra.

Rolfe, J. and Bennett, J. (1996), *Valuing International Rainforests: A Choice Modelling Approach*, Vanuatu Forest Conservation Research Report No. 12, School of Economics and Management, The University of New South Wales, Canberra.

Rolfe, J., Bennett, J., Blamey, J. (2000), *An Economic Evaluation of Broadscale Tree Clearing in the Desert Uplands Region of Queensland*, Choice Modelling Research Report No. 12, School of Economics and Management, The University of New South Wales, Canberra.

Rolfe, J., Bennett, J., Louviere, J. (1997), *Choice Modelling – An Exercise in Environmental Valuation,* Conference Paper, the Australian Agricultural and

Resource Economics Society, Gold Cost, Australia.

Scarpa, R. (2000), 'Contingent Valuation Versus Choice Experiments: Estimating the Benefits of Environmentally Sensitive Areas in Scotland: Comment', *Journal of Agricultural Economics*, 51(1): 122–128.

Stevens, J.H., Belkner, R., Dennis, D., Kittredge, D. and Willis, C. (2000), 'Comparison of Contingent Valuation and Conjoint Analysis in Ecosystem Management', *Ecological Economics*, 32(1): 63–74.

Stewart, D.W. and Shamdasani, P.N. (1990), *Focus Group: Theory and Practice*, Sage Publications, Newbury Park, California.

Zhang, J. and Hoffman, S.D. (1993), 'Discrete-Choice Logit Models: Testing the IIA Property', *Sociological Methods and Research*, 22(2): 193–213.

APPENDIX A: THE ORTHOGONAL FRACTIONAL FACTORIAL DESIGN

The orthogonal fractional factorial design can be obtained by the systematic draw from the factorial of possible combinations. The requirement is satisfied when each level of one factor occurs with each level of another factor with proportional frequencies (Green 1974). Rolfe and Bennett (1996) recommended allocating consistent numbers of levels across attributes if possible for the simplicity of a symmetrical orthogonal array. However, nothing is technically or practically wrong with varying numbers of levels between attributes, which leads to an asymmetrical orthogonal array, as long as each level will occur an equal number of times within each attribute, in each individual choice set and throughout the entire set of alternatives.

Louviere (1988), Adamowicz *et al*. (1994) and Bennett (1999) emphasised that separate choice sets of alternatives should be treated independently. They suggested that this requirement be achieved by adding the 'no-change' option – i.e. the 'current situation' option – to each and every choice set. Each choice set is then evaluated without regard to previously presented choice sets. The provision of the 'current situation' option allows respondents to state that they would prefer not to purchase any of hypothetical alternatives presented in the choice set. Carson *et al*. (1994) called for caution about the possibility that respondents may use the 'no-purchase' option as a means to avoid making difficult decisions. In that case, adding the 'no-change' option to each choice set can adversely influence on the value estimates.

APPENDIX B: MATHEMATICAL FORMULATION OF CHOICE MODELLING – THE DISCRETE CHOICE MODEL

The discrete choice model is usually called the conditional logit model, which McFadden (1974) derived from random utility. The conditional logit model concerns the effects of choice-specific attributes on the determinants of choice probabilities. To formulate the conditional logit model, consider the case of J mutually exclusive and collectively exhaustive options labelled arbitrarily, where the numberings cannot be taken to indicate order or any magnitude. Respondents will be making comparisons between options with the utility of the jth option. The indirect utility of the jth option can be represented by:

$$V_j = V(z_k^j) + \varepsilon_j \tag{6}$$

where $V(z_k^j)$ is the systematic component of utility and ε_j is a random unobservable component. The systematic component is assumed to be the same for all observations while the random component is unique to each consumer. Assuming $E(\varepsilon_j) = 0$, the probability P_j that the jth outcome is observed is:

$$\text{Prob}\,[V(z_k^j) > V(z_k^{j'})] \quad j = 1,2,\cdots,J \text{ for all } j' \neq j \tag{7}$$

The systematic component $V(z_k^j)$ can be expressed as the sum of combinations of attributes given by:

$$\begin{aligned} V(z_k^j) &= b_0 + b_1 z_1^j + b_2\, z_2^j + \cdots + b_K\, z_K^j \\ &= \sum_{k=1}^{K} b_k\, z_k^j \equiv bZ \end{aligned} \tag{8}$$

where z_k^j refers to the kth attribute of the jth option, and one of z_k represents the price, b_k is the weight or coefficient associated with an independent variable z_k, and bZ is the linear combination of attributes with a single parameter vector b. It is notable that the same attributes appear in the utility function for every choice with varying levels within each attribute. This is not a requirement of conditional logit models, but a common feature of most choice modelling applications.

The probability of an outcome in Equation 7 is a linear function of independent variables. The problem with the linear probability model specification is that bZ is used to approximate a probability P_j that is limited between from 0 to 1, whereas bZ is not so constrained. That is, P_j is non-linearly related to z_k as well as to the b_k. This means that ordinary least square procedures cannot be used to estimate the parameters. To solve this problem, Aldrich and Nelson (1984) have demonstrated derivation of non-linear probability specifications. To begin with, it is necessary to take the logistic function as a non-linear transformation function of Equation 7:

$$P_j = \text{Prob}\,[y_j = 1 \mid V(z_k^j)] = \exp(\text{bZ}) \Big/ \sum_{j=1}^{J} \exp(\text{bZ}) \qquad (9)$$

where y_j is the index of the choice made given that the random utility of the choice is $V(z_k^j)$. Next, the upper bound P_j can be estimated by taking the ratio $P_j / P^{j'}$. The ratio must be positive, and since P_j and $P^{j'}$ are constrained between 0 to 1, the ratio has no upper bound. The lower boundary of zero can be eliminated by taking the natural logarithm, ln (P_j/ Pj'), the value of which can be any real number from negative to positive infinity:

$$L_j = \ln\left(\frac{P_j}{P_{j'}}\right) = V(z_k^j) = b_0 + b_1 z_1^j + \cdots + b_K z_K^j \qquad (10)$$

The ultimate goal of applying choice modelling is to estimate the coefficients (i.e. b_k) from the logit model. The logit model represented by Equation 10 defines the logarithm of probabilities that a particular choice will be made. Yet the coefficients estimated also directly relate to the utility (Rolfe and Bennett 1996). The higher the utility of a particular attribute level in Option *j*, the higher the odds ratio – that is, the higher the probability that a respondent will choose the particular option. Again, each coefficient estimated represents the marginal contribution of an attribute to overall utility.

Logit models are estimated by the maximum likelihood technique. This can be achieved by using a non-linear maximisation program such as Limdep. The maximum likelihood estimation procedure has a number of desirable statistical properties. Among those, two aspects are notable (Pindyck and Rubinfeld 1991). First, parameter estimators are asymptotically consistent and efficient. Second, all parameter estimators are known to be normally distributed, so that the analog of the regression *t*-test can be applied. One can then apply the likelihood ratio test to determine whether a logit model provides an appropriate model specification.

APPENDIX C: TESTING FOR 'INDEPENDENCE OF IRRELEVANT ALTERNATIVES'

When an independence of irrelevant alternatives (IIA) violation is found for a conditional logit model, Morrison *et al.* (1999) suggested including socio-economic characteristics, attitudinal variables or questionnaire evaluation variables in the model. For the conditional logit model, these variables can be included in one or both of two ways. The first is through interactions with the alternative specific constants. These interactions reflect the effect of these variables on the choice probability that a respondent will choose either Option 2 or Option 3. The second is by interactions with the attributes in the choice sets. These interactions indicate that the non-attribute variables can modify the effects of attributes on the choice probability of an option. The interaction effects between non-attributes (e.g. socio-economic characteristics of individuals) and choice specific dummy variables or between non-attributes and attributes are not restricted by the orthogonal design.

If a conditional logit model containing interaction terms of non-attribute variables is found to violate the IIA property, it becomes necessary to estimate a more complex model that relaxes part of the IIA assumption. An example of this type of model is the nested logit model. The IIA assumption in the nested logit model is relaxed because the options are divided into subgroups across which the variance is allowed to differ while maintaining homoscedasticity within the groups (Greene 1997). However, it is still necessary to test a nested logit model for the IIA property because the IIA violations can occur due to the presence of close substitutes within the same branch in a choice set.

In non-nested logit models, all elemental nodes are connected to the root of the tree as in Figure 1. In nested logit models, construct nodes are introduced. Assume there are three choices in each choice set. Option 1 is the 'no-change' option and appears in every choice set. Combinations of attribute levels for Options 2 and 3 vary across choice sets. Option 1 is labelled as the 'current situation', whereas Options 2 and 3 are labelled as 'new options'. The nesting occurs when an individual chooses first, for example between Option 1 (current situation) and Option 2 or 3 (new options), and then, conditional on not choosing Option 1, chooses between Option 2 and Option 3. A tree relevant to this situation can be specified as in Figure 2.

Various methods of detecting violations of the IIA property have been proposed (Ben-Akiva and Lerman 1985; Zhang and Hoffman 1993). A Hausman and McFadden (1984) test is often used. To carry out the test, a specific option throughout choice sets is excluded. A logit model with the restricted set of options but with the same model specification as the model with all choices included is then estimated. One can next check whether there is any change between the two models, and evaluate whether the new coefficients are sufficiently similar to the original ones, so as to satisfy the IIA property. The test statistic for checking the IIA assumption is distributed as a chi-square variable. This type of test can be carried out using the Limdep statistical package.

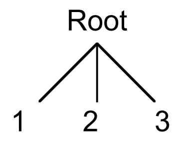


Figure 1. Tree diagram for a non-nested logit model (Model 1)

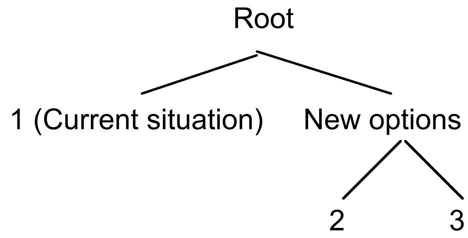


Figure 2. Tree diagram for a two-level nested logit model (Model 2)

Another test proposed by Hausman and McFadden (1984) is applicable to nested logit models, which allows the independence property to be tested directly. For example, IIA violations occur if Options 2 and 3 in Figure 2 have correlated unobservable factors.

In Figure 2, 'new options' is a construct node with parameter θ. This nested logit model assumes that the correlation between the errors of Options 2 and 3 is given by $(1 - \theta^2)$. If the hypothesis that $\theta = 1$ cannot be rejected, which can be tested using a likelihood ratio test comparing Models 1 and 2, then it can be concluded that Model 2 does not violate the IIA assumption. That is, testing of the IIA property is equivalent to determining whether the estimated value of θ is significantly different from 1.

There are other methods to test for the IIA assumption. Rolfe and Bennett (1996) suggested checking if a change in one attribute produces shifts in choice greater than expected. Adamowicz *et al*. (1994) noted that the cross effects should equal zero if the attributes are strictly independent, apart from the belief that the attributes of one option could influence the utility of another option. Morrison *et al*. (1999) introduced the universal logit model, which was also proposed by McFadden *et al*. (1977), to detect violations of the IIA assumption. In estimating a universal logit

model, called a ‘mother logit model’, the attributes of one option are entered into the utility function of a second option, and then the attributes of the second option into the utility function of the first option. If the model is found to be the more accurate model, the utility of one option can be considered to depend on the utility of other options, thus violating the IIA property.

15. Cost-benefit Analysis in Forestry Research

Steve Harrison

Cost-benefit analysis (CBA) is both a technique of analysis and a broad framework for economic evaluation of investments, for example, in development and infrastructure projects, and in research. It is applied to projects which run of a number of years, such that that discounting is used to bring costs and benefits to a comparable (present value) basis. The methodology is particularly relevant to forestry, which is a long-term investment, and forms one of the main socio-economic techniques in forestry research. As well, CBA is relevant to judging the desirability of research projects. In fact, some research funding agencies now require that a CBA of the research project be prepared as part of a grant application. This module first introduces the economic rationale for CBA. The procedures of DCF analysis and the various project performance criteria are then explained. The following section examines various practical complexities in applying CBA. Some remarks are then made on integration of non-market values in CBA. Concluding comments follow.

1. THE NATURE OF CBA

Cost-benefit analysis is used to examine the desirability of investment projects from an economic perspective. It is also relevant to estimating the desirability of infrastructure development, the economic impacts of new regulations, the benefits of conservation programs, and the expected payoff from research projects.

When CBA was first applied widely (it was mandated for watershed projects in the USA in the 1960s), the emphasis was on market or financial transactions, with noting but not valuation of 'intangibles'. Nowadays, an attempt is made to include all relevant costs and benefits of a project, in what is called *extended* or *social* CBA. The basic idea of social CBA is to determine whether investment projects are worthwhile from a social or taxpayer viewpoint, taking into account all the costs and revenues and positive and negative externalities they incur or generate.

In the field of forestry, social CBA is relevant to evaluation of public-sector funded projects such as community forestry projects, and government investment in industrial plantations and timber processing plants as well as forestry research programs. Each application of CBA is to some extent unique or one-off, but a number of common concepts and procedures are involved in applications. Prominent in these concepts are considerations of the value of money over time, identification of 'incremental cash flows' and definitions and estimation methods for cost and benefit flows. Discounted cash flow (DCF) analysis provides the computational method for deriving project performance criteria such as benefit-to-cost ratios, net present value and internal rate of return. When considering forestry plantations, the Faustmann formula for optimal economic rotation provides a framework for evaluation of plantation costs, revenues and externalities.

Cost-effectiveness analysis (CEA) is sometimes used as a simpler alternative to CBA. Here only the costs of a project are taken into account. Costs may be compared on a unit of physical benefit basis, e.g. dollars per amount of timber produced. CEA is also appropriate when alternative ways are being considered to achieve a constant output, e.g. to meet timber demands a decade from now. Many infrastructure projects have clearly defined outputs, which will be more-or-less constant regardless of how they are brought about provided design specifications are met, and hence are candidates for CEA. Public health and pest control projects could fall into this class. For example, CEA could be used to compare alternative strategies for control of a say a forest disease in a region or country. Once the decision has been made to eradicate a pest and disease, alternative methods could be compared by CEA. Note however that

CBA would be the relevant technique to use when determining whether it is in the national interest to eradicate the pest or disease in the first place.

Economists can play an important role by examining the economic implications of various alternative course of action, and pointing these out to policy-makers and administrators. Policy-makers (that is, the government) can combine this economic information with other information (e.g. about ecological values and community attitudes) and judgment and intuition to arrive at a decision. In other words, the input of the economics is information to augment other information already held or being gathered by policy- or decision-makers. Economists provide *decision-support* information. Usually, the economist does not make the decision, or take the consequences of that decision. But by pointing out the economic tradeoffs involved, they can assist managers in government or private enterprise to make better decisions.

Particularly where large public investment or conservation decisions are being considered, economists often work as part of a multidisciplinary team. To communicate effectively with other specialists, they need to have some appreciation of biology, sociology, planning, law and other areas.

2. THE ECONOMIC BASIS OF CBA

CBA is a widely used conceptual framework for applied microeconomic analysis and a practical methodology for project evaluation. Some form of CBA is invariably used to evaluate investments, although often in implicit rather than explicit form.

Economists typically take what is known as an *anthropocentric* or human-centred approach. Goods and services are valued in terms of what people are prepared to pay for them. The principle of *consumer sovereignty* is adopted, in which people are regarded as the best judge of what is good for them. This approach implies for example, that the preservation of species is important only insomuch as people place a value on species. An alternative approach would be a *biocentric* approach which would imply that species have value in their own right, regardless of whether the community wishes to spend money to protect them.

According to the Pareto principle, a project is desirable if it makes one or more people better off, and no-one worse off. In practice, projects invariable affect the income distribution in their area of influence, and typically some people become better off and some worse off. Hence a modified criterion – known as the Kaldor-Hicks criterion – was devised. This states that a project is desirable if it makes one or more people better off, and out of their gains they could potentially compensate any losers such that no-one is worse off. Note use of the term ‘potentially’ in this definition – the compensation may not actually take place. In general, it is the role of government to make adjustments to income distribution (such as taxes and welfare payments) to adjust to changes in an economy.

From a human-centred orientation, the desirability of a project may be assessed in terms of the changes in human welfare or satisfaction it creates. While welfare is often conceptualized in terms of what economists call ‘utility’, in practice this is normally measured in economic terms. Economic impacts of a project can be illustrated using market diagrams. Figure 1 is such a diagram, depicting the supply and demand curves (*S* and *D*) for a particular economic good (i.e. commodity or service). When q units of the good are produced and the price is p (in dollars per unit), the market is said to be in ‘equilibrium’. That is, at this price the market clears, and there is no shortage or deficit to drive the price up or down.

With reference to Figure 1, all consumers can purchase the good at price *p*, whereas some are prepared to pay higher prices (indicated by the demand curve for quantities less that *q*). The area *abc* represents the aggregate gain to consumers in terms what they would be prepared to pay but do not have to pay. This is called the *consumer surplus*. Similarly, some producers would be prepared to place goods on the market at a price of less than *P*, as indicated by the supply curve for quantities of less than *q*. The area *cbe* represents the aggregate gain

to producers, in terms of income they receive above what they would have been prepared to accept, and is known as the *producer surplus*. The sum of consumer and producer surplus, known as economic surplus for the market, represents the total community benefit from the market.

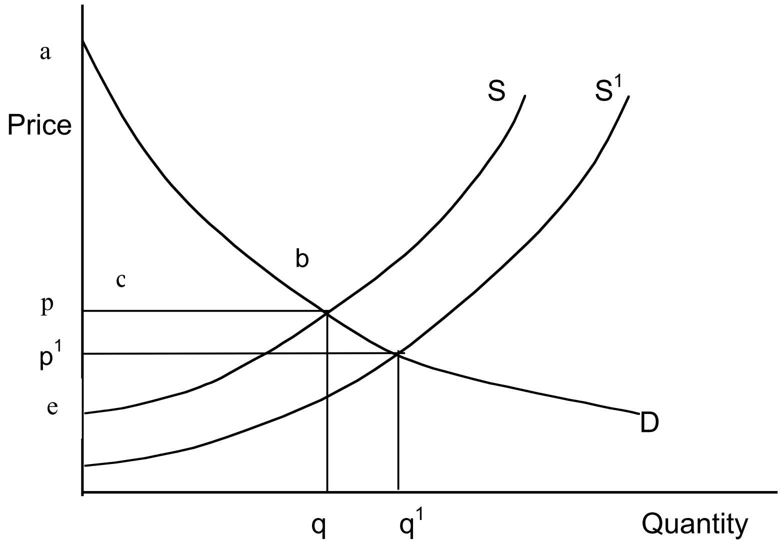


Figure 1. Market diagram illustrating producer and consumer surplus

Now suppose a project is carried out, which reduces the costs for producers. Suppose this shifts the supply curve (which according to economic theory is the short-run marginal cost curve) moves to the right, from *S* to S^1. As indicated in Figure 1, there is a new equilibrium price p^1 and quantity q^1. In this new situation, both the producer surplus and the consumer surplus increase. It is notable that not all the gains from lower production costs are appropriated by producers. Rather, the benefits are divided between producers and consumers, the relative shares depending on the proportionate changes in areas of the producer and consumer surplus, which in turn depend on the shapes (elasticities) of the supply and demand curves.

Although the economic surplus concept is confined to a single time period, it does shed light on how benefits from projects should be estimated. Producer surplus can be approximated by net revenue or profits to producers. Gains to consumers or 'consumer profit' can be estimated by what consumers are required to pay relative to what they would be willing (and able) to pay. *Willingness-to-pay* (WTP) is a critical indicator of value in economics

2. DISCOUNTED CASH FLOW ANALYSIS AND PROJECT PERFORMANCE CRITERIA

The technique known as discounted cash (DCF) analysis, which provides the computational mechanism for CBA, was described in Module 11. As indicated in that module, the economic performance criterion of net present value (NPV) is often used as a measure of project performance. A limitation of NPV is that it does not indicate a rate of return on investment, and in this context B/C ratios and the internal rate of return (if it exists and is unique) can be used to provide additional information. A variety of definitions of benefit-to-cost ratio may be devised, depending what items are included in the numerator (some measure of gross or net benefits) and the denominator (some measure of overhead and perhaps operating costs).

In this module, the emphasis is on the broader issues of defining the costs and benefits and other parameters relevant to CBA, so as to derive annual cash flows estimates for a project (from which DCF performance criteria can be calculated).

3. COMPLEXITIES IN COST-BENEFIT ANALYSIS

Cost-benefit analysis is used by a wide variety of people for a wide variety of tasks. The availability of powerful computer spreadsheets with build-in financial functions has greatly facilitated the application of CBA. At the same time, the ease with which the computational side of the technique can be applied has frequently led to a rather mechanistic approach being adopted. No two projects will have exactly the same characteristics, and it is unlikely that a standardised approach to CBA can ever be relied upon. Blind use of the technique can lead to results which do not truly represent the investment situation, and to poor information for decision-makers. The practitioner needs both to be aware of a number of complexities which often arise and the best ways of dealing with them. To some extent, successful application of CBA is an art rather than a science, and there is no substitute for experience in carrying out real-world applications. However, there are a number of complexities of which potential users can be made aware.

The concept of a project

The term ‘project’ will be used in the broadest sense throughout this module to mean any type of investment which affects costs and benefits over time, for an individual firm or industry or for the public sector. To be more precise, the terms ‘proposal’, ‘project’ and ‘program’ can be taken to mean different things. A proposal is a planned investment, prior to acceptance. A project is a specific type of investment, carried out by an individual or team usually from a single firm or government agency. A program consists of a number of related projects, such that various projects in a program may proceed in relative isolation from each other, to achieve related goals. For example, a forest industry development program may contain various individual tree-planting projects.

Public sector versus private sector orientation

Investment projects are carried out by both the private and the public sector. In the private sector, a landowner may wish to establish a forestry plantation. This will involve substantial initial outlays, with the main payoff taking place many years into the future. Similarly, a government (local, provincial or national) may wish to promote community forestry projects, providing funding for the establishment phase and perhaps taking an equity in the project to share revenue at a later time. Governments also support infrastructure projects, such as roads and schools, which indirectly enable revenue to be earned. They may also spend on hospitals, libraries and other community service obligations (CSOs), legitimately funded from taxation revenue. CBA is essentially a technique to be used for evaluation of public sector (government) projects. Discounted cash flow analysis techniques are also relevant to private sector investments, although a number of important differences arise in their application.

Project analysis from a private firm’s viewpoint is known as *financial* analysis, while social CBA is referred to as *economic* analysis. In this context, sometimes both the financial internal rate of return (FIRR) and the economic internal rate of return (EIRR) are estimated for the same project.

Individual agency versus multiple agency evaluation

Often projects are financed by more than one public agency. For example, community forestry projects may be supported by national and provincial or local government and by international loan agencies. A CBA may be carried out from an *individual agency* viewpoint (e.g. taking into account only costs and returns for the national government’s perspective) or from the viewpoint of *all investors* in a project.

Identifying the project and its variables and bounds

Experienced CBA practitioners sometimes comment that one of the most difficult tasks in carrying out any project evaluation is to determine the bounds to the project. Any change to a component of an economic system is likely to have an impact on various other components of that system. The further removed from the project these components are, the smaller the impact is

likely to be. As an example, consider the case of community rainforest reforestation. This is likely to have benefits in terms of creating a timber resource, and may help restore degraded farmland, improve wildlife habitat, protect streambanks and water quality, and provide employment in tree planting and maintenance. These may in turn have further impacts, e.g. reduced stream erosion may reduce loss of agricultural land, stream and water storage siltation and road maintenance, and improve fishing and enhance tourism. Some of these impacts are felt a considerable distance from the tree planting area, e.g. further down a watershed. Obviously, it is not possible to analyse the impacts of a local investment project on the 'whole world'. The task is to determine reasonable limits to the impacts. In essence, this means defining the technical scope of the project, and listing all of the variables which are likely to be affected, in terms of benefits received and costs imposed. As well, an attempt needs to be made to determine the nature of the relationships between relevant variables. At some point, it will be necessary to make the explicit assumption that some impacts which could be included are not being considered in the analysis.

Defining the 'with project' and 'without project' cases

When evaluating a project, it is necessary to define clearly the situations both with the project and in the absence of the project; the *incremental cash flows* arising from the project can then be identified. In this context, the present situation may not correspond to the 'without project' case. As an example, suppose an investment is to be made to establish a community-based forestry plantation. The 'with project' case will relate to the new plantation, while the 'without project' case will relate to the alternative use of the land. If the plantation were not established, the land may continue to be used as it is currently, or some other use may be intended. Similarly, other resources could be used differently, depending on whether the project went ahead. Thus the labour in site preparation and tree planting may be to some extent diverted from growing food crops, and hence implementing the project could result in reduced food crop production. When determining cash flows, it is not the costs and benefits of the plantation, but rather the differences in cost and benefits if the plantation is developed relative to the situation were it not developed, which should be taken into account in project evaluation.

Project designs, management options and environment scenarios

When setting up a CBA, it is necessary to define what alternative forms the project could take, to be compared in the analysis. A number of scenarios with respect to the physical, economic and legal and institutional environment may also need to be considered. This is all part of project definition.

Time-phased investment projects

Projects sometimes consist of several stages, some of which may not take place for a number of years into the future. The decision then arises as to whether to evaluate the overall project or only one or a small number of initial stages. A particular problem arises where infrastructure is 'oversized' in the first stage so that it is adequate for later states. For example, a particle board plant may be constructed in conjunction with a plantation forestry project, but may be designed to process more logs than are likely to be available in the near future, with a view to taking advantage of new plantings and greater particle board demand in the future. It is appropriate to attribute only that proportion of the overhead cost needed in the current stage of the project to the evaluation of this stage.

Relevant costs and benefits

When evaluating a project, it is necessary to include as far as possible all those costs and benefits to society associated with the project. It is a far from simple task to determine what cost and benefit variables should be included, and in fact, a major step in the analysis will be to identify which costs and which benefits are relevant. However, some general principles can be laid down. Project benefits include additional revenue generated, cost savings and positive social and environmental

externalities. Project costs include overhead and (discussed further below) operating costs. Operating costs include *explicit* and *implicit costs*, and negative social and environmental externalities. Explicit costs are those out-of-pocket expenses such as payments for labour (i.e. wages), fertilizer and seedlings. In contrast, implicit costs arise from use of a firm's own resources, such as labour, financial capital (money in the bank) and buildings, for which no actual payments are made. *Environmental and social externalities* of projects arise with respect to people other than those carrying out or having a financial interest in the project. For example, if plantation establishment leads to carbon sequestration or reduced sedimentation of water storages (a positive environmental externality) or reduces the number of lives lost in coastal towns from floods (a positive social externality), then these should be taken into account in the economic analysis. The valuation of environmental impacts presents many challenges. A low-cost approach which is intuitively appealing is to determine standard values for environmental conservation, which can be applied in rapid project appraisal. For example, a value could be placed on each hectare of native forest conserved, or each individual of a threatened species protected. A number of environmental databases have been developed to assist in benefit transfer methods. More advanced valuation approaches are discussed in another module.

Treatment of capital items in cash flows

Capital items may be divided into project capital outlays, working capital and salvage values at the end of the project's lie.

Capital outlays. Initial outlays and not depreciation allowances should be included in the cash flows. That is, the 'cash' flows of financial transactions should be recorded at the time they are made, rather than making periodic allowances for the services yielded by capital items.

Working capital. Often it is appropriate to make an allowance for working capital which is tied up during a project. This allowance (e.g. 2% of the capital outlays) can be treated as an outlay at the beginning of the project, with the full amount treated as an inflow at the end of the project's life.

Salvage value. Where plant and equipment are purchased for a project, an allowance for items on hand may be made as a capital inflow for the final year of project life, e.g. a scrap value of 10% of initial outlays. The value of standing trees at the end of the planning period could also be treated as a capital inflow.

Some special cases of project costs and benefits

Several cash flow items arise which warrant furrther comment.

Transfer payments. Payments from one group of stakeholders to another – such as taxes and subsidies – do not constitute any net gain or loss to society and hence should not be included in project cash flows in social cost-benefit analysis. (These amounts would however be relevant when evaluating a project from a private sector perspective.)

Interest payments. As mentioned earlier, the discounting procedure simulates interest payments. In any case, interest payments are transfers between borrowers and lenders, and so should not be included in social cost-benefit analysis.

Unpriced or non-market externalities spillover effects. It is desirable to include, as far as possible, social and environmental externalities in social CBA. (Externalities would not normally be considered when evaluating a project from a private sector perspective.) Typically, most of the social and environmental impacts of a project are not reflected in market transactions, and hence placing values on them becomes difficult. Some relevant valuation techniques were introduced in Modules 13 and 14. The term 'cash flow' seems somewhat at odds with the concept of unpriced values, but is used out of tradition; the term 'economic flows' may be more appropriate.

Standard lists of checklists of costs and benefits

It would seem desirable to develop lists of cost and benefit categories which can be

referred to when carrying out a CBA. However, projects dealing with different natural resources usually have different categories of costs and benefits, and even projects dealing with the same resource typically have a number of unique features. Hence checklists are a useful starting point, but should not be relied upon too heavily.

Determining the planning period

The number of years for which cash flows are estimated is referred to as the project's planning period or *planning horizon*. This depends on the planning horizon of the decision-maker, and the realistic life of the project. In the case of a forestry plantation, the planning horizon would normally be at least at least one rotation to final clearfell, but could be more than one forestry rotation. If there is no clear end point to the project – such as in the case of a limited production forest – then a period of the order of 20 to 30 years would normally be adopted. After about 30 years, the cash flows have little impact on DCF performance criteria. The Australian Department of Foreign Affairs and Trade, which commissions financial and economic evaluations of irrigation, water supply and other overseas projects for which Australian grant aid is provided, suggests 'a minimum of a 20 year planning horizon is generally considered appropriate' (AIDAB 1993). It is suggested here that, as a starting point, a period of 20 years be adopted, and that checks be carried out to determine whether this can be reduced or needs to be extended. It is to be noted that controversial issues can arise with respect to sustainable resource use and intergenerational equity relative to planning horizon and the discount rate.

In the case of research projects, it is necessary to predict how long the benefits will continue before they become replaced by new technology created by subsequent R&D. Also, it is necessary to make a judgment as to whether subsequent research will enhance or replace those from the current project; in the former case, the benefits could be assumed to continue while in the latter case they would terminate.

Choice of discount rate

The discount rate adopted in CBA has a major impact on estimated values of the performance criteria.[1] There is considerable debate in the CBA literature over the appropriate discount rate concept to adopt. The Department of Finance (1991, Ch. 5) noted the concepts of *social time preference rate* (STPR) corresponding to society's preference for present as against future consumption, and *social opportunity cost of capital* (SOC) corresponding to the rate of return on investment elsewhere in the economy. They recommended adoption of the project-specific cost of capital. This involves using the cost of borrowing, which is in most cases the long-term bond rate. A rate of 8% was recommended, comprising a risk-free rate of 5% and a risk margin of 3%.

Real and nominal interest rates

The interest or discount rate for social CBA is usually taken as the real rate, net of the rate of inflation. Consistent with this, the annual cash flow variables are measured in *constant* dollars, that is, they are not increased over time to take account of inflation. This of course does not mean that benefits and costs cannot increase over time due to other factors. An obvious difficulty arises from working with constant prices, in that in practice the rate of inflation may be greater for some project variables than others. If timber becomes scarce, the price increase may relatively rapid, such that a *real* price increase (i.e. an increase greater than the overall inflation rate) takes place. This could be incorporated in the analysis by either working in *current* prices throughout, or by allowing annual increases in timber prices equal to the difference in the rate of timber price increase less the inflation rate (i.e. the real price increase).

To determine the real rate of interest for CBA, the *nominal* rate must be adjusted to remove the inflation component. The way this is done depends on the relationship assumed between the real rate and the rate of inflation. In this context, two different

[1] As an illustration of this, a sum of $1M in 20 years time has a present value of $377,000 if the discount rate is 5%, but only $177,000 if the discount rate is 10% and only $26,000 if the discount rate is 20%.

models are possible. First, these may be assumed to be *additive*:

$$1 + n = 1 + i + f$$

where n is the nominal rate of interest (e.g. market or long-term bond rate), f is the inflation rate, and i is the real rate of interest. Here all rates are expressed on an annual basis, and in decimal form. For example, if the nominal interest rate is 11% and the inflation rate is 3% then

$$1 + 0.11 = 1 + i + 0.03$$

from which

$$i = 1.11 - 1.03 = 0.08 \text{ or } 8\%$$

More often, a *multiplicative* model is adopted, of the form

$$1 + n = (1 + i)(1 + f)$$

The solution to this interest rate formula, in terms of the real rate, is

$$i = \frac{1 + n}{1 + f} - 1$$

For example, if the nominal rate is 11% and the inflation rate is 4% then the real rate is

$$i = \frac{1 + 0.11}{1 + 0.04} - 1 = 0.0777 \text{ or } 7.77\%$$

It is to be noted that the real rate is slightly less under the multiplicative model than under the additive relationship. Use of the former is recommended in the recommendations for cost-benefit analysis laid down by the Department of Finance (1991).

For discounting purposes, a constant rate of interest is normally applied throughout the planning horizon. While the rate will no doubt vary over time, the current rate is usually adopted. However, if interest rates are expected to change in a particular direction in the short term, then this may be taken into account in the choice of rate. It is to be noted that because cash flows further removed in the future have less impact on present values than cash flows in the short term, predicting an appropriate rate over the first five to 10 years is more important than longer-term predictions.

Dealing with uncertainty in CBA

Typically, the annual cash flows of a long-term project are subject to a high level of uncertainty, particularly the future project benefits, and various procedures are available for taking this into account in CBA. Uncertainty can arise because of physical and financial factors, and also legal and institutional change. For some projects, changes in exchange rates with overseas countries will contribute to financial risk, and costs will be incurred with hedging or insurance against exchange rate depreciation.

Typically, a component for risk is included in the discount rate, on the grounds that a higher rate of return is needed to justify the greater project risk. Sometimes conservative estimates of project benefits are used; a systematic way of doing this is to compute *certainty equivalents* (the risk-free amounts which the decision-maker regards as equally acceptable to the higher but uncertain amounts). Normally, *sensitivity analysis* is carried out to determine what effect changes in parameter estimates would have on performance criterion such as NPV. An alternative which provides greater information is to apply *risk analysis*, in which probability distributions are attached to key variables or parameters, and the probability distribution of project performance (say NPV) is derived.

Allowing for slippage in economic performance of research projects

When applying CBA to research projects, it is usual to make allowance for the probability of research success (in terms of achieving the planned technical outputs and developing them into a usable technology package) and the level and timing of uptake of this technology by potential adopters, e.g. see Harrison *et al*. (1991). Notably, the cost of the development stage, and technology transfer (extension) also need to be included in project outlays.

Shadow pricing

In general, the prices of inputs and outputs used in CBA are those prevailing in markets at the time of the analysis. However, sometimes these may be regarded as unsuitable for the analysis, and may be replaced by *shadow prices*. 'A shadow price is an imputed valuation placed on a project input or output when a market price does not exist or is significantly distorted' (AIDAB 1993). Since natural resources are often underpriced (EPAC 1991; Harrison and Tisdell 1992), there is sometimes scope for shadow pricing in projects related to natural resources. When evaluating projects, a large premium (20% or more) is sometimes placed on wages earned (where there is high unemployment) or on export revenue (where there is a shortage of foreign exchange). These loadings may be justified on the basis that governments spend large amounts of money on job creation and export promotion activities.

Multiplier effects

It is not unusual for proponents of particular projects to argue for them on the basis of indirect of flow-on benefits, e.g. creation of local employment, more local spending hence improved local business activity, and improvements to local real-estate prices. Investment in a plywood or pulp mill could create a number of jobs and provide a substantial stimulus to a regional economy. These are the kinds of benefits which can be examined with inter-industry input-output analysis, in which employment, income and output multipliers are derived. While these benefits are real enough from a local viewpoint, the case for them is not so strong from an overall social perspective. The question needs to be asked: 'Would investment in one particular area create greater flow-on benefits than investment elsewhere?'. If not, then there may not be sufficient justification for estimating flow-on benefits for the particular area. As indicated by the Industry Commission in its study of water management (IC 1992), the 'wider beneficiaries' argument is not a strong one, and the main use of multiplier effects is to examine the adjustment problems when resource uses are curtailed in an area.

Effect on income distribution

Public sector projects will often have an effect on income distribution in the community. Some people are likely to be better off and others worse off. Intuitively, it would be appealing if a project improved the lot of the poor, as is likely with a community forestry project. A greater weight could be placed on income generated by a project which helped the poor than one which benefited mainly the rich. However, this would lead to difficult estimation problems, including interpersonal comparisons of utility. The economist has a role in identifying the likely impacts on income distribution. However, judgments about desirability of changes to income distributions 'are almost always most appropriately made by Government at the political level' (Department of Finance 1991).

Project evaluation versus project justification

This module would not be complete without sounding a warning about misuses of CBA. Given the complexities discussed above, it is easy inadvertently to use the technique incorrectly. More seriously, given that CBA is often used to demonstrate that a particular investment is worthwhile and to assist in obtaining funding, there can be a strong temptation to seek optimistic benefit estimates and hence ensure a positive NPV. That is, CBA can be used to generate misinformation rather than good information. Where agency goals to obtain approval for new projects are strong, considerable suasion may be placed on economists to come up with 'good figures'. How this is handled may become a difficult question of personal integrity versus loyalty to and peer acceptance in the organization. If the economist disagrees with the figures that are being proposed, then he or she has a responsibility to communicate clearly the basis of this disagreement.

4. INTEGRATION OF NON-MARKET ENVIRONMENTAL VALUES IN SOCIAL CBA

In general, if non-market values can be estimated, these can be factored into a CBA, alongside market values. They

become part of the annual incremental cash flows (or 'economic flows'), to be placed alongside financial flows, and traded off on a dollar for dollar basis. Thus, degradation of the environment and depletion of natural assets are included as part of the cost stream for a project, or protection of these values as part of the benefit stream. This is consistent with the economic concept of value as community willingness-to-pay for both market and environmental values.

Acceptability of non-market values and 'funny money' arguments

The inclusion of, and failure to include, non-market values in CBA are both often the subject of controversy. This raises the question of whether non-market values should be given lower weights than market values in CBA. In general, there does not seem justification for this discriminatory analysis, although there can be some cases where it might be acceptable.

Consumer versus producer surplus. Often a major component of environmental values are consumer surplus values, e.g. recreation values, landscape amenity, air quality. On the other hand, often the market values are producer values, e.g. from timber production. In that governments subsidize industry, it would seem that the shadow price of producer surplus is sometimes accepted as being higher than that of consumer surplus.

Negative signals to investors. The reservation of areas containing valuable exploitable natural resources, for conservation and recreation purposes, can provide negative signals to industry. It would appear that this has in some cases led companies to invest overseas rather than in the domestic economy, with loss of domestic production and employment. Again, this suggests there may be an argument for placing a higher weight on producer surplus than on consumer surplus.

Exports versus domestic products. Sometimes greater weight is placed on exports than domestic products and services. This is relected in tax concessions and other incentives to exporters, and is sometimes motivated by trade deficits and servicing of foreign debt. This suggests placing a higher weight or shadow price on exported commodities as against domestic goods and services.

Uncertainty and bias in value estimates. Precision of estimation of non-market values can be low. Generally, market values can be estimated more precisely, although this certainly is not always the case. The problem of uncertainty is exacerbated by a number of potential sources of bias which are recognized for some non-market valuation techniques, e.g. CVM. If there is a suspicion that non-market value estimates are upwardly biased, this could be a reason for downward adjustment.

It is to be noted that these are possible reasons for placing a lower weight on consumer surplus than producer surplus, but not for ignoring the latter (leaving it out of the analysis). Also, changes in levels of environmental services have can affect prices of market goods, e.g. higher agricultural costs, higher expenditures on environmental protection, higher costs of health services.

5. CONCLUDING COMMENTS

Cost-benefit analysis is both a conceptual framework which is used widely in economics, and a technique for project appraisal. As a framework, it provides guidelines on what cost, revenue and non-market variables to take into consideration in project appraisal and how to define and value these variables. As a technique, it has wide application in the evaluation of proposed public sector projects. CBA is relevant to a variety of economic issues in forestry, such as evaluation of government assistance programs to landholders, communities and industry, and evaluation of forestry research projects.

REFERENCES

AIDAB (Australian International Development Assistance Bureau) (1993), *Economic and Financial Evaluation of DIFF Projects*, Canberra.

Department of Finance (1991), *Handbook of Cost-Benefit Analysis*, AGPS, Canberra.

EPAC (Economic Planning Advisory Council) (1991), *Issues in the Pricing and Management of Natural Resources*, Background Paper No. 16, AGPS, Canberra.

Harrison, S.R., Dent, J.B. and Thornton, P.K. (1991), 'Assessment of Benefits and Costs of Risk-Related Research in Crop Production', in R.C. Muchow and J.A. Bellamy, eds., *Climatic Risk in Crop Production: Models and Management for the Semiarid Tropics and Subtropics*, C.A.B. International, Wallingford, pp. 491-510.

Harrison, S.R. and Tisdell, J.G. (1992), 'Reform of Natural Resource Pricing', *Micro-Economic Policy and Reform for International Competitiveness*, Economic Society, Gold Coast, pp. 241-253.

IC (Industry Commission) (1992), *Water Resources and Waste Water Disposal*, Report No. 26, AGPS, Canberra.

16. Using Financial Analysis Techniques in Forestry Research

John Herbohn

Previous modules introduced the principles and methods of Discounted Cash Flow (DCF) and Cost Benefit Analysis (CBA) in relation to forestry projects. The current module extends the discussion of DCF to include its use in financial analyses. The purpose of CBA is to determine whether investment projects (used in the widest sense) are worthwhile from a social perspective, taking into account all costs and benefits including positive and negative externalities. In contrast, the focal point of financial analysis is assessing the impact of the project of the wealth of the individual or firm considering the project. In this module, concepts and methods of financial analysis (FA) will be introduced, and then applied to a case study of financial evaluation of a forestry project.

1. INTRODUCTION

Cost-benefit analysis (CBA) and financial analysis (FA) have many elements in common, the key distinction being that the former is carried out from the perspective of the society or community while the latter is carried out from the perspective of the private interests of the investor. CBA is typically used in making decisions about public sector projects. Financial analysis typically is used as a technique in the capital budgeting decisions of firms and individuals. CBA would include all the impacts (both financial and non-financial) of a project to the organization or government undertaking the project and to society as a whole. Financial analyses is more narrowly concerned about the financial or direct cash flow impact of a project to an individual or business, with environmental externalilties such as pollution only included in a financial analyses to the extent that they impose cash flows such as fines or clean-up costs on the firm or individual. Similarly, social benefits such as increased employment or standard of living by local communities from the establishment of a timber mill (relevant in CBA) would not be included in a financial analysis because there are no direct cash flow consequences to the firm.

In many ways financial analysis can be considered as a sub-set of CBA. The direct cashflow consequences of a project to the organization undertaking the project (as identified in a financial analysis) must be identified as part of a CBA, although the distinction is seldom explicitly drawn between cashflows directly impacting on the organization and those impacting on other parties affected by the project.

Capital budgeting is a multi-faceted activity. There are several sequential stages in the process. For typical investment proposals of a large corporation, the distinctive stages in the capital budgeting process are depicted, in the form of a highly simplified flow chart, in Figure 1. The evaluation is carried out within the strategic planning of the company, and the screening of investment alternatives. Information from the financial analysis is integrated with other information and objectives of the company to arrive at a decision to accept or reject each project. Accepted projects must undergo implementation and monitoring procedures.

2. SOME KEY CONCEPTS IN FINANCIAL ANALYSIS

Incremental cash flows

The concept of an *incremental cash flow* was introduced in the previous module. Incremental cash flows are the cash *inflows* and *outflows* which will arise from the implementation of a particular project, and are estimated by comparing the cash inflows and outflows of the firm '*with*' the project, and those '*without*' the project. It is a 'marginal' or 'incremental' analysis comparing two situations.

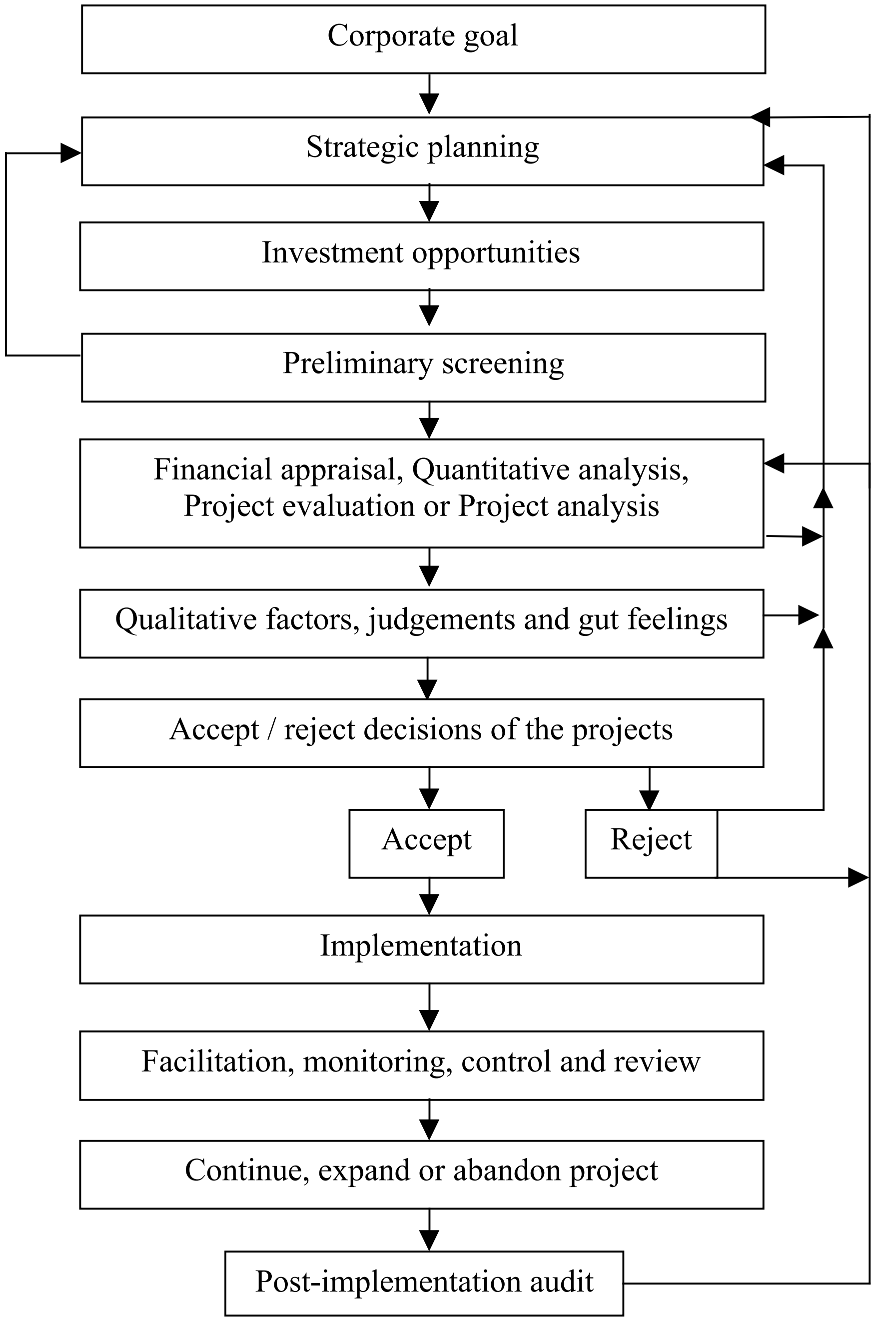


Figure 1. The capital budgeting process

For analytical purposes project cash flows may be separated into two categories: capital cash flows and operating cash flows. Capital cash flows may be disaggregated into three groups: (1) the initial investment, (2) additional 'middle-way' investments such as upgrades and increases in working capital investments, and (3) terminal flows. These are all cash flows and the distinction between them are only to facilitate the convenient identification of the different categories.

Synergistic effects and opportunity cost

All the indirect or synergistic effects of a project should be included in the cash flow calculation. Synergistic effects can be negative or positive. For example, suppose a land owner has 100 ha of plantations which are at a stage of requiring thinning. There is no current market for the thinnings and he will have to pay local people to thin his plantation. He is also considering buying a plant to make charcoal. If he buys the plant he could then use the thinnings from his plantations to make charcoal. In this case there is a positive 'synergistic effect' of investing in the charcoal plant i.e. the investment creates a use for a previous waste product.

The rationale for the incorporation of these indirect effects has its base in the 'opportunity cost' principle. When a firm undertakes a project, various resources will be used up and not available for other projects. The cost to the firm of not being able to use these resources for other projects is referred to as an 'opportunity cost'. The value of these resources should be measured in terms of their opportunity cost. The opportunity cost, in the context of capital budgeting, is the value of the most valuable alternative that is given up if the proposed investment project is undertaken. This opportunity cost should be included in the project's cash flows.

To consider another example, suppose a landholder is considering establishing 5 ha of Gmelina. The plantation will be established on land that he does not currently use but which he could rent out for 20,000 pesos per year. In this case the farmer will lose the opportunity to rent the land at 20,000 pesos per year if he establishes the plantation. This 'opportunity cost' should be included in the cash outflows of the project even though there has not been any actual cash payment. The reason is that opportunity cost of the space measures an extra cash flow that would be generated (for the firm) 'without' the project. On the other hand,suppose that the land has not been rented in the past and there is no intention to rent, sell or use for any other purpose in the future. In this case, there is no opportunity cost if the resource is used for the proposed project. Therefore, in this situation, the 20,000 pesos will not be included as a cash outflow.

3. TREATMENT OF KEY DCF VARIABLES

The basic principles of DCF analysis have been discussed previously. It is however useful to consider some of the key variables often found in DCF analysis in a little more detail.

Treatment of taxation payments

Tax is a cash payment to a government authority. How taxation payments are viewed differs depending on whether a CBA or FA is being undertaken. This is due to the different perspectives of each of these types of analysis. In CBA, in the perspective is that of the community, taxation payments are viewed simply as 'transfer payments' between one section of the community and another. They do not represent any net gain or loss to society, and hence are excluded from incremental cash flows.

With a financial analysis the focus is on the cashflow impacts on the individual or firm, hence tax payments need to be included in the analysis. If the project generates tax liabilities, then the tax payable is relevant to the project, and must be accounted for as a cash outflow. Corporate tax is a cash outflow. If the tax were levied on net cash inflow and paid at the same time cash was received, then the after-tax net cash flow would be easily calculated. However, tax is not based on net cash flow, but on taxable income.

Taxable income is defined by the relevant tax act and does not necessarily mean the same thing as net cash flow or even

accounting income or accounting profit. Taxable income is generally calculated by subtracting allowable deductions from assessable income. These terms are specific to particular tax acts, and are not easily dealt with in a general context. However, project evaluation needs to be able to accord some treatment to this calculation to determine after tax cash flows.

Treatment of depreciation

The tax definition of 'deductions' treats some non-cash items as allowable expenses. One such item frequently encountered in project analysis is asset depreciation. Depreciation is not a cash flow. It is an allocation of the initial cost of an asset over a number of accounting periods. Asset costs are allocated within accrual accounting systems so that they are matched over time against the income generated by the assets. That is, the initial cost of an asset is expected to benefit the firm over several years, hence the total initial cost is 'spread' over those future benefit years.

The actual per annum dollar amount of depreciation is only a notional amount. It does not represent the annual decline in value of the asset, it does not measure the value of the asset used up, and it does not measure the actual unit costs of the asset's services. A number of accounting methods can be used to calculate depreciation i.e. straight-line, reducing balance, 'sum of years digits' and units of production methods.

In project evaluation, what is relevant is not the accounting depreciation but the tax-allowable depreciation. The methods of calculating tax-allowable depreciation are often prescribed by the taxation legislation within a country. Sometimes the firm will have a choice between alternative prescribed methods, and in these cases the firm usually selects the method which will reduce the overall tax bill. The tax bill will be reduced if higher depreciation is claimed in the earlier years, thus delaying the payment of tax. The reducing balance method has this effect. Many national tax acts permit *accelerated depreciation* of equipment by allowing depreciation methods (defined in the tax act) which allow higher tax deductions in early years and lower deductions later. The Modified Accelerated Cost Recovery System (MACRS) in the USA. is an example.

It is important to understand that the interest in depreciation in financial analysis lies only in its tax effect, i.e. the depreciation tax shield or the reduction in taxes attributable to the depreciation allowance. If depreciation were not a tax deduction, it would not be considered in an NPV evaluation.

Treatment of financing flows

Treatment of financing flows is an area of confusion and sometimes a source of error in project analysis. It is important to distinguish between *project cash flows* and *financing cash flows*. For the purpose of identifying the relevant cash flows for project evaluation, the investment decision (project) must be separated from its financing decision. The financing investment decision involves deciding upon what proportion of the cash flows needed to fund the project are provided by debt holders and what proportion are provided by equity holders (i.e. from cash already held by the firm). The decision about the particular mixture of debt and equity used in financing the project is a management decision concerning the trade-off between financial risk and the cost of capital. The investment evaluation decision determines whether the project's discounted cash flows exceed the initial capital outlay (investment), and so adds (net present) value to the firm.

Generally, interest charges or any other financing costs such as dividends or loan repayments are not deducted in arriving at cash flows, because the interest is in the cash flow generated by the assets of the project. Interest is return to providers of debt capital. It is an expense against the income generated to equity holders, and as such is deducted in the determination of accounting profit. However, it is not included in project cash flow analysis, because the discount rate employed in the NPV calculation accommodates the required returns to both equity and debt providers. Therefore, inclusion of interest

charges in cash flow calculations essentially would result in a double counting of the interest cost.

Interest is tax deductible, and therefore provides a tax shield for any investment. This benefit is also accounted for in the discount rate, as the rate employed in project analysis is an 'after-tax' rate. Accordingly, tax savings on interest expenses are not included in project cash flow analysis.

Almost always there are exceptions to general rules or practices and this is the case for the general rule for the treatment of financing flows. There are situations where interest charges are explicitly incorporated into the cash flows. The question here is not whether it is right or wrong to incorporate financing flows into the cash flow analysis, but whether that incorporation is done correctly or incorrectly. If it is necessary to show the financing flows explicitly in the cash flows, as is often preferred by non-financial mangers and chief executives, it is quite possible to include them explicitly in a correct and consistent manner without distorting the final results.

In property investment analysis, 'property' cash flows are distinguished from 'equity' cash flows. Property cash flows are the equivalent of 'project' cash flows discussed in this module. Property cash flow calculations do not include loan repayments and interest charges as deductions. This approach is consistent with the general cash flow definition in this module. 'Equity' cash flow calculations deduct loan repayments and interest expenses.

One of the objectives of property investment analysis is to evaluate the return to the investor under various debt and tax situations. A mortgage applies specifically to one particular property, rather than being an unidentified part of the capital structure of a corporation. Some investment in property is to gain tax advantages from interest deductions associated with debt financing. For these reasons, in property investment analysis, equity cash flows are calculated after deducting loan repayments (principal plus interest) from other cash inflows to enable the effects of borrowing and taxation to be evaluated.

Inflation and consistent treatment of cash flows and discount rates

Inflation will have an effect on the expected cash flows of the project. Both cash inflows and outflows could be affected by inflation. Market rates, such as interest rates and equity returns, in general will also rise when the expected inflation rate is high. As the market rate rises the required rate of return by investors will also rise. To deal with inflation appropriately, the project analysis must recognize expected inflation in the forecast of future cash flows and use a discount rate that reflects investors' expectations of future inflation.

If all cash flows as well as the discount rate are equally affected by expected inflation, the net present value is the same whether the inflation is include or excluded. However, most projects will consist of a multitude of cash flow items over a number of years and it would be erroneous to assume that all of the cash flows would increase at exactly the same rate each year, or to assume the same effect on the discount rate. Some cash flows are unaffected by inflation while other cash flows are affected to varying degrees by inflation.

An outstanding example of the differential impact of inflation on a project's cash flow is the depreciation tax shield. Tax-allowable depreciation is totally unaffected by inflation. Depreciation tax shields are calculated on the basis of historical costs of the assets (cost at the time of their acquisition). Similarly, a long-term raw-material contract or the purchase of a commodity in the *forward* or *futures* markets may lock in the present prices thereby insulating the cash flow from inflationary effects.

Given the differential impact of inflation on different cash flow components, cash flow forecasts in *nominal* terms – incorporating the inflationary effect – have an advantage over cash flow forecasts in *real* terms – excluding the inflationary effects. That is, nominal cash flow forecasts can incorporate potentially different inflationary trends in product price, labour and material costs, and so on, into cash flow estimates by

applying different inflation rates for different components of the cash flow.

The required rate of return used for discounting cash flows is normally derived from observed market rates such as interest rates and return on equity. These observed market rates usually have the expected annual inflation rates built in and are usually quoted in *nominal* terms (as opposed to *real* terms).

Observed market rates expressed in nominal terms can, if necessary, be converted into their real terms using the algebraic relationship expressed in the Fisher effect. The Fisher equation is:

$$(1+n) = (1+r)(1+p) \qquad (1)$$

where, n = the annual nominal interest rate (expressed as a decimal value)
r = the annual real interest rate (expressed as a decimal value)
p = the expected annual inflation rate

From the above equation the real interest rate can be easily derived as

$$r = \frac{(1+n)}{(1+p)} - 1 \qquad (2)$$

Consistency in the discounted cash flow analysis requires that if the project's cash flows are in nominal terms then they should be discounted by nominal discount rates, and if the project's cash flows are in real terms they must be discounted by real discount rates. Real and nominal quantities cannot be mixed and matched.

The interest rate used for discounting cash flows is normally derived from observed market rates. In an efficient financial market, investors' required rate of return will include a component, (1+p), to compensate for expected inflation. The use of observed market-required rates then implies that inflation should incorporated into cash flows, to be consistent.

On rare occasions, real cash flows are appropriate in the analysis. In such situations, the real discount rate can be calculated from market (nominal) rates using the Fisher equation presented above.

4. THE DISCOUNT RATE

The discount rate used in an analysis greatly affects the NPV estimate. The choice of an appropriate discount rate thus becomes critical in the appraisal of any project. Differences in the focus of CBA and FA once again tend to mean different approaches are used in determining the appropriate discount rate. The focus on the individual or firm in financial analysis means that the discount rate is firm specific. In essence, it is the risk-adjusted rate that the firm or individual considers appropriate for the project being evaluated. Often this equates to the 'cost of capital' for the project. If a project were to be financed entirely by a loan, the cost of capital would be the rate at which the firm could borrow the money for the project. If this rate was say 9% per year, then this is the cost of capital and the appropriate discount rate would be 9%. This is a rather simplistic example and in real life things are much more complicated. As in CBA, a sensitivity analysis involving the application of a number of discount rates is usually undertaken, e.g. using the preferred rate plus or minus two percentage points.

What are the components of a discount rate?

In most standard financial analyses a risk adjusted discount rate (RADR) is used. Conceptually, the RADR has three components:

- *Risk free rate* to account for the time value of money (r).
- An *average risk premium* to compensate investors for the fact that the company's assets (or investments) are risky. This is a risk premium to account for the business risk of the firm's existing activities, being simply *the average risk premium for the firm*. This is denoted as 'u'.
- An *additional risk factor* (which could be zero, negative, or positive) to account for the difference in the risk between the firm's existing business

and the proposed project, which is denoted as 'a'.

Thus, conceptually, the RADR (which will be referred to as k) may be expressed in an algebraic form as:

$$k = r + u + a. \quad (3)$$

If the risk of the proposed project is same as the average risk of the firm's existing projects, the additional risk factor 'a' is zero. The required rate of return (or the RADR) to be employed as the discount rate for *projects of average risk*, is then 'r+u'.

Estimating the RADR

In estimating the RADR to be used in project evaluation, in practice, the three different components may not always be calculated separately and added. For example, if cost of capital is used to estimate the RADR, then a value of '$r + u$' is obtained, to which can be added a value to account for 'a' (which could be negative). On the other hand, if capital asset pricing model (CAPM) is used to estimate the RADR, then separate values for *r, u*, and *a* will be obtained.

The risk-free discount rate, 'r' may be arrived at by examining government bond yields or insured banks' term deposit rates. By examining data published in the financial press, a suitable figure for 'r' can easily be arrived at. When doing so, it is important to that use a bond, guaranteed loan or other financial asset with a similar term to that of the project. For example, if a project stretched over 5 years, then the rate on a 5-year government bond or a 5-year insured fixed deposit would be adopted.

The Average Risk Premium for the Firm, 'u' may be estimated using the firm's cost of capital. The weighted average cost of capital (WACC) is normally adopted for calculating this. Another approach to estimate 'u' is to use capital asset pricing model (CAPM). Illustration and application of these involve relatively lengthy calculations.

The Additional Risk factor 'a' is the difference between the average risk faced by the firm and the perceived risk of the proposed project.

5. ANALYSING PROJECT RISK: SENSITIVITY AND BREAKEVEN ANALYSIS

There are numerous ways to analyse projects for risk. Two of the most widely used are sensitivity analysis and breakeven analysis.

Sensitivity analysis

One of these is to evaluate the project under various scenarios in which selected parameters or variables are stepped through their pessimistic, most likely and optimistic values. In this analysis, often only one parameters at a time is changed. The resulting set of NPVs reveal to management which variables have a material impact on financial performance of the project. Management can then decide to either invest more time and effort in establishing more reliable forecasts for these parameters, or to abandon the project because of excessive risk.

The steps in sensitivity analysis are:

1. Calculate the project's net present value using the most likely estimate for each parameters.
2. Select from the set of parameters those which management feels may have an important bearing on the project.
3. Forecast pessimistic, most likely and optimistic values for each of these parameters over the life of the project.
4. Recalculate the project's net present value for each of the three levels of each parameters. Whilst each particular parameters is stepped through each of its three values, all other parameters are held at their most likely values.
5. Calculate the change in net present value for the pessimistic to the optimistic range of each parameters.
6. Identify parameters for which project financial performance is sensitive.

Break-even analysis

Break-even analysis is a special application of sensitivity analysis, which endeavours to find the levels of individual parameters at which the project's NPV is zero. For example, management may wish to know how low can the unit selling price can fall before the project becomes unsuccessful. If managers know that this 'cut-off' price is likely to eventuate, then they may decide not to proceed with the project.

In common with sensitivity analysis, parameters selected for break-even analysis are typically tested only one at a time. Within an Excel spreadsheet, the calculations can be carried out in three ways:

1. by creating a data table for a range of values and reading off the appropriate value at a zero NPV.
2. by trial-and-error substitution of variable values within a spreadsheet
3. by using the Excel 'Tools – Goal Seek' function.

Management can use the break-even results in two ways. They may decide to abandon the project prior to implementation if forecasts show that near break-even parameter levels can be expected to occur. Once a project is implemented, management can react to a worst-case scenario involving the investigated parameters, e.g. suspend production, try to make production more efficient, or adjust the unit selling price.

6. DETAILED EXAMPLE OF A FINANCIAL ANALYSIS OF A FORESTRY PROJECT

A case study of a forestry investment is now presented to illustrate the steps of financial analysis. The example concerns a company (lets call it Flores Venture Capital Ltd, or FVC Ltd) which is considering diversifying its operations into forestry. The company has obtained a consultant's report on the proposal, which involves the establishment of a plantation estate of 1000 ha in an area with suitably high rainfall and soil quality. To encourage investment in the region, a local government has offered the required land rent-free for the period of the project on the condition that a native species rather than an exotic pine is grown. The company has decided to establish a mixed species (eucalypt and rainforest species) plantation. The finance department of FVC Ltd has indicated that a rate of return of 7% is required for the project.

Step 1: Identifying the forestry system

In Module 3, it was demonstrated how the Delphi method could be used to develop estimates of key parameters. The example used was for a forestry project involving the use of two species for which little quantitative growth and harvest age data existed. In this module, it is demonstrated how a forestry financial evaluation can be undertaken using similar data. The example in Module 3 was based on a real-life Delphi survey and the associated financial analysis was highly complicated; the basic information from that example will be used here, but will be simplified in a number of ways.

Because there was no past experience in growing native species in the region, the consultant used the Delphi method to develop an appropriate silviculture (tree growing) system. Based on the results of this investigation, it is recommended that trees be planted at a density of 660 stems per hectare. It is expected that extensive weed control will need to be undertaken in the first year, with further weed control required in the second and third year. Pruning of the trees to ensure good form will be required in years 2, 4 and 6. A number of experts involved in the Delphi survey indicated that the amount of pruning required is difficult to estimate because no one knows how much branching will occur on the sites and this could range from minimal to requiring high labour inputs. The consultant has explained that pruning is crucial because it ensures that the trees produce a straight, knot-free log for which high prices can be obtained. It was also recommended that each pruning event should be certificated by an external party because this will increase the likelihood of being able to obtain a premium price for knot-free wood. A non-commercial thin is required at year 8, at which time 320 trees will be removed.

The first revenue from the plantation is predicted to come when a commercial thinning occurs at the end of year 17, at which time about 85 trees will be harvested. At year 26, 85 further trees will be cut and sold for telephone and electricity poles. The best 85 trees will be left to grow until year 34, when about half (42) will be cut for sawlogs. The remaining trees will be allowed to grow to year 60 when they will be harvested and sold as high quality veneer logs.

From the above information about the plantation system, the major cash outflow and inflow categories have been identified as listed in Table 1.

Table 1. Main cash categories and predicted timing

Cashflow category	Nature of cashflow	Timing (year)
Establishment (capital) costs	Planning and design	0
	Incidental clearing	0
	Site preparation and cultivation	0
	Cover crop establishment	0
	Pre-plant weed control	0
	Cost of plants	0
	Planting and refilling	0
	Post plant weed control	0
	Fertilizer	0
	Fencing	0
Maintenance costs	Post plant weed control (1)	1
	Post plant weed control (2)	2
	Post plant weed control (3)	3
	First prune(plus certification)	2
	Second prune (plus certification)	4
	Third prune (plus certification)	6
	Thinning – non commercial	8
Annual costs	Protection and management	
	Land rental (if applicable)	
Cash Inflows	Thinning revenue	18
	Revenue from poles	26
	Revenue from 1st harvest	34
	Revenue from 2nd harvest	60

Step 2: Estimating cash outflows

Estimates are now made of the likely quantum of the cash outflows associated with the categories of Table 1. This has drawn on information from a number of sources, e.g. quotes sought for the cost of establishing the plantation and for pruning costs based on past experience. Table 2 provides the financial estimates of each of these activities provided by the consultant. For convenience, all estimates are expressed on a per hectare basis.

Table 2. Cash outflows and timing associated with a two-species plantation

Cost group	Nature of cash outflow	Timing (yr)	Cost ($/ha)
Establishment costs	Planning and design	0	74
	Incidental clearing	0	158
	Site preparation and cultivation	0	265
	Cover crop establishment	0	88
	Pre-plant weed control	0	92
	Cost of plants	0	450
	Planting and refilling	0	645
	Post plant weed control	0	540
	Fertilizer	0	83
	Fencing	0	560
	Sub-total		2955
Maintenance costs	Post plant weed control (1)	1	1300
	Post plant weed control (2)	2	800
	Post plant weed control (3)	3	200
	First prune(plus certification)	2	800
	Second prune (plus certification)	4	800
	Third prune (plus certification)	6	800
	Thinning – non-commercial	8	500
Annual costs	Protection and management		40
	Land rental		0

Step 3: Estimating cash inflows

Cash inflows arise from the harvest of the trees. Harvest revenue is determined by the volume of timber (in m^3) produced multiplied by the stumpage price paid ($/$m^3$). For example, if a commercial thinning occurs at year 17 (as in the current example) and yields 170 m^3 of timber with an estimated sale (stumpage) price of $30/$m^3$, this will result in estimated cash inflows of $5100/ha.

Estimates of cash inflows are particularly difficult to make for forestry investments. The long production cycle means that it is difficult to estimate what stumpage prices will be many years into the future. In rare cases, long-term supply contracts may be signed with a guaranteed sale price. Even in these circumstances the uncertainty associated with harvest volumes means there remains considerable uncertainty when estimating cash inflows from harvests. Some revenues may be obtained from commercial thinnings part way through the production cycle, though stumpage price is usually low due to small diameter trees and low timber quality. Typically there is no market for thinnings of a very young age, in which case the thinning process results in a net cash outflow. This is the case in year 8 of the current example where a non-commercial thin is required costing an estimated $500.

The largest cash inflows from plantations will come at the end of the production cycle. In the current example, final harvest revenue takes place after 60 years, although another significant harvest occurs at 34 years. For this case study, estimates of harvest volumes and timing were collected as part of the Delphi survey undertaken by the consultants. These estimates, combined with estimates of future timber prices (in nominal dollars and also collected as part of the survey) can be used to estimate cash inflows (Table 3). The expert panel used in the Delphi survey thought high quality sawlogs of a native hardwood produced at year 34 should achieve a price of $200/$m^3$. Furthermore, it was thought that it was likely that logs harvested at 60 years would be suitable for the production of veneer, and attract a premium of 50% above the price of high quality sawlogs.

Table 3. Estimated cash inflows for 1000 ha plantation

Activity resulting in cash inflow	Year of harvest	No. stems/ha	Yield (m^3/ha)	Stumpage ($/$m^3$)	Revenue ($1000)
First thinning	17	170	170	$30	$5,100
Second thinning (poles)	26	85	--	$148 per pole	$12,580
First harvest (sawlogs)	34	42	100	$200	$20,000
Second harvest (sawlogs/veneer logs)	60	43	270	$300	$81,000

Step 4: Developing the financial model

For simplicity, the tax component of the analysis has been simplified. It is assumed that all cash outflows are fully allowable as deductions in the year that they are paid. In most years there is no revenue from the plantation against which to offset these tax losses. It is however assumed these losses can be claimed against income generated from other company operations, producing a tax benefit equivalent to the amount of the net cash outflows multiplied by the prevailing tax rate (30%).

A number of other assumptions have been made in the analysis that are worth mentioning. A nominal rate of interest of 7% has been used which includes a risk premium of 3%. Prices of all cash-flow items are assumed to increase over time at a rate equal to the underlying inflation rate, so that no adjustment of either cash inflows or outflows are needed.

The computer spreadsheet package Excel provides a convenient platform for financial analysis, to calculate key financial parameters such as NPV and IRR. Such a spreadsheet is presented as Table 4.

Step 5: Undertake a sensitivity analysis

Once the financial model has been set up, it is a simple task to examine at the effect of changes in parameter levels on the project performance criteria. This includes analysis with respect to required rate of return and with respect to parameters which are not under the control of the company.

The Excel Table function has been used to derive NPV values for a range of discount rates, and these have been plotted in Figure 2. Note that the IRR is the discount rate for which NPV is zero, i.e. the point where the curve in Figure 2 crosses the x-axis. This analysis suggests that the project is marginal in terms of financial acceptability. At the real rate of return of 7% required by management, the NPV for the project is -$58,214, and the IRR is just below 7% (6.96%). Figure 2 also presents the *land expectation value* (LEV) or *site value* for the project. This is the NPV for an infinite sequence of identical rotations, and is useful to compare forestry projects of unequal duration. The LEV will be considerably higher than the NPV for short rotations, but for long rotations (such as in the case study project) LEV will differ little from NPV.

Parameters outside the control of the company which are likely to have most effect on NPV have been identified, and pessimistic, most likely and optimistic levels identified for each, as in Table 5.

The spreadsheet used in the calculation of NPV and IRR for the most likely values in Step 4 has been used to recalculate these values for the optimistic and pessimistic values for each of the parameters of Table 6. Only one variable is changed at a time, while all the other variables are held at their most likely values. The ‘Scenario’ function in Excel allows multiple scenarios to be developed and the results reported in a table in a separate spreadsheet. This function has been used to undertake the sensitivity analysis, results of which are reported in Table 6.

Table 4. NPV calculations for forestry investment, with 7% discount rate ($1000s)

CASH FLOW ITEM	Timing of cash flow (end of year)																
	0	1	2	3	4	5	6	7	8	9-16	17	18-25	26	27-33	34	35-59	60
Annual protection and management		-40	-40	-40	-40	-40	-40	-40	-40	-40	-40	-40	-40	-40	-40	-40	-40
ESTABLISHMENT COSTS (Capital Costs)	-2955																
OPERATING COSTS																	
Post plant weed control		-1300	-800	-200													
First prune(plus certification)			-600														
Second prune (plus certification)					-600												
Third prune (plus certification)							-600										
Thinning (non-commercial)									-500								
Land rental (capitalised @ 4% of $2000/ha)																	
TOTAL OPERATING EXPENSES		-1340	-1440	-240	-640	-40	-640	-40	-540	-40	-40	-40	-40	-40	-40	-40	-40
OPERATING REVENUE																	
Thinning revenue 1 (year 17)											5100						
Thinning revenue 2 (year 26)													12580				
Harvest Revenue 1 (year 34)															20000		
Harvest Revenue 2 (year 60)																	81000
Total OPERATING REVENUE	0	0	0	0	0	0	0	0	0	0	5100	0	12580	0	20000	0	81000
TAX PAID OR TAX BENEFIT (30%)	886.5	402	432	72	192	12	192	12	162	12	-1518	12	-3762	12	-5988	12	-24288
NET CASHFLOWS (Operating revenue - capital outlays - operating costs)	-2068.5	-938	-1008	-168	-448	-28	-448	-28	-378	-28	3452	-28	8778	-28	13972	-28	56672
NET PRESENT VALUE OF CASH OUTFLOWS AT DISCOUNT RATE SPECIFIED	-58.214																
INTERNAL RATE OF RETURN	6.96%																

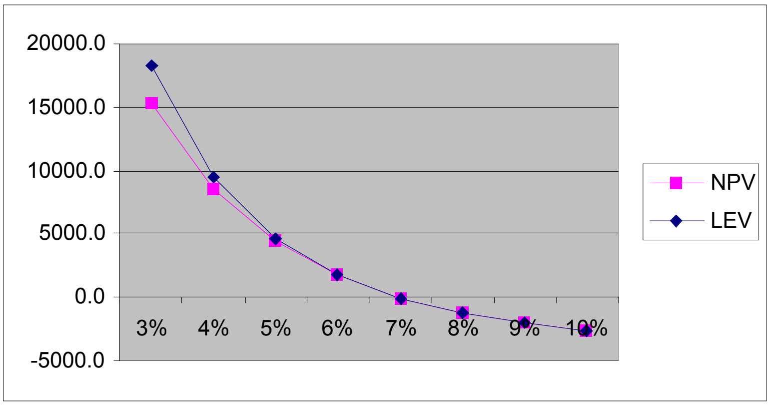


Figure 2. NPV Profile of FVC Ltd. forestry investment ($1000)

Table 5. Parameters selected for sensitivity analysis

Uncontrollable variable	Pessimistic ($1000)	Most likely ($1000)	Optimistic ($1000)
Stumpage price, first thinning ($/m^3)	25	30	35
Stumpage price, poles ($/pole)	110	148	200
Stumpage price, first harvest ($/m^3)	100	200	300
Stumpage price, second harvest ($/m^3)	150	300	450
Yield, first thinning (m^3)	120	170	190
Yield, poles (number)	70	85	85
Yield, first harvest) (m^3)	80	100	150
Yield, second harvest (m^3)	220	270	350
Establishment costs ($)	3455	2955	2655
Post plant weed control, year 1 ($)	1800	1300	1000
Post plant weed control, year ($)	1100	800	600
Post plant weed control, year 3 ($)	400	200	0
Pruning, years 2, 4, 6 ($)	1000	600	500
Thinning costs, year 8 ($)	800	500	400

Table 6. NPVs for forestry investment for pessimistic and optimistic parameter levels ($1000)

Parameter	Pessimistic estimate	Optimistic estimate	Range	Rank
Stumpage price, first thinning ($/m^3)	-247	130	377	10
Stumpage price, poles ($/pole)	-448	475	922	4
Stumpage price, first harvest ($/m^3)	-760	643	1403	1
Stumpage price, second harvest ($/m^3)	-547	431	978	3
Yield, first thinning (m^3)	-391	75	465	9
Yield, poles (number)	-326	-58	268	12
Yield, first harvest) (m^3)	-339	643	982	2
Yield, second harvest (m^3)	-239	232	471	8
Establishment costs ($)	-408	152	560	6
Post plant weed control, year 1 ($)	-385	138	523	7
Post plant weed control, year ($)	-242	64	306	11
Post plant weed control, year 3 ($)	-172	56	229	13
Pruning, years 2, 4, 6 ($)	-703	103	806	5
Thinning costs, year 8 ($)	-180	-17	163	14

The results of the sensitivity analysis could be now used by the company as a guide to which parameters need to be investigated further. From the sensitivity analysis it is clear that the stumpage prices for the first and second harvests and for poles, along with the yield for the first harvest, are the parameters that have the greatest effect on project NPV. The company may investigate ways to reduce the impact of uncertainty with respect to these parameters, such as through investing further resources into developing more accurate yield predictions or investigating further projections of future timber prices. It may also use the existing data to undertake an investigation of the three most critical parameters at a greater number of values within the range from pessimistic to optimistic.

7. CONCLUDING COMMENTS

This module has shown that the approach required for financial analysis differs somewhat from that of cost-benefit analysis. Key differences arise in treatment of social and environmental externalities and transfer payments such as taxation. In corporate project evaluation, often real discount rates are used, with allowance for differing inflation rates with respect to cost and revenue items. The discount rate may also differ from that employed in CBA.

The case study illustrates a stepwise application of a spreadsheet package to perform financial analysis including determining performance criteria such as NPV and IRR and carrying out sensitivity analysis is relatively straightforward, providing reliable cost and revenue data can be obtained for a forestry project.

17. Forestry Applications of Linear Programming

Steve Harrison

Linear programming (LP) is a highly versatile mathematical optimization technique which has found wide use in management and economics. It is used both as a research technique and as a planning tool, particularly at the individual firm and industry level. In general, LP is designed to maximize or minimize a linear objective function subject to a set of linear constraints. Linear programming is one of a group of techniques which may be referred to as *mathematical programming*. Other related techniques are goal programming, mixed integer programming and quadratic programming. Some typical applications of linear programming include:

1) determining the most profitable combination of enterprise or activity levels for a business firm with limited supplies of various resources
2) determining the most profitable investment portfolio, given the amount of investment capital available, rates of return on various stocks, bonds and other 'paper assets', and limits on high-risk investments.
3) formulating mixtures to combine ingredients such that a required overall composition of the mix is satisfied at least cost. Important applications are fuel and fertilizer blending and determination of livestock rations or supplementary feeds.
4) scheduling the various tasks in a construction project so as to complete the overall project in minimal time or at minimal cost
5) determining the location and size of storage facilities and processing plants together with the distribution pattern, so as to minimize the total of transport, storage and processing costs.

As illustrated by this list, the range of applications of linear programming is indeed wide, and this is one of the most widely used operations research techniques. In this module, the algebraic formulation of LP models in explained, and a simplistic decision problem is formulated in a linear programming framework, and is then solved graphically to illustrate the basic principles of the technique. Computer solution of this problem using the *simplex* method is then demonstrated, with reference to setting up the model on a spreadsheet and interpreting the output. Further applications to transshipment modeling, capital budgeting and goal programming are then illustrated.

1. ALGEBRAIC FORMULATION OF LINEAR PROGRAMMING PROBLEMS

In mathematical terms, the purpose of linear programming is to optimize a linear *objective function* subject to a set of linear *constraints*. The objective function may, for example, express the profit from a combination of enterprises or activities, the *cost* of a combination of ingredients, or the *time* required to complete a series of tasks. In these contexts, optimization may mean either *maximization* of profits or *minimization* of costs or time. The nature of linear programming will be illustrated with reference to a profit maximization problem. A highly simplified example has been chosen deliberately, so as to illustrate the technique without undue computational distraction. More realistic applications will be developed later in this module.

Example 1

A cabinet-maker produces dining room suites and grandfather clocks out of Australian red cedar timber. He can obtain annual supplies of up to 600 linear metres (lm) of red cedar (one linear metre is equivalent to one metre in length by one

metre in width by one centimetre in depth). Up to 4000 hours of labour a year are available for operations such as sawing, joining and polishing. Each dining room suite requires 12 lm of timber, and each grandfather clock requires 7.5 lm of timber. One hundred hours of labour are required to produce a dining room suite, and 40 hours to produce a grandfather clock. The profit from each dining room suite is $900, and that from each grandfather clock is $450. The cabinet-maker wishes to know how many units of each furniture line to produce in order to maximize profits. Formulate this decision problem as a linear programming model.

Formulation of the model

Before this problem can be solved by graphical or other means, it must be expressed in algebraic form, i.e. as a model. The first step is to introduce an 'x' notation for the decision variables, here numbers of suites and clocks to be produced. Thus we let x_1 = number of dining room suites produced, and x_2 = number of grandfather clocks produced.

It is now possible to formulate an objective function which in this case is an equation defining total profit. The term 'profit' is used loosely here, in that the figures of $900 and $450 are more correctly called 'gross margins' for the two activities, i.e. they are returns net of variable but not fixed costs. In obtaining these figures, allowance is made for allocatable costs such as materials, labour and marketing costs, but not for overheads such as rent on premises or rates, depreciation of equipment, and accountancy. In this module the term net revenues will be used and the objective function will be referred to as a revenue function. If each dining room suite has a net revenue of $900, then the total of net revenues from producing (and selling) x_1 dining room suites will be 900 x_1 dollars. Similarly, the total net revenue from producing x_2 grandfather clocks will be 450 x_2. If the symbol Z is used to represent total net revenue, the objective function may be written as

$$Z = 900\,x_1 + 450\,x_2$$

The objective can now be identified more precisely as finding those values of x_1 and x_2 for which Z is a maximum, bearing in mind the restrictions on production imposed by limited supplies of timber and labour.

Resource restrictions also can be expressed in algebraic form. If x_1 dining room suites are produced, each requiring 12 lm of red cedar timber, then dining room suites will consume a total of 12 x_1 lm of timber. Similarly, if x_2 grandfather clocks are produced these will consume 7.5 x_2 lm of timber. The total amount of timber consumed cannot exceed the supply, so the production plan is constrained by the inequality expression

$$12\,x_1 + 7.5\,x_2 \le 600$$

The left-hand-side of this expression indicates the amount of timber which will be used for any production policy (combination of x_1 and x_2 levels); the right-hand-side indicates timber supply. This timber constraint ensures that the demand for timber cannot exceed the supply; any production plan consuming more timber would violates this constraint and would therefore be *infeasible*.

Similar reasoning can be applied to derive a labour constraint. Since suites and clocks require 100 and 40 manhours of labour respectively, and since the labour supply is 4000 manhours, feasible levels of x_1 and x_2 are bounded by

$$100\,x_1 + 40\,x_2 \le 4000$$

Two further constraints are necessary to define the decision problem fully. These are that the numbers of suites and clocks produced cannot be negative

i.e. $x_1 \ge 0$ and $x_2 \ge 0$

Non-negativity constraints may at first appear unnecessary in a practical sense; after all, it is not possible to produce negative numbers of suites or clocks. However, they must be included for mathematical completeness, to delineate fully the feasible region of production. The

cabinet-maker's decision problem may now be summarized as a linear programming model with an objective function, two activities (dining room suites and grandfather clocks), two resource constraints (timber and labour) and two non-negativity constraints, as follows:

$$\text{maximize revenue } Z = 900\, x_1 + 450\, x_2$$

subject to the resource constraints

$$12\, x_1 + 7.5\, x_2 \leq 600$$

$$100\, x_1 + 40\, x_2 \leq 4000$$

and non-negativity constraints

$$x_1 \geq 0, \;\; x_2 \geq 0$$

General algebraic formulation

For the more mathematically inclined, the above problem formulation may be generalized in abstract terms for greater numbers of activities and constraints. Suppose levels are to be determined for n *activities* (x_1, x_2, ... x_n) yielding individual *net revenues* c_1, c_2, ... c_n. Limited supplies are available of m *resources*, the supply levels being b_1, b_2, ... b_m. Each activity uses fixed amounts of resources; in particular, the requirement of resource i by one unit of activity j is a_{ij}. The elements of the matrix of a_{ij} values (for i=1 to m and j=1 to n) are known as *technical or input-output coefficients*. The linear programming model may now be written as

$$\text{maximize } Z = c_1\, x_1 + c_2\, x_2 + \ldots + c_n\, x_n$$

subject to the linear resource constraints

$$a_{11}\, x_1 + a_{12}\, x_2 + \ldots + a_{1n}\, x_n \leq b_1$$
$$a_{21}\, x_1 + a_{22}\, x_2 + \ldots + a_{2n}\, x_n \leq b$$
$$\vdots$$
$$a_{m1}\, x_1 + a_{m2}\, x_2 + \ldots + a_{mn}\, x_n \leq b_m$$

and the non-negativity constraints

$$x_1 \geq 0,\; x_2 \geq 0,\; \ldots\; x_n \geq 0.$$

For readers familiar with matrix algebra, the decision problem may be expressed more compactly in matrix notation, as

$$\text{maximize } Z = C^T X$$

$$\text{subject to } A\, X \leq B \text{ and } X \geq 0$$

where C is a vector of activity net revenues (and C^T is the transpose of C) ;
X is a vector of activity levels;
A is a matrix of input-output coefficients;
B is a vector of resource supplies;
and 0 is the null vector.

2. GRAPHICAL SOLUTION OF RESOURCE ALLOCATION MODELS

Having expressed the production planning problem as an algebraic model, it is now time to derive optimal levels of the decision or policy variables. For simple problems such as this (with two activities), solution by graphical means is possible. Graphical presentation provides a number of practical insights into the nature of the decision problem and its solution. For the above example, a graph is set up as in Figure 1, in which the number of dining room suites (x_1) is measured on the horizontal axis and the number of grandfather clocks (x_2) is recorded on the vertical axis. The resource and non-negativity constraints are entered as straight lines on this graph, delineating the *feasible region*.

To fix the position of the timber constraint line, note that the maximal number of suites which could be produced if all timber were devoted to suites is the total timber supply divided by the per unit demand, i.e. x_1 = 600/12 or 50. No timber would then be available for production of clocks, i.e. x_2 = 0. On the other hand, if no timber were devoted to suites (x_1 = 0) then up to x_2 = 600/7.5 or 80 clocks could be produced before the supply of timber became exhausted. The two end points of the timber constraint – (0,80) and (50,0) – are drawn in Figure 1. If half the timber were used for suites and half for clocks then up to 25 suites and 40 clocks could be produced, corresponding to midway on the straight line joining the above end points. In fact, any point on the straight line between (0,80) and (50,0) corresponds to a production plan in which exactly 600 lm of timber are used.

Any combination of x_1 and x_2 levels on or to the left of (below) this line is feasible in terms of the timber constraint $12x_1 + 7.5x_2 \leq 600$. Any point to the right would use more than 600 lm of timber and violate this constraint.

Applying similar reasoning to that above, the labour constraint $100x_1 + 40x_2 \leq 4000$ can be drawn as a straight line connecting $x_1 = 4000/100 = 40$ and $x_2 = 0$ on the horizontal axis to $x_2 = 4000/40 = 100$ and $x_1 = 0$ on the vertical axis, as in Figure 2. Any production plan represented by an (x_1,x_2) pair on or to the left of this line is feasible in terms of labour use, while any production plan represented by a point to the right is infeasible. The combined effect of the timber and labour constraints is to restrict the feasible region to on or to the left of the line BCD in Figure 2.

It was noted earlier that in addition to resource constraints, non-negativity constraints must be imposed on the feasible region, these being of the form $x_1 \geq 0$ and $x_2 \geq 0$. In Figure 2, the constraint $x_1 \geq 0$ confines solutions to on or to the right of the vertical axis, while $x_2 \geq 0$ means any point on or above the horizontal axis. The feasible region of production is now defined fully; it is the irregular area ABCD. Any numbers of suites and clocks corresponding to an (x_1,x_2) pair within or on the boundary of this region is feasible in that it will not use more timber or labour than is available. If a point corresponding to an (x_1,x_2) pair is not on the production frontier BCD then some timber and labour remain unused.

Having defined the feasible region it is now possible to determine the most profitable production policy, i.e. the policy for which the objective function Z takes as large a value as possible. Revenue considerations are introduced by drawing lines on the graph joining equally profitable (x_1,x_2) combinations, called *isorevenue lines*. The location of the initial isorevenue line is arbitrary. Suppose, for example, the policy of producing 30 dining room suites and no clocks were chosen; the total net revenue would then be

$$Z = 900(30) + 450(0) = 27,000 \text{ dollars}$$

This same level of revenue could be achieved by producing no suites and 27,000 / 450 or 60 clocks. In fact, any combination of (x_1, x_2) values along the line from (30,0) to (0,60) corresponds to a total net revenue of \$27,000. The \$27,000 isorevenue line is drawn as a broken line in Figure 3.

Various other isorevenue lines could be drawn, as illustrated in Figure 3. Each of these other lines is parallel to the \$27,000 isorevenue line. The higher the revenue the further the line from the origin. For example, the \$31,500 isorevenue line corresponds to the production of 35 suites or 70 clocks or any combination along the straight line joining these points. The most profitable plan would be indicated by the (x_1,x_2) pair touching the highest possible isorevenue line. A parallel shift to the right for the isorevenue line reveals that this must be at point C. Reading from the graph, the co-ordinates for point C correspond to an optimal plan of producing approximately 22 dining room suites and 44 grandfather clocks. The total net revenue from this plan is approximately 900(22) + 450(44) or 39,600 dollars.

It is to be noted that the above solution is approximate only, having been read imprecisely from a graph. Procedures are available for finding a more precise solution; this is to produce 44.44 clocks and 22.22 suites. However, in practice only whole numbers of suites and clocks can be produced, so these numbers could be rounded downwards to 44 and 22 respectively.

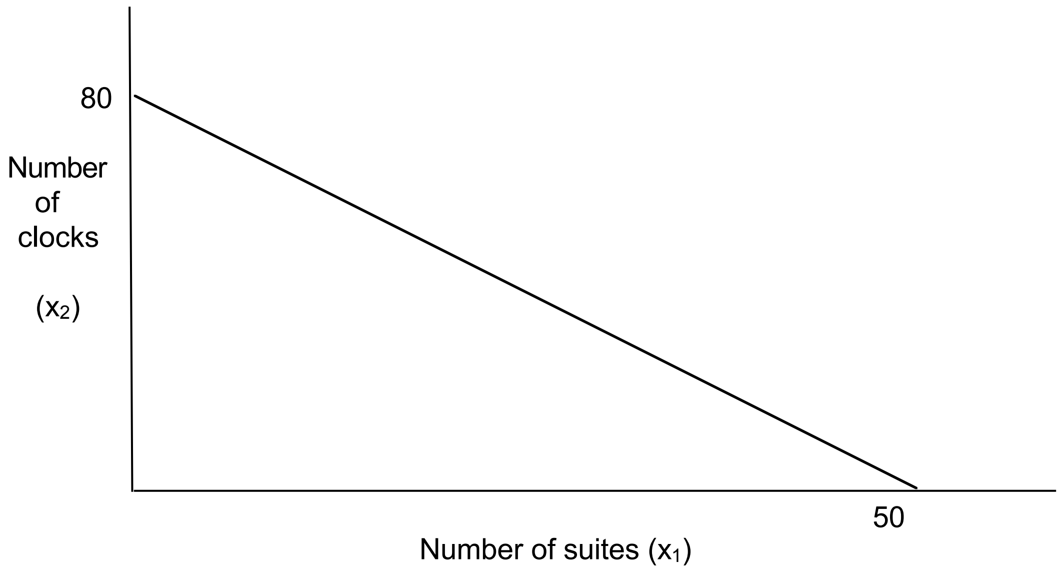


Figure1. Graphical representation of timber constraint

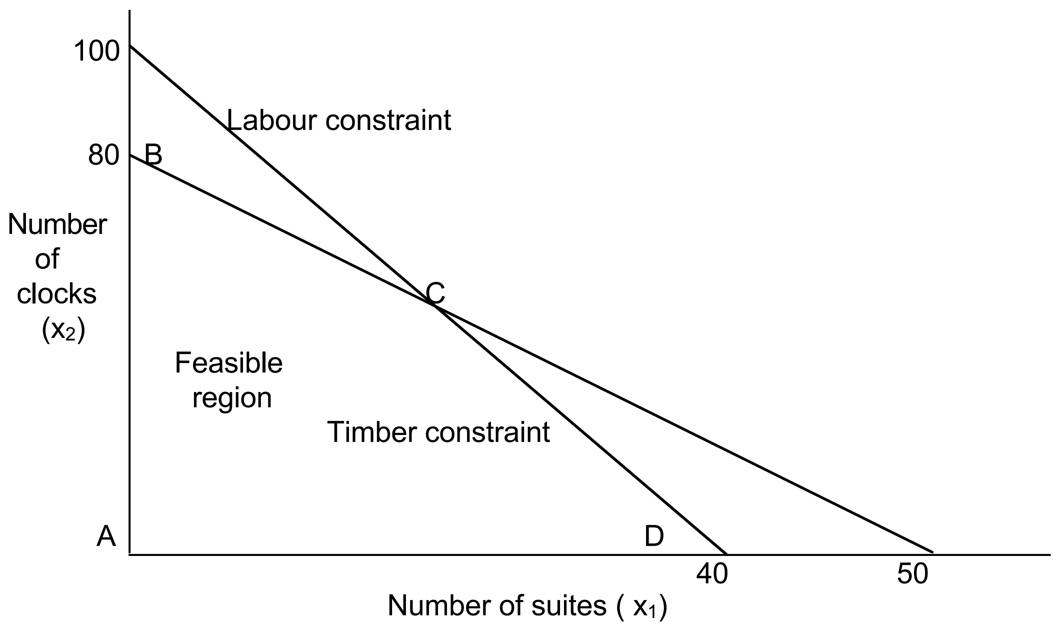


Figure 2. Feasible region of production (ABCD)

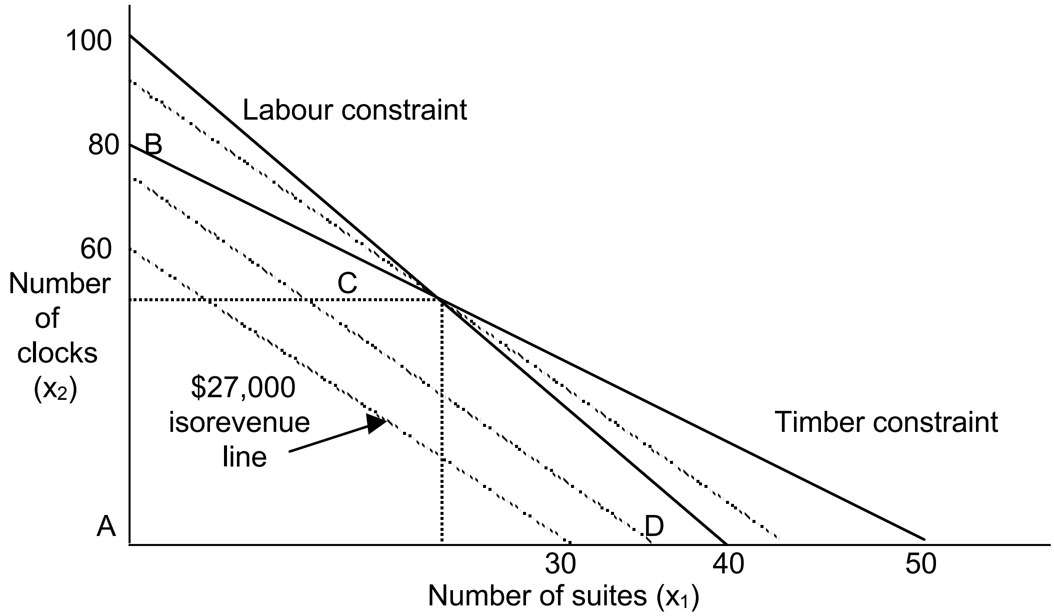


Figure 3. Isorevenue lines and optimal plan

Sensitivity or stability analysis

The graphical solution procedure illustrated above has identified the most profitable combination of activity levels for given net revenues, resource supplies, and so on. However, in practice the concept of a single best policy is of limited value; it is desirable to know under what circumstances this policy remains optimal, and how it would change if any of the parameters were varied. Decision makers frequently require information about the sensitivity or stability of the optimal plan with respect to changes in net revenues, resource supplies or input-output coefficients. Two relatively simple but highly useful forms of sensitivity analysis are *net revenue ranging* and calculation of *shadow prices*.

Net revenue ranging

The fact that point C is optimal is due to the relative net revenues of clocks and suites. If the relativity between net revenues were to change then the slope of the isorevenue line would change, and a different plan could become optimal. For example, suppose the net revenue of suites remains at $900 per unit, but that for clocks rises to $600. The $27,000 isorevenue line will now run from (30,0) on the horizontal axis to (0,45) on the vertical axis, as in Figure 4. A parallel shift outwards in the isorevenue line reveals that the optimal plan is now at point B, i.e. production of 80 clocks and no suites, with a total net revenue of 80 times $600 or $48,000.

By examining different slopes of isorevenue lines on a graph such as Figure 4, it is possible to determine within what range the net revenue of one activity can vary for a fixed net revenue of the other activity, while the current production plan remains optimal. The slope of any line on the (x_1,x_2) co-ordinate axis system is defined as the change in level of the variable on the vertical axis divided by the corresponding change in the variable on the horizontal axis. For the labour constraint, there is a decline in x_2 of 100 for an increase in x_1 of 40, and hence a slope of -100/40 or -2.5. Similarly, the slope of the timber constraint is -80/50 or -1.6. The slope of the isorevenue line is the negative of the ratio of net revenue of the activity on the horizontal axis to net revenue of the activity on the vertical axis, i.e. slope of isorevenue line

$$= - \frac{\text{net revenue of suites}}{\text{net revenue of clocks}} \quad (\text{or } - c_1/c_2)$$

$$= - \$900 / \$450 = -2$$

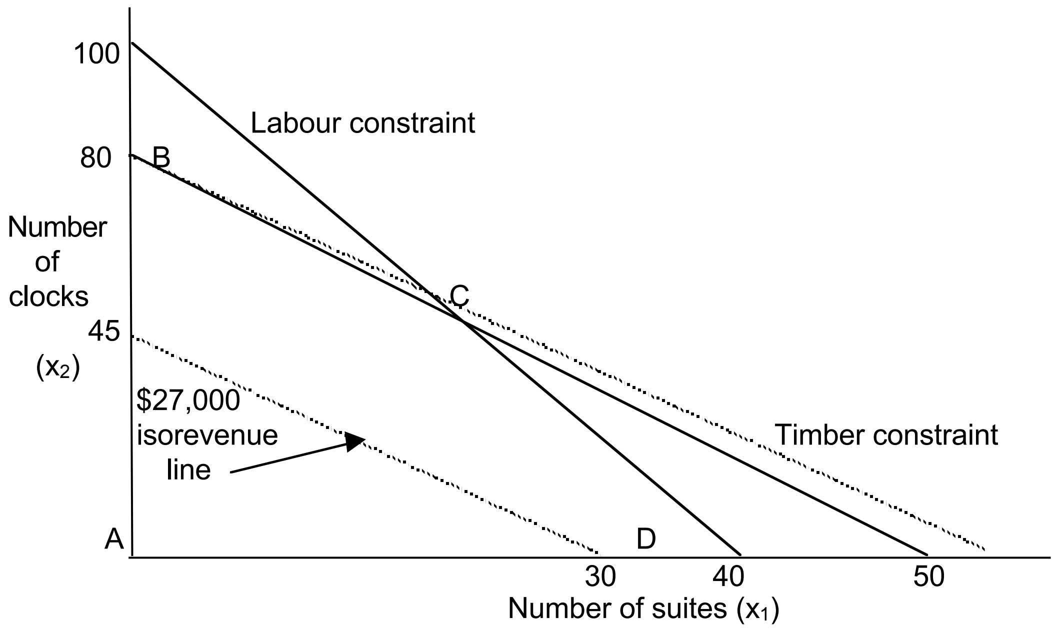


Figure 4. Determination of the optimal plan when the net revenue of clocks is increased to $600

Provided the slope of the isorevenue line remains between -2.5 and -1.6, the point furtherest from the origin touched by an isorevenue line will be point C. Now suppose the net revenue per suite is fixed by a sales contract at $900, but that for clocks is uncertain, depending on demand conditions. For point C to remain optimal, the net revenue of clocks must be such that the net revenue ratio remains between -2.5 and -1.6,

i.e. $-2.5 \le -c_1 / c_2 \le -1.6$

i.e. $2.5 \ge -c_1 / c_2 \ge 1.6$

Solving the two parts of this inequality expression for c_2, it is found that

$$c_2 \ge c_1 / 2.5 = \$900/2.5 = \$360, \text{ and}$$

$$c_2 \le c_1 / 1.6 = \$900/1.6 = \$562.50.$$

That is, the net revenue for clocks can vary in the range $360 to $562.50 without optimal activity levels changing.

A special case exists where the slope of the isorevenue line is exactly equal to that of a linear segment on the boundary of the feasible region. For example, if the net revenues of suites and clocks were $900 and $360 respectively then the slope of the isorevenue line would be -2.5, exactly matching that of segment CD on the boundary of the feasible region. In this case, any point along CD would be equally profitable; there is no unique optimal policy. Further, small variations in net revenues may result in large changes to the optimal policy. This, of course, is a an unusual situation in that it would only be by coincidence that the slopes of the isorevenue line and one of the constraint lines were exactly equal.

Shadow prices

The computer solution method for LP problems involves selecting activities to be brought into the basis (solution) only if their net revenue exceeds their opportunity cost. When an LP problem is solved, a number of activities may be absent from the optimal solution. These include real or production activities, and also what are called *disposal activities* or resource non-use activities. The latter are a device to convert inequalities into equations to provide an initial feasible solution for the simplex solution method. In this context, the constraints for the decision solved graphically above can

be written as

$$12\,x_1 + 7.5\,x_2 + 1\,x_3 = 600$$

$$100\,x_1 + 40\,x_2 + 1\,x_4 = 4000$$

where x_3 and x_4 are non-use or 'disposal' of timber and labour respectively.

In the case of a maximization problem, for a real activity, the shadow price is the amount by which the net revenue would be reduced if one unit of this activity were forced into the solution. Equivalently, it is the amount by which the net revenue would have to increase before this activity entered the optimal solution. For a disposal activity, the shadow price is the cost of forcing one unit of a resource into disposal or non-use, or equivalently the amount by which overall revenue would increase if another unit of the resource were available. The calculation of shadow prices will not be illustrated here, but the concept will be discussed further with respect to computer output.

3. APPLYING THE SIMPLEX METHOD TO RESOURCE ALLOCATION MODELS

The above method of solving linear programming problems is quick and simple when there are only two activities, even if there are many constraints. But if there are three or more activities, then the two-dimensional graphical approach breaks down. However, a mathematical procedure known as the simplex method has been devised which can be used to solve LP problems regardless of the numbers of activities and constraints (even when there are 1000 or more of each). The mathematical basis of the simplex solution *algorithm* involves advanced matrix algebra which will not be explained here. While it is desirable to have an understanding of the economic logic of the solution procedure, to use the technique it is only necessary to be able prepare the data (objective function, activities and constraints) for input to a computer package and to interpret the computer output.

Invariably, linear programming problems are solved using a computer package. A variety of computer programs are available for solving LP problems. Provided the problem is not too large – not more than a couple of hundred activities and constraints – Excel Solver may be used to obtain the optimal solution and shadow prices. For illustration purposes, an extension of the cabinet-maker's decision problem in which there are three activities and three constraints is presented below, and the method of using Excel Solver is demonstrated.

Example 2

Suppose the cabinet-maker of Example 1 can produce roll-top desks as well as dining room suites and grandfather clocks. Also, he wishes to distinguish between general labour (for sawing, joining and polishing) and specialized woodcarving labour used in decorating his furniture. Supplies of timber and general labour are as in Example 1, while 700 hrs per year of woodcarving labour are available. The net revenue from each desk is \$600. Suites and clocks require timber and general labour as in Example 1, plus five and eight hours of carving labour respectively per unit. Each desk requires 8 lm of timber, 60 hrs of carpentry labour and 7 hrs of carving labour. Assuming that the cabinet maker's objective is to maximize net revenue, determine the optimal production plan, net revenue ranges and shadow prices.

Solution using Excel Solver

Table 1 presents the linear programming formulation for this problem, set up on a spreadsheet ready for solution using Solver. The activities are represented across the columns of the tableau (columns B to D), and the resources down the rows (rows 5 to 7), with the objective function as row 8. The *technical coefficients* form the body of the tableau (cell B5 through to D7). Column G lists the initial resource supplies or constraint right-hand-sides.

Table 1. Extended cabinet maker's problem in spreadsheet format

	A	B	C	D	E	F	G
1	Extended cabinet maker's decision problem						
2							
3	Constraint or objective	Dining room suites	Grandfather clocks	Roll-top desks	Resource use	Sign	Resource supply
4	Activity level	0	0	0			
5	Timber	12	7.5	8	0	≤	600
6	General labour (hrs)	100	40	60	0	≤	4000
7	Carving labour (hrs)	5	8	7	0	≤	700
8	Net revenue	900	450	600	0		

Prior to using Excel Solver, it is necessary to introduce a row for activity levels (row 4); these activity levels are initially set at zero. It is also necessary to introduce a column for resource use (column E). As well, a column for signs of the constraints is introduced (column F), in this case containing only '≤' signs.[1]

The most complex step in setting up the initial tableau is to enter formulae in the 'Resource use' column, i.e. column E:

1. the resource use for the timber constraint (cell E5) is entered as the formula '=SUMPRODUCT(B$4:D$4,B5:D5)'. The cell value initially takes a level of zero, because the activity levels in row 4 are zero. Note that absolute cell references are required for row 4 (represented by the dollar sign before the row number).
2. the contents of cell E5 are then copied to cell E6 through to cell E8. The coefficients in cell E5 to E7 represent 'resource use' with respect to the resource constraints, while the value in cell E8 is the level of the objective function. Initial values in these cells are again zero.

Once these data have been entered onto the spreadsheet, Solver can be called up to further set up the problem for solution. Solver is to be found under the Tools menu of Excel. (If it is not currently available, seek assistance on how to access it.) Note that the general form of the Solver window is:

Set Target Cell: Equal to: By Changing Cells: Subject to the Constraints:

First declare the target cell. This is the objective function, and is chosen by clicking in the total net revenue cell, here E8. Next check that the maximization option of the 'Equal to' row has been chosen, with a dot in the 'Max' circle. Cells B4 to D4 are selected as the changing cells, i.e. the levels of the decision or x variables.

It is next necessary to add constraints. First click on the 'Add' button. When adding constraints, select corresponding cells in columns E and G. In the 'Add constraints' dialogue box, under 'Cell reference', click on the cell range E5 to E7, then under 'Constraint' click on cells G5 to G7. The sign '≤' is automatically selected as the default. Click on the Add or OK button to confirm that this constraint is to be added.

It is also necessary to add non-negativity constraints. Hence click on the Add button again, then select cells B4 to D4. In this case, the sign must be changed to '≥', and the value of zero must be entered on the right-hand-side (under 'Constraint'). The OK

[1] If not available in the Excel version, these symbols may be copied and pasted from a wordprocessor file. These signs are included for readability of the tableau only, and could be replaced by text, e.g. 'le' for 'less than or equal to'.

button is now pressed to confirm addition of this constraint. The constraints will now appear in the dialogue box. The Change button may be used to modify these constraints if there are any apparent errors.

The Solve option may now be used to obtain the optimal product mix. When clicking on the Solve button, the spreadsheet should change to that of Table 2. (If this is not the case, then there are errors in the tableau or Solver specifications such as errors in constraints which need to be tracked down.) The optimal solution is to produce 30.76 clocks and 46.15 desks (these levels could be truncated to whole numbers). The total net revenue is $41,538. All of the timber and general labour is used, but about 130 hours of carving labour are unused (or in disposal).

A little further information can be gleaned from through sensitivity or post-optimality analysis. This is not generated automatically – the 'Sensitivity' option must be chosen under the 'Reports' options in the 'Solver Results' dialogue box. It is possible to move between the spreadsheet and the reports by clicking on buttons near the bottom of the computer screen. Table 3 presents 'Sensitivity Report 1' for this decision problem. In the first section of this table, in the Reduced Gradient column, it is found that the net revenue of dining room suites would have to increase by a little over $69 for this activity to enter the basis or solution. In the second section, under Lagrange Multiplier, it is found that the reduction in total net revenue would be $23.08 if one unit of timber were removed and $6.92 if one hour of general labour were removed. These are the amounts the cabinet-maker could afford to pay for additional units of these resources and still be able to use them profitably.

As indicated by Example 2, inclusion of additional activities and constraints paves the way to more realistic formulation of decision problems. Of course, if there are large numbers of activities and constraints then there will be large numbers of columns and rows in the simplex tableau, and checking of the tableau becomes more critical. Even small linear programming problems are usually solved on a computer, and commercial computer packages of varying levels of sophistication are available. Some of these packages perform a more detailed post-optimality analysis than that presented above.

4. PRACTICAL DIFFICULTIES IN LINEAR PROGRAMMING ANALYSIS

The linear programming technique as outlined above is mathematically elegant, computer packages to derive optimal policies are readily available, and the technique is widely used. However, the user needs to be aware of a few theoretical limitations and occasional computational difficulties.

Correct perception of the role of linear programming

Linear programming is capable of simultaneously evaluating the profitability of large numbers of activities in the presence of numerous constraints. It can solve decision problems that are beyond the power of intuition, pencil-and-paper methods and formal budgeting. However, the role of this powerful decision aid is often poorly understood. The analyst carrying out a linear programming study in not usually the person responsible for making policy decisions and taking the consequences of these decisions. It should not be expected that the optimal plan derived from a single linear programming analysis will be implemented in precise detail by management. Decision-makers are likely to have particular hunches and preconceptions as to the best course of action. Results generated by a linear programming study typically are used to reinforce or challenge existing views and tentative plans, i.e. to provide decision support. Further, decision-makers often wish to ask a number of 'what if' type questions, e.g. 'What if suites can be produced with 90 rather than 100 hours of labour?' or 'What if the net revenue of clocks is $500 rather than $450?'.

Table 2. LP spreadsheet after solution by Solver

	A	B	C	D	E	F	G
1	Extended cabinet-maker's decision problem						
2							
3	Constraint or objective	Dining room suites	Grandfather clocks	Roll-top desks	Resource use	Sign	Resource supply
4	Activity level	0	30.769231	46.1538			
5	Timber	12	7.5	8	600	≤	600
6	General labour (hrs)	100	40	60	4000	≤	4000
7	Carving labour (hrs)	5	8	7	569.231	≤	700
8	Net revenue	900	450	600	41538.5		

Table 3. Sensitivity report for cabinet-maker's decision problem, as generated by Excel Solver

Cell	Name	Final Value	Reduced Gradient
$B$4	Activity level Dining room suites	0	-69.23083027
$C$4	Activity level Grandfather clocks	30.76923077	0
$D$4	Activity level Roll-top desks	46.15384615	0
Cell	Name	Final Value	Lagrange Multiplier
$E$5	Timber Resource use	600	23.07692308
$E$6	General labour (hrs) Resource use	4000	6.923076923
$E$7	Carving labour (hrs) Resource use	569.2307692	0

Results of post-optimality analysis and further computer solution runs will help to shed light on these questions. LP tends to be used interactively to explore a number of variations to a basic model in a single session on the computer, and to produce a good deal of information about optimal activity levels and sensitivity of profits or costs to variations in parameter levels (data estimates or assumptions). The broad picture which is built up of the alternatives and consequences can be of considerable help in guiding decision making.

Assumptions of the technique

The linear programming model presented above has implicitly relied upon a number of assumptions, the acceptability of which may be questioned in particular applications.

Planning horizon. In production planning applications of linear programming, net revenues make no allowance for overhead costs. That is, the firm is assumed to possess a fixed asset structure and the analysis concentrates on how current resources should be deployed to maximize profits in the short term. The examples which have been presented all rely on single-period or static models. In the longer term the firm may acquire more land, factories, machines and so on. Planning these capital acquisitions over time is a problem of a different order to maximizing gross margins for a single production period.

Single-valued expectations. It has been assumed that gross margins, resource supplies and input-output coefficients are known with certainty. In practice these are often estimated with a good deal of guesswork. Further, single point estimates are made when in reality it would be more

meaningful to specify a probability distribution of values, particularly in the case of net revenues.

Linear objective function. The objective function of the decision maker is assumed to be a linear function, and to depend on net revenues or input costs alone. Particularly in enterprise planning applications, the decision-maker may be concerned about both level of income and variability of income over time. Activities which have relatively low expected net revenues may be chosen because of low risk associated with this revenue. Suppose in Example 2 that suites have an assured market outlet at a pre-arranged price but clocks are sold directly by the manufacturer, at prices which can vary widely depending on the whim of the market. He may then produce more suites than indicated in the 'optimal' plan because of the lower uncertainty about the net revenue of suites.

Additivity within activities. The amount of each resource used per unit of each activity is assumed to be constant regardless of the level at which the activity is conducted. If one suite requires 12 lm of timber and 100 hrs of labour then 20 suites require 240 lm of timber and 2000 hrs of labour. In practice, increasing or decreasing returns to variable factors are common, e.g. as more suites are produced the labour input per suite may be reduced due to more streamlined production.

Fixed resource proportions. The input-output coefficients are constant within each activity vector, e.g. fixed proportions of timber and labour are used to produce each suite. It may be possible to use timber more efficiently, including using offcuts, if greater time is taken in carpentry and joining.

Independence of activities. It is assumed that the level at which any activity is conducted has no effect on input levels or revenue of any other activity, i.e. there is no complementarity between activities. If offcuts from production of suites could be used in production of clocks, then the timber requirement of the two activities would be less than that of either activity in isolation.

Divisibility. Resource inputs and activity levels are assumed to take a continuous range of fractional units. In Example 1, the 4000 hours of labour may be supplied by two full-time tradesmen, and any labour supply level which is not a multiple of 2000 hrs may not be practicable. Also, the solution obtained for this decision problem includes 30.77 clocks and 46.15 desks. However, clocks and desks can only be produced in whole numbers. An approximate integer solution can be obtained by truncating the final values of the decision variables, which here would yield 30 clocks and 46 desks. But we cannot be certain that this is the optimal integer plan.

In practice, procedures have been devised to overcome most if not all of the limitations implied by these assumptions, though at the cost of greater tableau complexity. Multi-period models may be constructed to allow for expansion in the fixed asset base over time. Variations in factor proportions may be readily accommodated, e.g. if labour and timber can be combined in different proportions to produce suites then two or more suite activities, each having different input-output coefficient vectors, may be defined. If complementarity exists between two activities then these may be combined into a single activity producing two products. Optimal integer levels may be obtained using the mixed integer programming technique. The point to be made here is that one should not be discouraged from using the linear programming by what appear superficially to limitations of the technique.

Some other potential problems

Sometimes a linear programming tableau is constructed which possesses either no solution or no unique optimal solution. For example, consider the following models:

Model 1

maximize $Z = x_1 + x_2$
subject to $2\,x_1 + x_2 \geq 7$
$2\,x_1 + x_2 \leq 6$
and $x_1 \geq 0,\ x_2 \geq 0$

Model 2

maximize $Z = x_1 + x_2$
subject to $2\,x_1 + 3\,x_2 \geq 7$
$4\,x_1 + x_2 \geq 6$
and $x_1 \geq 0,\ x_2 \geq 0$

The constraints of Model 1 are contradictory since $2\,x_1 + x_2$ cannot be at the same time greater than or equal to seven, and less than or equal to six, so this problem has no feasible solution. The objective function of Model 2 can be increased without limit, since no upper bounds are imposed on the values that x_1 and x_2 can take; this problem is said to be unbounded. The usual cause of these unsolvable problems is an error in specification of the model. For example, it is simply not possible to have unlimited resource supplies as implied by an unbounded problem.

Other problems can arise during solution such as rounding errors and 'degeneracy'. Specific procedures are available to overcome these. They will not be discussed further here.

5. OPTIMAL TRANSPORTATION AND LOCATIONAL EFFICIENCY MODELS

The purpose of a transportation model is usually to determine the transport or 'shipping' allocation or a commodity such that quantities are moved from various origins to various destinations at minimum cost.

Example 3

Suppose log timber is to be transported from various production regions to various markets in a small country, as in Table 4. The two domestic supply regions – central and islands – can produce an annual turnoff of 100,000 and 120,000 cubic metres (m^3) respectively, and up to 200,000 m^3 can be obtained annually as imports. Demands are 200,000 m^3 per year in the northern industrial region, 100,000 m^3/year in the western region and 70,000 m^3/year in the south and east region.

Solution

An initial transportation table is set up as in Table 4. This summarises the timber supplies from the various origins and demands at the various destinations. Transport costs from origin to destination in dollars per cubic metre are as indicated in the body of the table, e.g. \$40/m^3 for 'shipping' from the central region to the northern industrial region.

The optimal transport allocation can be obtained by linear programming. If we let

Z = total transport cost
x_{ij} = number of units transported from origin i to destination j (1000m^3)
c_{ij} = cost of transport from origin i to destination j (\$/m^3)

then this decision problem may be represented algebraically as:

minimize overall transport cost

$$Z = 40x_{11} + 25x_{12} + 20x_{13} + 30x_{21} + \ldots + 60x_{33}$$

subject to supply and demand constraints

$$x_{11} + x_{12} + x_{13} \leq 100$$
$$x_{21} + x_{22} + x_{23} \leq 120$$
$$x_{31} + x_{32} + x_{33} \leq 200$$
$$x_{11} + x_{21} + x_{31} \geq 200$$
$$x_{12} + x_{22} + x_{32} \geq 100$$
$$x_{13} + x_{23} + x_{33} \geq 70$$

plus non-negativity constraints on all nine transport quantities ($x_{11} \geq 0, \ldots x_{33} \geq 0$).

Since the quantities are in terms of thousands of cubic metres, the objective function will be expressed in terms of \$1000s.

Table 4. Log supplies, demands and transport costs

Source	Destination			Supply (1000 m^3)
	Northern industrial region	Western region	South and east	
Central region	*40*	*25*	*20*	100
Islands	*30*	*20*	*35*	120
Imports	*100*	*100*	*60*	200
Demand (1000m^3)	200	100	70	

The initial spreadsheet for this model is set up as in Table 5. The transport quantities are initially set to zero (cells B6 to D8). A column for sums of supply allocations (column E) is added, with elements set as the supply sum for each origin, e.g. cell E6 is set equal to the sums of the values of cells B6 to D6. Similarly, a row for demand allocations (row 9) is added, with elements equal to the sums allocated for the various destinations, e.g. the value in cell B9 is set equal to the sum of values in cells B6 to B8. Transport costs are entered as a separate block (cells B12 to D14). A total transport cost (objective function) cell is set up) with the formula

'=SUMPRODUCT(B6:D8,B12:D14)'.

Solver is now called up under the Tools menu, and the various entries are made in the dialog box. The target cell is set as cell C16, and the 'Min' option is chosen. The changing cells are selected as the range B6:D8. The constraints can now be entered. The column vector E6:E8 is set less than or equal to F6:F8 for supply constraints, the row vector B9:D9 is set greater than B10:D10 for the demand constraints, and the block of changing cells B6:D8 is set greater than or equal to zero.

On solving this LP problem, Table 6 is obtained. No timber is to be transported from the central region to the northern region (allocation in cell B6 = x_{11} = 0), 100 units (thousand cubic metres) are transported from the central region to the western region (allocation in cell C6 = x_{12} = 100) and so on. Although up to 200 units of timber may be imported, the demands are satisfied by imports of 150 units (cell E8).

A sensitivity table could be generated, and this would indicate for example the amount by which transport costs for each non-used shipment path would have to fall before transport through this path would become warranted (would reduce overall cost).

Transhipment models

The structure of transhipment models is similar to that of transportation models, except that intermediate destinations are added where product is stored or processed. For example, timber could be produced in a number of areas, processed into plywood at a specific location, and then transported to other areas for marketing. Changes in volume may take place with storage or processing, in which case it is necessary to standardize the volume units. In the case of processing, since the optimal shipment path will depend on processing cost at the various processing locations, but processing cost will vary with volume of throughput, it is usually necessary to adopt a trial-and-error solution approach.

6. CAPITAL BUDGETING OR PORTFOLIO SELECTION MODELS

Another useful application of linear programming is to assist in the selection of investment projects. In this application, the investment projects are treated as separate activities, and the objective function is defined in terms of net present values.

Table 5. Initial spreadsheet for timber transportation model

	A	B	C	D	E	F
1	Timber transportation model					
2						
3	Source		Destination			
4		Northern industrial region	Western region	South and east	Sum of allocations	Supply
5						
6	Central region	0	0	0	0	100
7	Islands	0	0	0	0	120
8	Imports	0	0	0	0	200
9	Sum of allocations	0	0	0		
10	Demand	200	100	70		
11						
12	Central region	40	25	20		
13	Islands	30	20	35		
14	Imports	100	100	60		
15						
16	Total transport cost =		0			

Table 6. Optimal log transport allocation

	A	B	C	D	E	F
1	Timber transportation model					
2						
3	Source		Destination			
4		Northern industrial region	Western region	South and east	Sum of allocations	Supply
5						
6	Central region	0	100	0	100	100
7	Islands	120	0	0	120	120
8	Imports	80	0	70	150	200
9	Sum of allocations	200	100	70		
10	Demand	200	100	70		
11						
12		40	25	20		
13		30	20	35		
14		100	100	60		
15						
16	Total transport cost =		18300			

Example 4

Three plantation type options are available for 200 ha of community land, viz. fast-growing 'softwood' species (e.g. Gmelina)[2], hardwoods (e.g. Mahogany) or mixed species planting. Resource constraints include land, labour, and capital in the first two years of plantation establishment. (The technical coefficients are provided in Table 7.) Over a 20-year period, two rotations of softwoods or one rotation of the other plantation types could be grown. $400,000 in capital is available for use in the first two years, and may be supplemented by borrowing of up to $50,000 in each of these years. The natural resource management agency wishes to determine what combination of the plantation types would provide the greatest aggregate NPV.

Solution

The initial spreadsheet for this decision problem is provided as Table 7. This is in a similar format to the earlier resource allocation model, except that a few features have been added:

1. Provision is made to transfer unused capital from the first year to the second. This takes up or uses capital in the first year, but makes capital available in the second year, which is achieved by having a +1 coefficient for the first year and a –1 coefficient for the second year.

2. Borrowing activities are introduced, with units of $1000. These are 'supply' activities, with negative coefficients (-1s) in the relevant capital rows.

3. The objective function coefficients for the borrowing activities are the loan interest and redemption costs, in year 20 currency equivalents. As an approximation, these are taken as negative NPVs equal to the amounts borrowed.[3]

4. Constraint rows are introduced for the borrowing activities, with +1 coefficients in the columns of the borrowing activities, so as to limit the amounts borrowed.

When this problem is solved, the spreadsheet of Table 8 is obtained. The optimal investment is to choose softwoods only, plant all 200 ha, and borrow $50,000 in each year. Not all of the available labour is used. An aggregate NPV of $5.9M is predicted.

Extensions to this portfolio selection model include specifying dependent or mutually exclusive projects, and specifying that if a project is to be introduced then the level must be sufficiently high to warrant the tooling up, i.e. imposing a minimum threshold level. For example, it might be decided that at most two of the three plantation type options considered above can be adopted, and that the minimum area for any plantation type is 50 ha. These requirements can easily be built into the model if mixed-integer programming – available within Excel Solver – is used.

Portfolio selection models can also be designed to take risk on asset returns into account. For example, the variance as well as the expected payoff for each investment may be estimated, and quadratic programming used to determine the optimal investment portfolio given the decision-maker's degree of risk aversion.

---

[2] Gmelina is not technically a softwood, but has similar wood properties.

[3] Calculation of objective function coefficient for the borrowing activity is rather complex. This is the present value of the interest plus redemption payments at the end of 20 years, and will depend on when the repayment is made and what is the inflation rate. If it is assumed that the loan interest rate is equal to the inflation rate, then a simple approximation may be made. If a loan were taken out at the beginning of a year, at say 15%, and paid at the end of the year with interest (i.e. $1.15 repaid for each dollar valued), the present value of the repayment would be $1.15/1.15 or $1. Similarly, if a loan were repaid after 20 years, and interest were compounded, the repayment would have a present value of $1.

Table 7. Initial tableau for forestry investment portfolio decision problem

	A	B	C	D	E	G	H	H	I	J
1	Portfolio selection model									
2										
3	Constraint or objective	Soft-woods	Natives	Mixed species	Transfer capital, yr 1->2	Borrow capital, year 1 ($1000)	Borrow capital, year 2 ($1000)	Res-ource use	Sign	Resource supply
4										
5	Activity level	0	0	0	0	0	0			
6	Land (ha)	1	1	1				0	≤	200
7	Labour (person days)	45	30	35				0	≤	10000
8	Capital, year 1 ($1000)	2	2.5	2.3	1	-1		0	≤	400
9	Capital, year 2 ($1000)	0.5	0.5	0.7	-1		-1	0	≤	0
10	Max. loan, year 1					1		0	≤	50
11	Max. loan, year 2						1	0	≤	50
12	NPV ($1000)	30	28	24		-1	-1	0		

Table 8. Final tableau for forestry investment portfolio decision problem

	A	B	C	D	E	G	H	H	I	J
1	Portfolio selection model									
2										
3	Constraint or objective	Soft-woods	Natives	Mixed species	Transfer capital, yr 1->2	Borrow capital, year 1 ($1000)	Borrow capital, year 2 ($1000)	Res-ource use	Sign	Resource supply
4										
5	Activity level	200	0	0	50	50	50			
6	Land (ha)	1	1	1				200	≤	200
7	Labour (person days)	45	30	35				9000	≤	10000
8	Capital, year 1 ($1000)	2	2.5	2.3	1	-1		400	≤	400
9	Capital, year 2 ($1000)	0.5	0.5	0.7	-1		-1	0	≤	0
10	Max. loan, year 1					1		50	≤	50
11	Max. loan, year 2						1	50	≤	50
12	NPV ($1000)	30	28	24		-1	-1	5900		

7. GOAL PROGRAMMING

A more systematic way of dealing with multiple goals is through *goal programming* (GP). Here, a number of goals are identified, and a target or *aspiration level* is specified for each. *Deviational activities* are then introduced in the tableau to allow for underachievement or overachievement of goals relative to the aspiration levels. In particular, an *underachievement activity* is introduced for each aspiration floor, and an *overachievement activity* is introduced for each aspiration ceiling. A series of constraints is then set up for the goals. In the objective function, the coefficients are not the activity payoffs. Rather, they are the costs of underachievement and overachievement of goals relative to the aspiration levels. The objective function now

states that the sum of shortfalls and over-runs, with appropriate coefficients for each, be minimized.

Because of the presence of deviational variables in goal programming, the constraints are referred to as *soft constraints*, c.f. the *hard constraints* for resource limits. Even if some or all the aspiration levels cannot be met, the problem will not be found to be infeasible (with an error message when an attempt is made to find the optimal solution). Rather, levels of the deviational variables in the solution will indicate the extent to which it is not possible to meet the goal targets.

The question arises as to what coefficients or weights to place on the deviational activities. Shortfalls can be expressed in a variety of different units, such as dollars, rate of return as a percentage, jobs and hectares. One approach would be to place a weight on each deviational variable to represent the cost of goal under-achievement or overachievement. Applying a system of weights to these is known as *weighted goal programming*. In practice, expressing the deviations in a common unit – say dollars – may be difficult and unnecessary.

Often a priority ordering can be established between goal dimensions, referred to as *lexicographic* or *preemptive* goal programming. For example, the requirement may be to achieve the target dividend rate first, then NPV payoff, then job creation, then the environmental requirement. This could be achieved by placing weights of different orders of magnitude on the shortfall activities. However, software has been developed in which the priority ordering is stated as input data and the solution algorithm enforces this priority ordering of goals. Often, it is possible to solve lexicographic goal programming problems by using an appropriate system of weighting on deviational variables (e.g. $1M per unit shortfall for top priority goals, $0.1M for second priority goals, and so on.)

Example 5

A government wishes to support local development projects with employment and environmental benefits. Three projects are identified which will contributed to these goals, viz. road construction, tree planting and public housing. Each kilometre of road construction provides 200 jobs (person months), has net carbon release of 2 tonnes, and requires expenditure of $100,000. Each hectare of tree planting generates 3 jobs, sequesters 5 tonnes of carbon per year and requires $2000 of capital. Each house constructed provides 20 jobs and requires expenditure of $30,000. Determine the government program which meets these goals most closely, if equal weights are attached to the three goals.

Solution

The initial tableau is set out as Table 9. Note that shortfall activities are included for jobs and carbon sequestration (with coefficients of +1 in the relevant constraint rows), and an expenditure over-run activity is included (with coefficient of +1 in the expenditure row). The objective is to minimize the sum of deviations. Note that the constraints are in inequality form, rather than forcing exact equalities. The constraints are set up in the usual way, i.e. the formula

'=SUMPRODUCT(B$4:G$4,B5:G5)

is entered in cell H5, and copied into cells H6 to H8. For the objective function, this in effect means summing the numbers of units for the three deviation variables (values in cells E4 to G4). The objective function is one of minimization.

The solution to this problem is reported in the spreadsheet of Table 10. The recommended portfolio includes about 3.5 km of road construction and a little over 100 ha of tree planting. The jobs and carbon sequestration goals are met, but there is an expenditure over-run of a little over $50,000.

f any goal were considered more important than another, this could be reflected by weights of different orders of magnitude in the objective function. For example, if the expenditure control was a critical goal, a weight of say 1000 could be entered in cell G8.

Table 9. Initial tableau of local government jobs and environment program

	A	B	C	D	E	F	G	H	I	J
1	Goal programming example									
2										
3	Constraint or activity level	Road constn. (km)	Tree planting (ha)	House constn. (houses)	Job short-fall	Carbon seqn. shortfall	Expen diture over-run	Res. use	Sign	Resource supply
4	Activity level	0	0	0	0	0	0			
5	Jobs (person months)	200	3	20	1			0	_	1000
6	Carbon sequestration (tonnes)	-2	5	0		1		0	_	500
7	Expenditure ($1000)	100	2	30			-1	0	≤	500
8	Deviation				1	1	1	0		

Table 10. Final tableau for local government jobs and environment program

	A	B	C	D	E	F	G	H	I	J
1	Goal programming example									
2										
3	Constraint or activity level	Road constn. (km)	Tree planting (ha)	House constn. (houses)	Job short-fall	Carbon seqn. shortfall	Expen diture over-run	Res. use	Sign	Resource supply
4	Activity level	3.48	101.39	0	0	0	50.70			
5	Jobs (person months)	200	3	20	1			1000	_	1000
6	Carbon sequestration (tonnes)	-2	5	0		1		500	_	500
7	Expenditure ($1000)	100	2	30			-1	500	≤	500
8	Deviation				1	1	1	50.70		

8. CONCLUDING COMMENTS

In this module, the highly versatile technique of linear programming has been introduced with reference to a simple profit maximization problem involving two production activities. This decision problem can be expressed as an algebraic model, and solved graphically, providing insights into the nature of the resource allocation problem. The simplex method, which is quite general in that it can solve linear programming problems containing any numbers of activities and constraints, and is readily available in computer packages. Sensitivity analysis has been shown to provide further information about solution stability and opportunities for short-run expansion of production.

Extensions to the linear programming technique to handle more complex applications include transportation modeling, portfolio selection and goal programming. The technique has great potential for modeling decision situations in forest management and policy.

FURTHER READING

Dayananda, D., Irons, R., Harrison, S., Herbohn, J. and Rowland, P. (in press), *Capital Budgeting: Financial Appraisal of Investment Projects*, Cambridge University Press, Cambridge.

Dent, J.B., Harrison, S.R. and Woodford, K.B. (1986), *Farm Planning with Linear Programming*: *Concept and Practice*, Butterworth, Sydney.

Pannell, D. (1997), *Introduction to Practical Linear Programming*, Wiley, New York.

18. Team Building, Grant Seeking and Project Administration

Steve Harrison and John Herbohn

The various research techniques introduced in earlier modules need to be viewed in the broader context of how they will be supported and applied. Doing research is a little like running a business. No matter how good the ideas are and how technically proficient one might be at doing the job, the success of the business depends largely on the structures that are put into place to get the job done. To draw an analogy, doing research in isolation is the business equivalent of running a small manufacturing business as a sole trader. To expand the business beyond this one needs to work with or employ other people (i.e. build or join a team) and capital is needed to do this (i.e. research grants), and this capital needs to be used efficiently and effectively and in accordance with the guidelines of those people (c.f. banks) that provided it (i.e. project management). The researcher then needs to produce a product and sell this product into the market place to generate a return to the business (i.e. completing research and publishing it). This module discusses the issues of research team building, obtaining funds and subsequent project management. In the following module, issues associated with documenting and disseminating the research findings are discussed.

1. WORKING IN RESEARCH TEAMS

Nowadays, it is increasingly difficult for an individual working largely on their own to carry out high level, high impact original research. Much of the successful research effort arises from highly focussed and cohesive research teams. Hence it is appropriate to consider how teams are formed and maintained, and the strengths and weaknesses of team research.

Team building and maintenance

Research teams are often formally brought together for particular tasks, e.g. public inquiries. The interest in this module is with more informal teams, driven by mutual interest. Teams sometimes emerge spontaneously, where common interests become apparent. Research funding arrangements also favour team building. For example, the cooperative research centre (CRC) funding arrangements of the Federal government in Australia have been responsible for providing long-term funding to bring together teams of up to about 200 researchers, in different institutions and different states, to focus research in related themes. An example is the Rainforest CRC in Queensland, which has programs on rainforest ecology, entomology, threatened and threatening species, water quality, tourism, reforestation and other areas. The ACIAR model is also a valuable contribution to team building. Large grants which continue for a number of years, and provide funding for travel of researchers to meetings and field sites, and for research officers, are critical for teamwork. Success in obtaining large grants is critical to success in team formation and continuation.

In that the research grant scene is difficult to break into, and favours researchers with an established track record, it is often an advisable strategy for more junior faculty members to team up with experienced researchers.

While a university research team will normally be coordinated by established academics, research officers or research assistants[1] and postgraduate students (which can be overlapping roles) play an important role in these teams. They can bring a large amount of energy, initiative and dedication, and in practical terms are a cost-effective way of adding value to research funding. The involvement of high-calibre postgraduate students can make a major difference to research output. Often potential postgraduates have little

[1] Higher education workers or HEWs in the Australian University vernacular.

experience in designing research projects and are looking for guidance on what types of projects are feasible. One strategy for obtaining postgraduate students to work in areas of interest in our Rainforest CRC programs has been to develop a portfolio or potential research projects, and to provide this to people who make inquiries about postgraduate study. Joining an established team also provides advantages for a post-graduate student. Importantly it gives them access to team members with a range of skills and backgrounds. If reseachers have interesting projects available, and some funds available to support the research, then this can be attractive to postgraduate students looking for a suitable thesis project. If the postgraduate students have a scholarship, and some top-up funding can be provided in the form of research assistanct pay, this further improves the arrangement and hence the ability to attract capable people.

Whether teams continue to exist depends critically on effort put in by individual members, continuation of a mutual research interest, and success in obtaining grants and securing publications.

Advantages of working in teams

Working in a research team can have a number of advantages:

Specialisation. A team benefits from having members with specialist skills, e.g. team members can specialise in technical areas (such as silviculture), in developing research ideas, in grant seeking and budget preparation, survey methods and statistical analysis, computing skills, text generation and editing, report writing ability, promotion of research activities and findings and so on. It is unlikely that a single individual will be sufficiently multi-talented to cover all relevant areas at a high level of competency.

Critical mass. Some advantages exist from having a number of people available to take part in research activities. This allows progress to continue when other work pressures create unavoidable diversions for some of the team. The variety of skills becomes apparent to funding bodies and to agencies seeking consultancy services. There is a greater probability that someone in the team (as distinct from an individual working alone) will be aware of highly relevant literature and of experts who should be contacted. There is a greater probability that calls for grant applications will be detected and that proposals can be put together to meet deadlines. A critical mass of people can also lead to a critical mass of funding, whereby one or more full-time research officer can be employed, and this can have major benefits in terms of research continuity, record keeping, communication with client groups, and outreach (extension) activities.

Synergies. Perhaps the greatest advantage of research teams is the synergies which result between members. Members can ‘bounce research ideas’ off each other, comment on each other’s work and writings, challenge unsound thinking (an important validation of research), and provoke deeper thinking on a topic. These synergies can lead to more thorough analysis and writing-up, and to identification of further promising areas of research.

Strategic alliances. Some members of a team may be members of other professional organisations and research units, expanding the range of research contacts. Some may be members of journal editorial boards, or have good contacts with journal editors, or be members of departmental or external committees which judge research fund applications, and so on.

Grant success. The combined research capability and track record of a research group tends to more impressive than that of individuals, which can be important for attracting interesting work and research funding.

Publication success. Working as a team which includes researchers with a strong publication ‘track record’ can lead to higher success rates in publication of journal articles and book chapters.

Disadvantages of working in teams

There are some disadvantages in working in teams. Teams are made up of individuals with a range of personal characteristics and

idiosyncrasies. Successful teams that work well together are invariably composed of people who get on well together and enjoy working collaboratively. However, even in the most successful of teams there will be personal conflict and tension at times. To make teams work, group members must be willing to compromise and be able to put in place mechanisms to deal with negative characteristics of other team members, i.e. be tolerant. If this is not done, many of the benefits of teams can be lost and much time and energy is wasted in conflict situations.

Mentoring of junior team members

In research teams, there is a requirement of more established researchers to provide encouragement, support and training for more junior members. This can be viewed as in part a social responsibility in academia. However, it is also important from a self-interest viewpoint, of having capable and productive colleagues with a favourable attitude to project work who are likely to remain with a project rather than seek alternative work. Some practical mentoring steps include:

- making research resources available wherever possible, e.g. funds for fieldwork.
- being reasonably accessible to discuss research activities.
- providing a sympathetic listener and counselor when personal problems arise.
- committing team members to seminar or conference presentations.
- providing referee support and review comments on grant (and scholarship) applications.
- including team members in joint applications for research funding.
- directing members to useful contacts or relevant reading.
- providing comments on draft papers – within days not weeks – and making suggestions for publication outlets
- including members as authors in research papers. More established researchers often have better access to publication outlets. Developing confidence to prepare and submit papers can take some time and training.
- providing positive reinforcement wherever possible, e.g. making a point of commenting on successful achievements, such as papers published. (In some cases, it will be possible to provide financial incentives.)
- should the opportunity arise, providing funding support for key events, such as an international conference.
- providing positive confidential referee reports when these are needed.

Where younger project workers are postgraduate students, the supervision and mentoring roles tend to overlap. Supervisors and mentors can leave a lasting impression on those that they work with. However, there is a wide range of performance by higher degree supervisors, from totally inert to totally 'switched on'. In team situations it is thus critical to ensure that strong supervision and ready help is provided to students. Postgraduates share information readily with other existing and potential postgraduates. If sound supervision is provided, then the word of mouth advertising by existing and past students can be a highly effective means of encouraging new postgraduates to join the team. The opposite also applies.

2. STRATEGIES FOR OBTAINING RESEARCH GRANTS

Adequate funding is a critical ingredient to achieving research objectives. Even with sufficient time, skills and commitment, it is difficult to carry out sound research without sufficient money for fieldwork travel and equipment, research assistance, computer equipment, postage and printing, purchase

of research books, and so on. Attending conferences can be critical for keeping abreast with what is happening in a research field, and conferences are invariably expensive to attend. While one's employment agency generally has some accessible research funds, the extent of paper work for a modest grant can be discouraging, and there may be limited flexibility in how the funds can be used. These are best thought of as 'seed funds', to assist in developing research concepts and proposals. All this leads to the conclusion that obtaining external research grants is critical to a healthy research program.

Grant seeking involves identifying opportunities for funding, developing research proposals, and frequently making modifications or preparing rejoinders in response to feedback from the funding body or appraisers of the application.
In some cases, preparing research proposals can take several weeks of work – for literature search, drafting the proposal, collecting cost data and preparing the budget, and so on - possibly as much time as it takes to write a research paper. Hence grant seeking needs to be approached strategically, and not entered into light-heartedly. It is wize to make an estimate of the probability of success before launching into development of a research proposal, but bearing in mind that a proposal rejected by one agency may be reworked for another. Given the amount of effort that can go into developing a convincing research proposal, it has been commented that one has to do the research before obtaining the grant. Following on from this (tongue in cheek) view, when a grant is received, it is used to carry out the research needed to apply for the next grant!

Choosing the research area

It is obviously preferable to choose a research area with which the team has expertise and is comfortable, to take advantage of their comparative advantage. Sometimes, a tradeoff arises between choosing areas of greatest interest and areas for which funding is most readily obtainable. Sometimes, involvement in projects that are well funded but not quite in the area of interest of the researchers can be advantageous; a researcher or team with small amounts of funding has limited flexibility in what projects then can choose. Many grants have some degree of flexibility in what can be done with the funds once they are awarded, which may allow opportunity to explore issues of particular interest. Sometimes, modifications to project objectives will be agreed upon, depending on what lines of research turn out to be feasible.

Identifying funding sources and prospects for success

A surprisingly large number of potential research funding sources exist – including local, national and international bodies – and new ones continue to arise. This suggests that it is necessary to be vigilant about identifying what grant are potentially accessible, and to establish a database of grant sources and application deadlines. Most funding bodies have one or two calls a year, with specific submission dates.

A key strategy is to get to know the more promising agencies which are sources in relation to one's research interests, and to understand their priority funding areas. (This is sometimes disclosed on a Web site.) It is also helpful to know other idiosyncrasies of the funding body, e.g. the type of emphasis they like in project proposals, and what they like and dislike. Some funding agencies are particularly keen to see any related research highlighted in the proposal; some like inter-agency grant arrangements; some place great emphasis on having a technology transfer component in the research; some have a highly applied focus (e.g. 'getting trees in the ground'). Some do not like to see involvement of postgraduate students in the proposed research activities, taking the view that such a proposal is a disguised form of application for a student scholarship.

Choosing which grants to apply for is an economic problem, of cost of developing the proposal versus expected payoff in research dollars. It is worth bearing in mind that large grants are not necessarily more difficult to obtain than small ones, and that the amount of effort involved in obtaining small grants is sometimes a questionable

use of time.

It is helpful to know what sizes of grants are supported. Sometimes successful applicants will be given the full amount they request, and sometimes amounts will be scaled down, which creates problems when a specific set of objectives and activities in proposed. For some of the more prestigious grants, the probability of success is discouragingly low. For example, only about 20% of applications for Australian Research Council grants are successful.

Being known to the agency can be a critical factor in grant success. Some agencies have a rather narrow researcher base, and like to continue funding people who have a record of achieving their research objectives. Successful researchers often receive invitations to apply for grants.

It would appear that funding agencies are highly risk averse with respect to whom they grant funds. In one case, a funding agency commented that if they received a particularly interesting application from an unknown researcher, they would invite someone well known by the agency to carry out the proposed research! This reinforces the view that it is difficult to break into the research funding circle, and that success in obtaining grants favours further success. One way to break into this circle is to prepare joint applications with more established researchers. The fact that a credible research team is being put up in the proposal is also important in having the agency look favourably on the request.

The probability of grant success can be increased greatly by including established researchers in the application. Usually, grant success is related (directly or indirectly) to track record. For example, the Australian Research Council (ARC) council recently placed 50% weight in scoring applications on the proposal and 50% on the track record of the applicants.

Preparing the proposal

Funding applications tend to take a relatively standard layout. Some of the things which are often included are: background and literature review, research questions, aims and objectives, research method, significance of the research, technical outputs and practical outcomes, technology transfer method, timetable, Gantt chart indicating predecessor activities and milestones, budget, justification for budget items, references, ethical clearance statement, statement of employer's support, and applicants' track records.

The typical weak spots in applications are: lack of a theoretical or conceptual model underlying the proposed research; the proposed research method is too vague and lacking in detail; and the justification for budget is weak (e.g. it says how the money will be spent rather than why it is required). Once a proposal has been prepared, it is useful to review the text in these areas of often-encountered weaknesses.

Some expenditure areas tend to be well accepted while other raise suspicions of the appraisers. The more acceptable expenditure areas: research assistance, fieldwork equipment and travel, interview surveys, purchase of research books, maintenance (phone, fax, photocopying, postage), and project workshops. Agencies funding international projects are generally supportive of project workshops (planning, training, end-of-project). Requests for funding of conferences (particularly overseas ones) and notebook computers are likely to be viewed critically. Sometimes the travel budget is considered excessive. Some agencies give applicant the opportunity to place priorities on the various budget items, as A, B, etc. In such cases, it is wize to place the lower priorities on the items which are generally less favourably viewed. Where expenditure items have to be prioritised, amounts granted may be less than the total amount applied for.

A useful recent development is that many funding bodies have a staged application procedure. An initial expression of research interest of not more than about two pages in length may be called. It the proposed work is judged competent and fits within the priority areas of the funding bodies, then the applicant is invited to submit an expanded application. In some cases, there may be three or four stages to the application, the downside of which is that it may take about two years for the project to finally obtain approval.

It often happens that grant applications are 'recycled'. That is, if an application is rejected by one funding body, it is reworked and submitted in response to the call for applications by another agency. This means that a further payback can be obtained from the sunk costs of development of an unsuccessful research proposal, and helps justify the large effort expended in chasing prestigious but low probability grants. In the case of large grants (e.g. ACIAR grants) several weeks may be required on the proposal and even the budget may take a number of weeks to prepare. This should be considered as part of the research process, not as dead time. That is, the project proposal should be considered as a working document, which is part of the research process, and which can if necessary be recycled for another funding opportunity, or form the basis of a publication. Persistence in grant seeking often pays off.

A mistake which is sometimes made in preparing the project budget is to ask for too little funds, in the hope that this will increase the probability of securing a grant. This means that if the application is successful, the researcher is stuck with trying to achieve promised outputs with inadequate resources. We were caught in this way in a recent study of tribal villages in India, but fortunately managed to obtain supplementary funding from another source.

It is a valuable step to obtain a peer review of a grant proposal. Some organisations arrange this as a matter of course. Someone who has not been involved in the preparation of the proposal is more likely to detect omissions, unclear statements and statements which could be misunderstood or cause a negative reaction. It is not uncommon for people promise too much, i.e. to propose to undertake more than they could reasonably achieve with the funds and time available.

Ethical clearance is increasingly being required for grant applications, even to undertake surveys (e.g. of households or farmers).

Responding to requests for more information or modifications

Funding bodies vary in their decision processes. In some cases, a detailed proposal is prepared, and an accept/reject decision is made, with no interaction with the applicant. At the other extreme, an officer from the funding body interacts repeatedly with the applicant to steer the project document and budget through the various approval stages. ACIAR is excellent in providing support for applicants in this regard. In most cases, some amount of questioning of the proposal arise, with opportunity for revisions or clarifications. When proposals are sent to external reviewers, there may be an opportunity for the applicants to prepare a rejoinder to the appraisers' comments.

In the case of projects involving researchers from more than one agency, of from two countries, there may be negotiations on the research scope to satisfy the needs of each agency or country.

Generally, when a project and grant is approved in principle, it is necessary to obtain statements of agreement from the various parties. For example, the office of research of a university normally has to sign off for a project which involves university staff and resources. Concerns can arise about ethical considerations (clearance from an ethics committee may be required) and any restrictions which are imposed by the funding body on publication of research findings.

3. RESEARCH ADMINISTRATION

In any research project, there is a considerable amount of administration and reporting (in addition disseminating information about the research and its achievements). These tasks involve planning of research activities, management of people and funds, monitoring project progress, and reporting to officers of the host organisation and funding bodies.

Planning research activities

Poor planning can result in both wasted funds and wasted time – both of which are important commodities for a researcher.

The need to plan data collection and analysis activities has been discussed elsewhere. There is also a similar need to plan the expenditure of funds. Where travel is involved, particularly international travel, a great deal of time can be spent in making travel arrangements. If field visits and research trips are not planned well they can involve a considerable waste of time. For instance, a visit to a field site to talk to community members can be a complete waste of time if the key people being sought are away.

Personnel management

In general, team research is a cooperative and collegial activity, rather than a management-and-line-worker process. This is particularly the case when the team members have confidence in each other's abilities and dedication. Team management then involves taking a close interest in colleagues' activities, and providing encouragement and support where possible, rather than attempting to make directives. Research tends to be most productive when there is an open and inquiring atmosphere and people have the opportunity to explore ideas. Coordination then takes place through planning meetings and informal discussion.

Managing the budget

This can be one of the most difficult areas in management of research projects. In one sense, the funds for a project are fully allocated at the time of funding approval, hence the task is to ensure that the funds are used in the way contracted between the researchers and the funding body. Fortunately, more flexibility than this usually exists, and modifications to project objectives as well as unforeseen circumstances alter the way in which funds are used.

Where requests for supplementary funding by team participants are made, it is desirable to meet these to the extent possible given the project commitments and finance constraint.

Making claims for research expenses, and acquittal for funds advanced or disbursed can be an onerous task for researchers, but is a necessary part of public accountability.

Project reporting

Invariably, there is a requirement for an end-of-project report to the funding body. For projects with a duration of more than one year, there is usually annual reporting of expenditure and of performance in relation to milestones. When field trips are made, there may be a requirement for individual trip reports. (This is a requirement of ACIAR, for example.) If project workshops are conducted, there will be an expectation of some form of workshop proceedings or report. There may be an expectation of production of a technical monograph at the end of the project. In general, the funding body will want to ensure that the money is well spent, and produces useful findings or 'makes a difference' to a target group of stakeholders, and provides favourable publicity for the activities of the funding body.

Project monitoring and staying on track in terms of budget, outputs and timetable

Experience indicates that there is often 'slippage' in projects relative to research intentions. While some delays and expenditure over-runs can be tolerated, the project objectives and milestones need to be kept in sight. There will be adverse impacts on researchers if a project does not achieve its objectives, such as: questions asked by funding bodies about annual reports; imposition of a project review by an independent expert group; early termination of a project; more difficulty in obtaining further funding.

Dealing with project reviews

Successful performance in reviews is important for securing further funding. Reviews sometimes arise because the funding body is not happy about the way a project is progressing, so need to be handled with care. There is a need to provide details of all achievements of the project (e.g. research technical outputs and publications, technology transfer, capacity building) of difficulties encountered, and of modifications of objectives.

4. CONCLUDING COMMENTS

There are no guaranteed strategies for research success, but experience indicates a number of lessons on what measures are likely to assist. Success is a matter of both opportunities and effort. It has been said that ‘genius is 10% inspiration and 90% perspiration’. Success in research requires commitment and persistence. In that much interesting research findings and ideas fail to appear in print, if one critical tip can be given, it is to become a habitual writer.

5. DISCUSSION QUESTIONS

1. What are the promising sources of research funds for your organisation?

2. Which journals would you target for research outputs?

3. To what extent are you involved in mentoring colleagues, and through what measures?

4. Do you undertake research in a teamwork situation? If so, what special skills do individuals in your team possess?

5. Consider the following text passage. Suppose you are given the task of providing editorial comments to the author. What aspects of the passage would you suggest be reworded?

   ESTABLISHMENT

   ‘Brown (1988) notes that plantation establishment was not good; there were significant seedling losses, due to the hot temperatures.

   Therefore, although the seedlings were very cheap, the plantation was not expected to be profitable.’

19. Report Writing and Publication Strategy

Steve Harrison and John Herbohn

In an academic world, it is critical to develop a research 'track record', a key indicator of which is the publications list. This can influence job satisfaction, promotion opportunities, and success in obtaining grants and consultancies. Publications are probably the single most important means by which researchers in universities, colleges and research institutions are evaluated. Publishing research results sends a signal to potential research funders and employers that the researcher is capable of bringing a research project to a conclusion and can produce a tangible outcome. Peer review provides a critical validation of the research methods and findings. Publications are also a major element in the transfer of technology – such as new processes and new understanding of system behaviour and management techniques – to potential users. Particularly when grant funding has been obtained, publication is expected, and is the requirement for further funding. Writing papers is a critical task for researchers. Many people do excellent conceptualisation, literature reviews, statistical analysis and other research activities, but fail to document their findings in a non-perishable and widely available form. Some people are 'writerholics' who write through compulsion; others are strongly disinclined to 'put pen to paper' and will always find something else pressing to do rather than write up their research. Writing is to some extent a matter of habit. However, some tips can be given to improve writing skills. While success in research is to a large extent a matter of commitment and perseverance, a number of strategies may be employed to increase effectiveness. Different strategies work best for different people, but a number of observations may be made as to how to be more effective in this endeavor. This module examines various aspects writing up and publishing research findings. The observations made here represent to some extent the personal views and experiences of the authors, and are designed to provoke interest and discussion on research publication strategies.

1. WRITING GENRE

The researcher must bear in mind the style of writing which is most appropriate for the particular type or output. For example, when writing a thesis, there is little pressure on space, but a strong imperative to demonstrate that all relevant aspects have been considered and all assumptions made clear (a defensive form of writing). There is also need for a considerable amount of *motivataion* (saying why material is presented) and *signposting* (saying how material is set out). Journal articles on the other hand are expected to be highly organised and concise; journals are expensive to produce and have tight space limitations. Differences exist between different types of journals; science journals tend to follow the 'aims – materials – method – results' format, while journals in the social sciences and humanities are a little more expansive. Working papers and conference papers tend to be written more quickly with less finishing-off effort than journal articles. When writing textbooks or other teaching materials, it is generally necessary to provide extensive explanation to make sure that readers can understand the concepts being presented, and to reinforce this with a large number of examples and perhaps case studies. Material which is intended for a multidisciplinary audience must contain relatively simple explanations. In the case of extension material, this generally has to be written in a more popular and catchy style.

2. CRAFTING A RESEARCH PAPER

A research paper – e.g. journal article, conference paper, discussion paper – tends to have a relatively standard structure. This varies somewhat between journals, and disciplines, so the following suggestions won't find precisely for every case, but are based on experience for a variety of publication outlets.

Typical sections of a research paper are: an abstract (of about 150 words, sometimes followed by keywords and journal classification codes[1]), acknowledgments (usually either a footnote to the first page or a section before the references), introduction, research method, a number of sections forming the body of the paper, discussion, concluding comments, references, and perhaps one or more appendix.

The *Abstract* should be a brief overview of what the paper is about, a note on the methodololgy, and a concise summary of the main findings, such that a potential reader can decide whether to read the paper in detail. The *Introduction* should indicate what the paper is about in general terms, perhaps review some relevant literature, perhaps mention areas of controversy of important information gaps, indicate in general terms what methodology has been adopted, and state how the paper is set out (the sequence of sections). The amount of detail will vary between short communications and longer papers. These two sections (Abstract and Introduction) are often badly crafted in draft papers.

Literature review material may appear in a separate section, but often this is contained both in the Introduction and throughout other sections of the paper. It is important to make clear the *Methodology* which has been adopted, often as a separate section. If the paper reports survey findings, then the study area, sampling frame, survey method, response rate, and method of data analysis would normally be indicated. In a review paper, there may be no need for a Methodology section.

It is difficult to lay down guidelines as to what structure the body of the paper will take. If results are being presented – e.g. from experiments, modelling or surveys – then a critical reviewer would normally look to see if there was any statistical analysis (such as *t*-tests, regression analysis or ANOVA). In qualitative research or review papers, it may not be possible to provide this type of analysis. The main results of the research are normally presented in the body of the paper, sometimes as a separate section.

The *Discussion* section provides an opportunity to expand on research results, and any limitations of the methodology and findings, and to highlight the main implications for management or policy.

The *Conclusions* section tends to summarise and remind the reader of the main impressions they should take away from reading the paper. This section is often brief, and the heading *Conclusion* (singular) or *Concluding comments* is sometimes used, to signify the section concludes the paper, and does not necessarily present detailed conclusions from the research. A guiding principle for this section is 'say what you found, not what you did' – don't repeat comments on the methodology or paper layout.

Appendices are useful when the author wants to include information which is more technical and could be useful for a specialist reader, or which is necessary to support the main arguments, but would interrupt the flow of text if included in the body of the paper. This keeps the paper concise, but provides an opportunity to include the additional information.

3. HANDLING REFERENCES

The handling of references can present many difficulties, so is worth considering separately here. In general, references in text should indicate author's name and year only; full references should be compiled at the end of the paper (not provided as footnotes).

There are many formats adopted by journals – an almost unlimited combination of sequences and syntax – and the writer has little choice but to follow house styles. A critical requirement is to obtain all necessary information when initially capturing references; this may include both volume and *issue* for journals, and *city* (not country) of publication for books and reports. It is extremely painful and time wasting to have to chase up the original sources because incomplete information was recorded initially. A recent personal experience of this was when a journal

[1] For example, in economics the JEL Classification Codes are often included.

insisted on stating the number of pages in books and reports referenced.

In the case of personal communications, it is desirable to provide information about the informants claim to expertise – designation, employer, location – in the references section. Anagrams are often preferable for text references, especially for agencies with long convoluted names – provided full details are provided in the reference section.[2]

The use of indirect references – of the form 'Brown (1988, cited in Bloggs, 2001)' – is to be avoided if at all possible, since it indicates a lack of effort to chase down the original source.

While 'Anon' was commonly used in the past for unknown authors, this tends not to be acceptable now. Sometimes the name of the agency is cited, if there is no information about the particular persons who did the writing.

Increasingly, material from the Web is cited in research papers. Reviewers of submitted papers tend to be rather critical of the Web as an information source, since much of the material posted has undergone little or no validation (peer review). It is desirable to indicate the full http address, the agency to which it applies where possible, as well as the date when the site was accessed.

For convenience, references may be kept in a reference database, such as Endnote. This allows them to be conveniently grouped under topics, and extracted in a variety of formats, to suit the style of the publication medium.[3]

4. FURTHER WRITING HINTS

While choice of words is a personal thing, some guidelines can be laid down on what presentation is likely to be well received by editors and reviewers. While some authors like to use extravagant language, in general a guiding principle is to keep the text as simple, straightforward and uncluttered as possible, both visually and in terms of the message conveyed.

There are a number of words commonly used in conversation which are to be avoided or minimized in written material. These include:

* vague terms, e.g. 'very' ('highly' may be a better choice), 'fairly', 'quite', 'different' (as an adjective – 'various' is often a better choice).

* terms that have well accepted meanings which differ from the context in which they are being used, and are sometimes used excessively in written material, e.g. 'good' (means a product, or holy), 'certain' (can often be replaced by 'particular' or 'specific') 'significant' (a term which has a specific meaning in statistics relating to probability levels in hypothesis tests – 'considerable' or 'major' may be better choices).

* terms which are misused in everyday language and can be irritating to the reader, e.g. 'cheap' (not an economic term, and often taken to mean 'of low quality'), 'quality' (used far too often and without any qualification of 'high' or 'low' quality)

* uninformative terms, such as 'results' (it is usually possible to indicate what type of results are being provided)

* personal pronouns, such as 'we' and 'you'. These can be used to effect for a reader-friendly ambience in a textbook, but in a journal article are rather too informal.

* inappropriate use of adjectives, such as in the expression 'cold temperature', 'cheap prices' and 'early age'. Here the noun is a number, not a characteristic, e.g. air temperature is a number between –70 and +50 approximately, in celsius), but the adjective is designed to describe a characteristic. In each case here, 'low' would be a better adjective.

[2] In Australia, the Department of Employment, Education, Training and Youth Affairs – which was responsible for the points system for grading university publications – was much more simply referred to in text as DEETYA. Like many public sector agencies, the name has been changed at least once since this mouthful.

[3] This can also save time chasing up a reference from another document on the computer hard drive, e.g. using Windows Explorer – Tools – Search – Containing text.

A construction to be avoided is to commence a new paragraph with an expression designed to link sentences in the same paragraph, such as ‘Therefore’, ‘Hence’, ‘As a result’ and ‘However’.

The wording of section headings should be sufficient to convey to the reader what the section covers, i.e. should be relatively ‘stand-alone’ or convey context. It is better to have long headings (even running on to a second line) than overly brief (including one-word) headings. It is undesirable to have two headings running; normally, there should be some substantive or at least signposting text between headings.

References are normally cited in past tense, since the author may no longer have the same view, or may not even be alive. It would not make much sense to say ‘Faustmann (1844) argues that the optimal rotation is …’.[4]

Graphs and tables require particular attention. It is important to lable the axes or columns with the name of the variable (not just the units, e.g. ‘(%)’. All columns of a table, including the first, should have a caption. A table must have a title, and be situated in document at a point after which it has been mentioned in the text. (Some journals require tables and figures to be grouped at the end of a paper when it is submitted.) A table should have a minimum number of horizontal lines – usually at the top and bottom of the row containing the column captions and at the bottom of the last row of the table. Vertical lines are to be avoided if at all possible. Following these guidelines will result in an uncluttered appearance.

A number of style guides are available, which provide advice on both writing style and syntax, e.g. the Cambridge guide and the AusInfo guide. While these are recommended reference works, it must be borne in mind that each journal has its own in-house style. Also, the style of writing tends to vary between disciplines. One interesting example is that in the sciences it is a guideline to include descriptive material about tables in the title, whereas in the social sciences the principle is to keep table titles as concise as possible and confine all explanatory material to the text or a table footnote.

[4] Here ‘argues’ should be changed to past tense, but ‘is’ should remain in present tense.

5. FORMS OF PUBLICATIONS

There are various forms of papers, and to some these forms can be stages in the production of a journal article.

Conference papers

This is often the first form of publication of research or review findings. A commitment to a conference presentation provides a deadline which can increase a researcher’s focus – one needs to prepare something useful to say. A conference is a made-to-order feedback mechanism, though often the most useful comments will be obtained outside the formal sessions. Audience members may draw attention to other related research. Often the proceedings will be published, in which case these are claimable publications. Sometimes conferences result in an invitation to submit a paper to a journal.

In-house discussion papers and monographs

These are sometimes used as a first run, or as a more durable form for an unpublished conference paper, to establish claim in the field, publicise research and obtain feedback. These have higher status if subjected to peer review process.

Journal articles

In general, the greatest publication credit is obtained for having papers published in refereed journals. This can be a time consuming and exacting task. Choice of an appropriate journal is critical. Ideally, articles will be submitted to ‘double blind review’ by two or more reviewers, i.e. where the authors does not know who the reviewers are, and the reviewers do not know who the authors are. Sometimes single blind review, where the author’s names are disclosed to the reviewers, is appropriate; the reviewer’s then know ‘where the authors are coming from’ with their views and can take this into account. Sometimes papers rejected in one journal

will be accepted by another.

Book chapters

These are another well respected form of publication, though often ranked not quite as highly as journal articles. The review process for books typically is not as exacting as for journals.

Authored and edited books

Authored books tend to be regarded more highly than collections of chapters individually authored by contributors. Also, books with a mainly research focus tend to be regarded more highly than textbooks. Notably, the weight given to books depends to some extent on the idiosyncrasies of the appraiser. For instance some people say that nothing new ever appears in text books, and that the only worthwhile publications are those presenting original research in journal articles. Others regard books with publishers of high standing as important publications.

Consultancy reports

To some extent, taking on consultancies is competitive with undertaking research projects. In an academic environment, consultancy reports typically are not considered as 'publications' in terms of the narrow criteria which bring academic recognition, funding and promotion. However, taking part in consultancy projects can be invaluable experience, can provide an important service to industry or community, can lead to development of new professional contacts, and can lead to subsequent production of papers for publication. Hence it appears desirable to undertake some consultancy work, but not at such a high level as to prevent other more traditional university research.

Extension papers

These can perform a useful technology transfer and service function. Unfortunately, in an academic environment, they are given little if any credit.

6. GETTING ITEMS PUBLISHED

Even when a researcher writes highly original and interesting papers, publication is by no means assured, and there is a good deal of strategy involved in having papers accepted for publication. Choice of journal is important, since research quality is assessed by quality of the journal in which it is published and it is therefore necessary to aim for as journal of as high standing as possible. Importance of high-standing, high impact, high citation rate journals. This does involve to some extent maintaining a clear disciplinary focus in one's research program. While it is critical for junior faculty staff to build up the length of their publication record, the importance of publishing in high-standing, high impact, high citation rate journals cannot be overemphasised. At the same time, the journal should be clearly within the researcher's discipline area. In establishing a research reputation, it is necessary to be seen as a specialist in a particular area, and not a generalist publishing over a wide spectrum of outlets – 'the jack of all trades is a master of none'! Further comments on choice of journal are provided in later sections of this module.

Sometimes it is difficult to know if a paper one has written is publishable. In this situation, it is helpful if an opinion can be obtained from an experienced publisher, such as a departmental mentor. Often the write will not recognize the publication potential of their work.

Sometimes a publication outlet can be found for concept papers. This means for example that a project proposal can be worked up into a publication. Obviously, a lot of research often goes into a research application, so this is not an unreasonable outcome.

There can be advantages in joint authorship. To the extent that length of publication list is important, it can be beneficial to have a large number of jointly authored papers, rather than a small number of sole-authored papers. Also, the combined efforts of two or more researchers can result in stronger papers and hence acceptance in better journals.

When seeking to have books published, it is usually necessary to prepare a book proposal for a publishing house. Most

publishers can provide a set of instructions on the kinds of information they require to make a judgment about accepting a book. This information will include details on the authors, competing books, special features, amount if international material, the targeted market and potential sales, expected length, extent of graphics, potential reviewers, possibility for translation into other languages, and various other things. It is usually necessary to submit a few sample chapters, though some publishers will want to see a draft of the overall manuscript before making a commitment. The information sought by the publishers is usually designed to give them an indication of the market size and production cost. In some cases, the publisher will have a preference for being provided with camera-ready copy; this reduces their production costs, hence allowing smaller print runs, and greatly speeds up the publication process.

The manuscript will sometimes be sent for expert review. Almost invariable, a copy editor will make style corrections. Once the page proofs of the text are finalised, the authors will need to prepare an index. This usually involves preparing a list of items, sub-items and page numbers, most easily done on a computer and using an alphabetical sort. Modern wordprocessor packages have the capability to form an index, through selecting appropriate words in the document files, though it is not clear that this actually saves time in preparation of an index.

Nowadays, Powerpoint slides are widely used for conference, symposium and workshop presentations. While these are occasionally published, it should be kept in mind that Powerpoint slides have little information content except in the presence of the author providing a commentary on them. If one has prepared a Powerpoint display, but not written a formal paper, it may be possible to capture the essence of the presentation by use of a cassette (or microcassette) recorder.

It is interesting to observe habits of highly successful researchers. One person with a notably impressive publication record devotes a block of time each day to writing (typically 5am to 10 am, before going to the office), spends a good deal of time contemplating and conceptualising, keeps collections of notes and ideas on various papers under development, does not use a computer (but has research typing support), writes relatively short papers, produces a large number of discussion papers which are later developed into journal articles, and is able to bring papers to a publishable state with remarkably few drafts. Unfortunately, some of these habits are not particularly feasible for most of us, but there are certainly lessons to be learned from successful researchers.

A useful principle to keep in mind is to become a habitual writer, e.g. type up field notes and summaries on discussions, script lectures, write discussion papers, commit to conferences. Writing becomes easier, and confidence is gained about writing ability, with practice.

7. CHOOSING THE RIGHT JOURNAL

Choosing the appropriate journal to publish an article in is one of the most important decisions that are made in the publication process. Not all journals are equal in terms of quality and prestige. The next section deals with two the key issues associated with selecting a journal, i.e. (1) matching the type of article to the publication outlet, and (2) determining the quality of a journal.

Matching the article to the journal

Journals are selective in what they publish. They usually publish material on a particular discipline area such as forestry, environmental management, or nuclear physics. It seems almost self-evident that it is important to select a journal that publishes material on the topic of your paper. The general discipline area of a journal is usually evident from its title. For instance, it is obvious that the *Journal of Forestry* publishes papers in the general area of forestry. What may not be clear is the specific focus of a journal. For instance, the discipline of forestry covers areas such as silviculture, menstruation, forest economics, extension and ecology, and there are many forestry journals in print which sever particular specializations in the discipline. Some journals only publish papers in quite specific specialty areas

within a discipline. Particular journals also often favour particular styles of articles. Some journals publish only highly quantitative papers which involve large data sets obtained from rigorous experiments. Other journals publish material that is qualitative in nature or more policy orientated, while other journals publish mainly review material. It does not matter how good a paper is, it is unlikely to be accepted by a journal if it is not consistent with the fields and style of material that the journal publishes.

It is critical to spend some time determining which journals are likely to be interested in the article that you are preparing. It is preferable to do this research at an early stage of preparing the paper. An attempt should be made to keep the style of the paper as consistent as possible with that used in the target journal. This reduces the likelihood of the paper being rejected on the basis of inconsistency with the style and nature of the journal.

The easiest way to identify whether a journal is appropriate is to read past issues and see if papers have been published in the field of the your paper and which use similar methodologies. Usually you will refer to a number of past studies in a manuscript and a sound guide to suitable journals is to look at what journals have been the sources of the key references of your paper. If still uncertain, it is advisable to contact the editor of a journal to see if they are interested in publishing the type of paper you are preparing.

As a general rule, try to publish in the best journal that will accept your work. Saying this, it is also important to be realistic about the quality of the work that you produce. Some pieces of research that you produce will be better than others. Some articles will be good enough to get into top international journals while other pieces of work might be more suited to regional journals. Comparing your work with articles published in the journal under consideration is probably the best way to whether it is of suitable quality. Important things to keep in mind include the size of your data set compared with other studies, the uniqueness of the work, thoroughness of analysis, the types of statistical techniques used, the level of theory development and extent of referencing to recent relevant literature. Top journals seldom publish research or review articles that have methodological flaws or lack a sound theoretical framework.

Usually the Chief Editor of a journal acts as the ‘gatekeeper’ and will not allow unsuitable papers to proceed to the review process, although this is not always the case. Publication can be delayed substantially by submitting a paper to the wrong type of journal, particularly if it goes into the review process before being deemed to be unsuitable.

It is almost inevitable that at some stage a paper one submits will be rejected. Do not feel discouraged – this happens to everyone – even the best of researchers. Aim to submit the article to the next journal on the list that you have identified. When a paper is rejected, the editor is likely to provide some feedback on what was considered unsatisfactory in it. The rejection may have been on the basis of the article not being consistent with fields and style of the journal. In other cases the paper may have gone to review and was rejected on the basis of reviewers’ comments. If this was the case then the editor is likely to forward the reviewers’ comments as well as some comments of their own. The comments of the editor and reviewers are often useful feedback and should be addressed as far a possible before resubmitting the paper to another journal. It is likely that the next reviewers will have some similar views to the previous reviewers. Before submission, it is important to rework the paper following the particular style requirements of the new target journal.

Quality vs quantity – which is more important?

The point has been made elsewhere that some avenues for publication are better than others. Generally, journal articles are considered to be better than book chapters, which are in turn considered to be much better that published conference and discussion papers. While this is true, it is also true that some journals are much better than others. It is also true that something is better than nothing.

The question of quantity versus quality is an important one and bears some discussion. Developing a record of publication is probably the most important aspect of developing a track record. Often when people are assessing a person from their resume or curriculum vitae (CV) – for example, for a job or in appraising a grant application – they will skim the initial information such as degrees and then flip straight to the list of publications of the applicant. It is thus critical to act strategically when submitting material for publication.

The relative importance of quality and quantity of publications differ depending on the stage of career. Assessors for grants and those on selection committees for jobs look for different things depending on the career stage of the applicant. For instance, an early career researcher or academic would not be expected to have 10 or 15 papers in leading journals. In reality, early career people often have no publications in top journals and assessors and selection panels understand this. What these panels often look for is potential. In these situations, the adage 'something is better than nothing' is highly appropriate. Panels look for 'indicators' or 'flags' that an applicant has potential, and actual publications (no matter what they are) are preferred to none. An applicant who has published several conference papers and maybe has a journal article in a third tier journal will be preferred to one who has no publications.

Getting journal articles into top journals can be a frustratingly long process. The review process for top journals is particularly rigorous and the research on which an article is based must be of high quality. The time between commencing research and having it published in a top journal can be six or seven years and sometimes longer. Such delays are less critical for established researchers because they invariably have a number of projects running simultaneously and which are at different stages of completion. For an early career researcher, presentation of preliminary research results at conferences is an effective way of obtaining publications.

In the case of early career researchers the balance between quality and quantity of publications is strongly in favour of quantity. As stated before – something is better than nothing. This does not mean that quality is not important – it simply reflects the fact that quality publications take time to produce and have published. It is important for early career researchers to produce some publications in the intervening period. A reasonable target is to produce at least two published papers a year, or more if they are multi-authored.[5]

Quality and type of publications become increasingly important with more senior positions, particularly in faculty positions. In the most senior positions, such as Associate Professor and Professor, quality is critical. Quantity is also important but more as a necessary minimum amount.

In Australia there is some degree of conflicting messages being sent in terms of quality versus quantity of publications. Promotions committees place a strong emphasis on quality of publications, particularly at the more senior levels. This conflicts with the funding formula for universities, which draws no distinction between the quality of refereed publications. The situation is different in the UK where much more emphasis is placed on the impact of research and this is how universities are assessed. Academics are assessed on the impact of their top articles, not on volume of publication.

8. ASSESSING THE QUALITY OF JOURNALS

Since journals vary in standing, it is worth spending some time exploring journal quality in more detail. Publishing in particular journals carries a substantial amount of prestige. How then does one gauge how 'good' a journal really is? Some suggestions are now made on how to determine the quality of journals, with specific comment on key forestry journals.

[5] The average number of published papers per person in schools of the University of Queensland is between one and two (based on a count of 1/n papers when there are n authors).

'Quick and dirty' guides

The simplest guide to what journals are best to publish in is to look at the journals cited in the references of one's paper. In any field there are often a number of key articles that are regularly cited – the journals in which these are published are usually the leading ones in the field.

A reasonable guide to the quality and impact of a journal is to look at who publishes it. Journals published by academic publishers such as Springer Verlag, Cambridge University Press, CABI, Blackwells, Urban and Fisher, and Elsevier are generally of a high standard and have a high wide readership and impact because of the distribution and marketing support of the publishing houses. Journals published and distributed by professional organizations and University departments are much more variable in quality and impact.

Another way of gauging what journals are highly regarded is to look at what other people in your field consider to be important. For instance, an excellent guide to what forestry journals are highly regarded in the USA can be obtained by looking at the profiles of US forestry faculty contained in the various web pages of leading forestry departments. The practice in the USA is for staff to list only their high quality or most significant publications in the lists. A quick perusal of these pages provides a sound indication about what journals are highly regarded. For instance, faculty commonly highlight *Forest Science*, *Forest Ecology and Management, Canadian Journal of Forest Research* and the *Journal of Forestry*. Publications in journals such as *Society and Natural Resources*, *Soil Science Society of America Journal* and *Forest Products Journal* are also commonly listed.

Journal citation and impact indices

One of the best – and certainly the most objective – measure of the quality and impact of journals is provided by citation and impact statistics. The ISI Journal Citation Reports are the leading source of information about citation rates and impact of journals. This is a web-based service which can be found at http://jcrweb.com//. Two databases are available, one for science journals and the other for social science journals. Within these databases, journals are further classified in subject categories. 'Forestry' is one of the discipline areas in the biological sciences. Within the social sciences list is 'Environmental studies'. Information on citation rates and impact of 29 journals is available within the forestry group (Table 1).

The journals can be sorted in order of total number of citations and impact factor. Figures 1 and 2 provide screenshots of the ISI JCR for forestry journals demonstrating sorting in total cites (i.e. citations) and impact factors. These rankings provide an effective guide of the top journals in a particular field and hence the best ones in which to publish. Importantly, reference can be made to the ranking of the journal to justify statements about its quality when preparing a CV or grant application.

Journals can also be sorted on the based of Impact Factor, Immediacy Index, Cited Half Life and Citing Half Life (Figure 3). Details of the basis of how these indices are derived is provided in Figure 4.

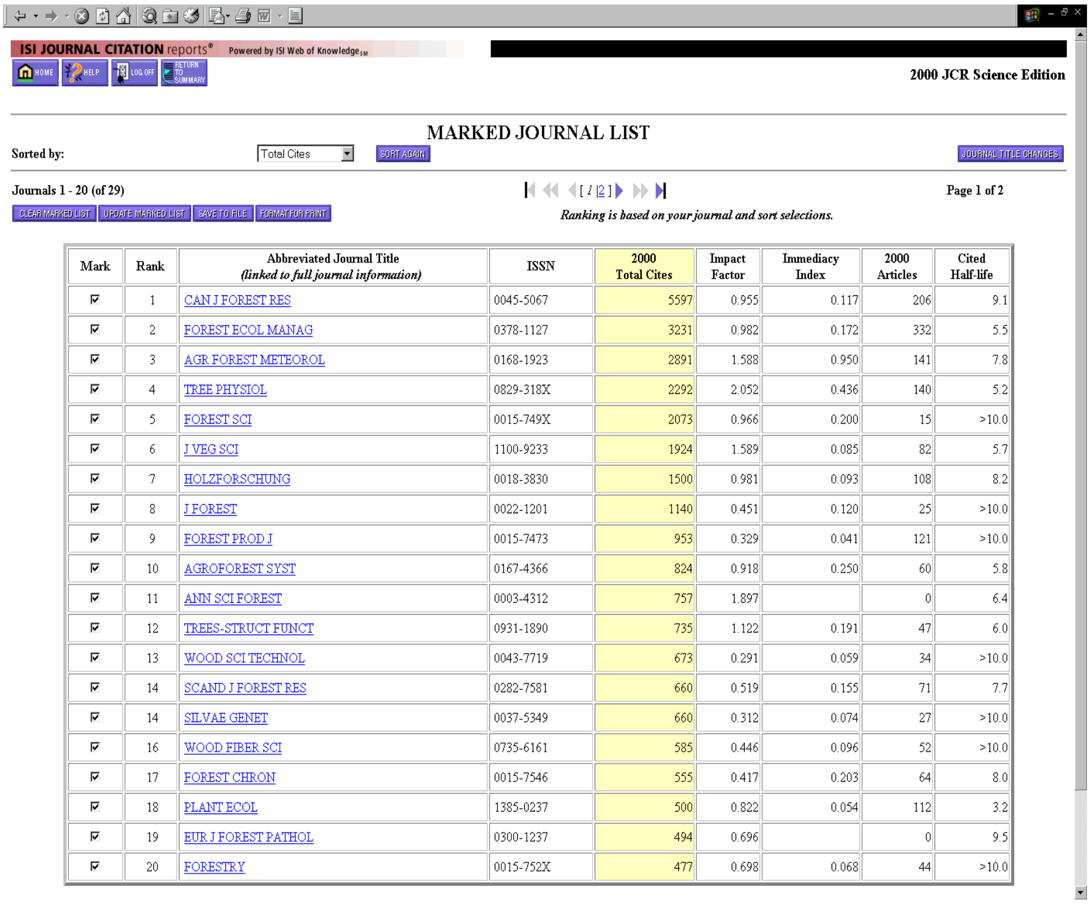
9. SEEING IT THROUGH

Bringing a paper to publication stage can be a long and frustrating process, and calls for considerable persistence. Some highly original thinkers and capable researchers move on to the next topic too soon, and don't do the 'hard yard' in report writing. Even when a paper is written and is suitable for a journal, this is not the end of the story. The 80%: 20% rule should be kept in mind: finishing off the last 20% of a paper production process takes 80% of the effort.

Notably, in publication success breeds success. When a researcher has a strong track record, invitations to contribute chapters in books and papers for regular and special issues of journals are likely to follow.

Table 1. Forestry journals listed in the ISI Journal Citation Reports

Abbreviated journal title	Year 2000 total cites	Impact factor	Immediacy index	Year 2000 articles	Cited half-life
AGR FOREST METEOROL	2891	1.588	0.950	141	7.8
AGROFOREST SYST	824	0.918	0.250	60	5.8
AI APPLICATIONS	87	0.500	99.999	0	
ALLG FORST JAGDZTG	141	0.239	0.031	32	6.9
ANN FOR SCI	53	0.576	0.203	69	
ANN SCI FOREST	757	1.897	99.999	0	6.4
CAN J FOREST RES	5597	0.955	0.117	206	9.1
EUR J FOREST PATHOL	494	0.696	99.999	0	9.5
FOREST CHRON	555	0.417	0.203	64	8.0
FOREST ECOL MANAG	3231	0.982	0.172	332	5.5
FOREST PATHOL	2		0.000	34	
FOREST PROD J	953	0.329	0.041	121	>10.0
FOREST SCI	2073	0.966	0.200	15	>10.0
FORESTRY	477	0.698	0.068	44	>10.0
FORSTWISS CENTRALBL	132	0.263	0.036	28	10.0
HOLZFORSCHUNG	1500	0.981	0.093	108	8.2
IAWA J	426	0.738	0.065	31	9.1
INT J WILDLAND FIRE	188	0.400	0.000	21	6.0
J FOREST	1140	0.451	0.120	25	>10.0
J VEG SCI	1924	1.589	0.085	82	5.7
NAT AREA J	330	0.452	0.095	42	6.8
NEW FOREST	184	0.417	0.027	37	5.6
PLANT ECOL	500	0.822	0.054	112	3.2
SCAND J FOREST RES	660	0.519	0.155	71	7.7
SILVAE GENET	660	0.312	0.074	27	>10.0
TREE PHYSIOL	2292	2.052	0.436	140	5.2
TREES-STRUCT FUNCT	735	1.122	0.191	47	6.0
WOOD FIBER SCI	585	0.446	0.096	52	>10.0
WOOD SCI TECHNOL	673	0.291	0.059	34	>10.0



ISI JOURNAL CITATION reports® Powered by ISI Web of Knowledge

HOME HELP LOG OFF RETURN TO SUMMARY

2000 JCR Science Edition

MARKED JOURNAL LIST

Sorted by: Total Cites SORT AGAIN JOURNAL TITLE CHANGES

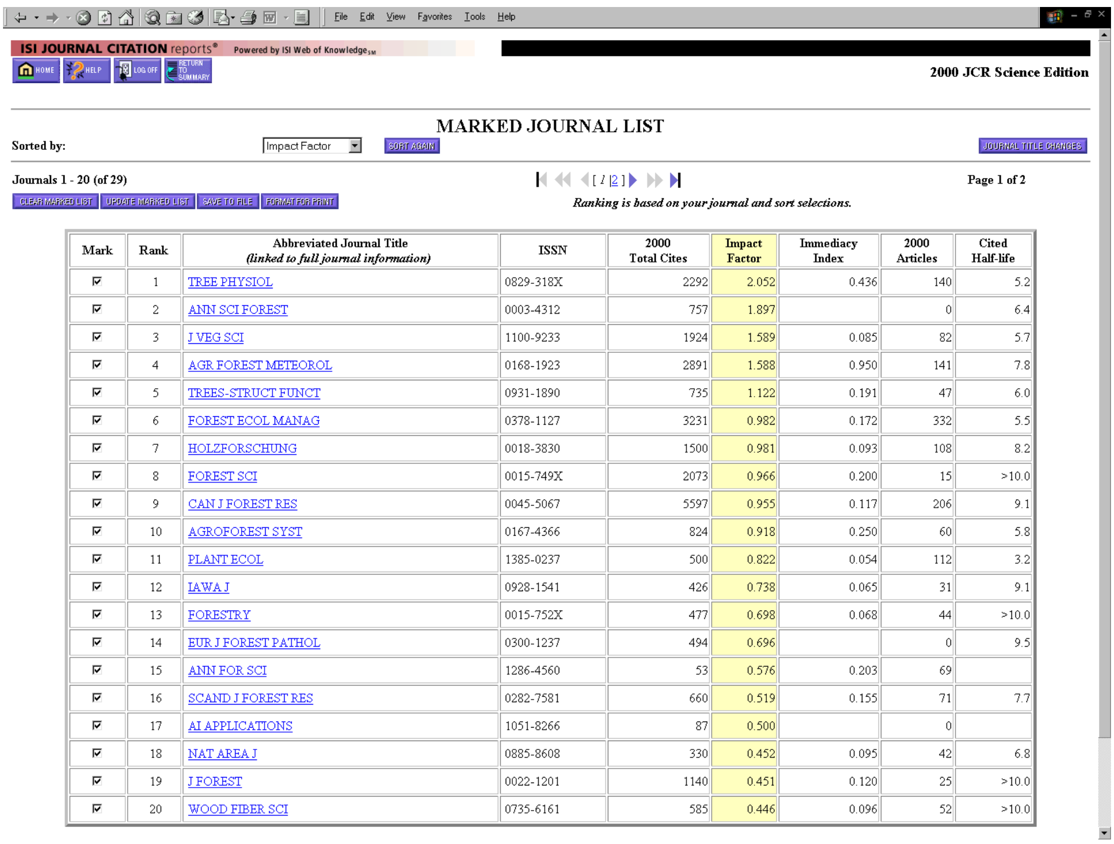
Journals 1 - 20 (of 29) [1 2] Page 1 of 2

CLEAR MARKED LIST UPDATE MARKED LIST SAVE TO FILE FORMAT FOR PRINT

Ranking is based on your journal and sort selections.

Mark	Rank	Abbreviated Journal Title *(linked to full journal information)*	ISSN	2000 Total Cites	Impact Factor	Immediacy Index	2000 Articles	Cited Half-life
☑	1	CAN J FOREST RES	0045-5067	5597	0.955	0.117	206	9.1
☑	2	FOREST ECOL MANAG	0378-1127	3231	0.982	0.172	332	5.5
☑	3	AGR FOREST METEOROL	0168-1923	2891	1.588	0.950	141	7.8
☑	4	TREE PHYSIOL	0829-318X	2292	2.052	0.436	140	5.2
☑	5	FOREST SCI	0015-749X	2073	0.966	0.200	15	>10.0
☑	6	J VEG SCI	1100-9233	1924	1.589	0.085	82	5.7
☑	7	HOLZFORSCHUNG	0018-3830	1500	0.981	0.093	108	8.2
☑	8	J FOREST	0022-1201	1140	0.451	0.120	25	>10.0
☑	9	FOREST PROD J	0015-7473	953	0.329	0.041	121	>10.0
☑	10	AGROFOREST SYST	0167-4366	824	0.918	0.250	60	5.8
☑	11	ANN SCI FOREST	0003-4312	757	1.897		0	6.4
☑	12	TREES-STRUCT FUNCT	0931-1890	735	1.122	0.191	47	6.0
☑	13	WOOD SCI TECHNOL	0043-7719	673	0.291	0.059	34	>10.0
☑	14	SCAND J FOREST RES	0282-7581	660	0.519	0.155	71	7.7
☑	14	SILVAE GENET	0037-5349	660	0.312	0.074	27	>10.0
☑	16	WOOD FIBER SCI	0735-6161	585	0.446	0.096	52	>10.0
☑	17	FOREST CHRON	0015-7546	555	0.417	0.203	64	8.0
☑	18	PLANT ECOL	1385-0237	500	0.822	0.054	112	3.2
☑	19	EUR J FOREST PATHOL	0300-1237	494	0.696		0	9.5
☑	20	FORESTRY	0015-752X	477	0.698	0.068	44	>10.0

Figure 1. Sorting on basis of total cites



File Edit View Favorites Tools Help

ISI JOURNAL CITATION reports® Powered by ISI Web of Knowledge

HOME HELP LOG OFF RETURN TO SUMMARY

2000 JCR Science Edition

MARKED JOURNAL LIST

Sorted by: Impact Factor SORT AGAIN JOURNAL TITLE CHANGES

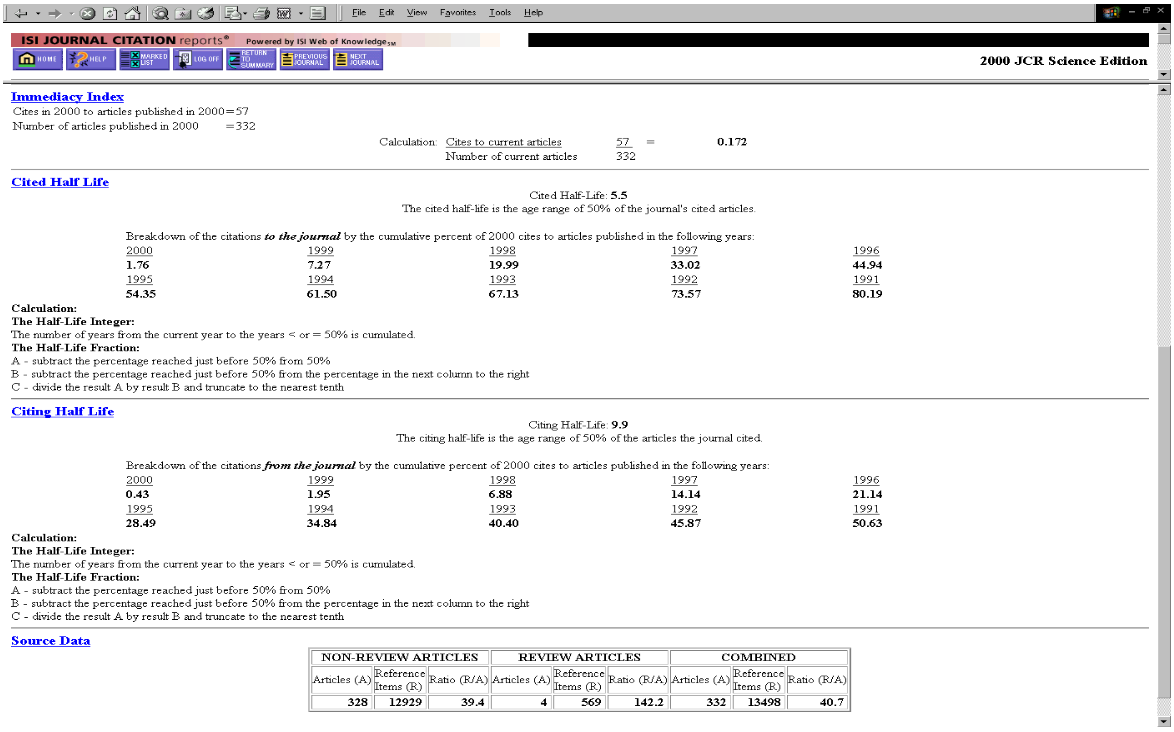
Journals 1 - 20 (of 29) [1 2] Page 1 of 2

CLEAR MARKED LIST UPDATE MARKED LIST SAVE TO FILE FORMAT FOR PRINT

Ranking is based on your journal and sort selections.

Mark	Rank	Abbreviated Journal Title *(linked to full journal information)*	ISSN	2000 Total Cites	Impact Factor	Immediacy Index	2000 Articles	Cited Half-life
☑	1	TREE PHYSIOL	0829-318X	2292	2.052	0.436	140	5.2
☑	2	ANN SCI FOREST	0003-4312	757	1.897		0	6.4
☑	3	J VEG SCI	1100-9233	1924	1.589	0.085	82	5.7
☑	4	AGR FOREST METEOROL	0168-1923	2891	1.588	0.950	141	7.8
☑	5	TREES-STRUCT FUNCT	0931-1890	735	1.122	0.191	47	6.0
☑	6	FOREST ECOL MANAG	0378-1127	3231	0.982	0.172	332	5.5
☑	7	HOLZFORSCHUNG	0018-3830	1500	0.981	0.093	108	8.2
☑	8	FOREST SCI	0015-749X	2073	0.966	0.200	15	>10.0
☑	9	CAN J FOREST RES	0045-5067	5597	0.955	0.117	206	9.1
☑	10	AGROFOREST SYST	0167-4366	824	0.918	0.250	60	5.8
☑	11	PLANT ECOL	1385-0237	500	0.822	0.054	112	3.2
☑	12	IAWA J	0928-1541	426	0.738	0.065	31	9.1
☑	13	FORESTRY	0015-752X	477	0.698	0.068	44	>10.0
☑	14	EUR J FOREST PATHOL	0300-1237	494	0.696		0	9.5
☑	15	ANN FOR SCI	1286-4560	53	0.576	0.203	69	
☑	16	SCAND J FOREST RES	0282-7581	660	0.519	0.155	71	7.7
☑	17	AI APPLICATIONS	1051-8266	87	0.500		0	
☑	18	NAT AREA J	0885-8608	330	0.452	0.095	42	6.8
☑	19	J FOREST	0022-1201	1140	0.451	0.120	25	>10.0
☑	20	WOOD FIBER SCI	0735-6161	585	0.446	0.096	52	>10.0

Figure 2. Sorting on the basis of impact factor



ISI JOURNAL CITATION reports® Powered by ISI Web of Knowledge

2000 JCR Science Edition

Immediacy Index

Cites in 2000 to articles published in 2000=57
Number of articles published in 2000 =332

Calculation: Cites to current articles / Number of current articles = 57 / 332 = **0.172**

Cited Half Life

Cited Half-Life: **5.5**
The cited half-life is the age range of 50% of the journal's cited articles.

Breakdown of the citations *to the journal* by the cumulative percent of 2000 cites to articles published in the following years:

2000	1999	1998	1997	1996
1.76	**7.27**	**19.99**	**33.02**	**44.94**
1995	1994	1993	1992	1991
54.35	**61.50**	**67.13**	**73.57**	**80.19**

Calculation:
The Half-Life Integer:
The number of years from the current year to the years < or = 50% is cumulated.
The Half-Life Fraction:
A - subtract the percentage reached just before 50% from 50%
B - subtract the percentage reached just before 50% from the percentage in the next column to the right
C - divide the result A by result B and truncate to the nearest tenth

Citing Half Life

Citing Half-Life: **9.9**
The citing half-life is the age range of 50% of the articles the journal cited.

Breakdown of the citations *from the journal* by the cumulative percent of 2000 cites to articles published in the following years:

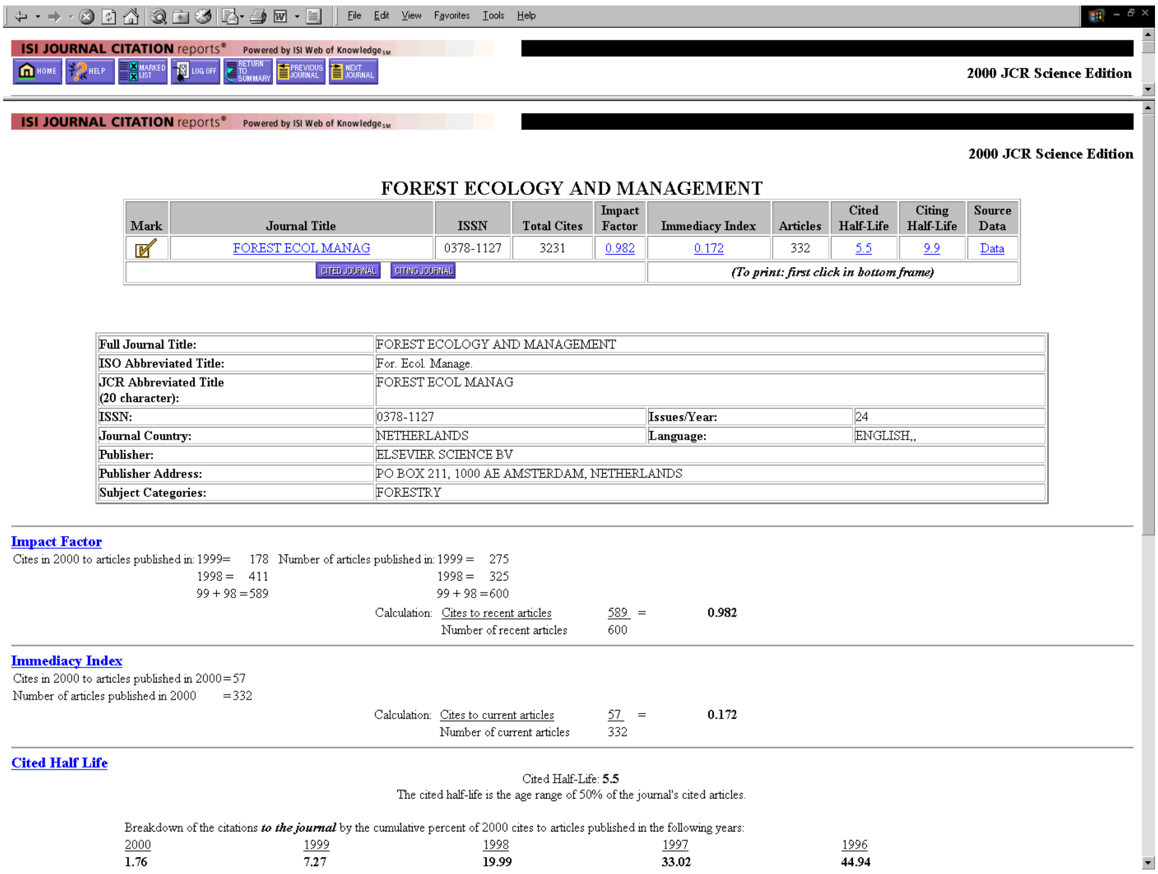
2000	1999	1998	1997	1996
0.43	**1.95**	**6.88**	**14.14**	**21.14**
1995	1994	1993	1992	1991
28.49	**34.84**	**40.40**	**45.87**	**50.63**

Calculation:
The Half-Life Integer:
The number of years from the current year to the years < or = 50% is cumulated.
The Half-Life Fraction:
A - subtract the percentage reached just before 50% from 50%
B - subtract the percentage reached just before 50% from the percentage in the next column to the right
C - divide the result A by result B and truncate to the nearest tenth

Source Data

NON-REVIEW ARTICLES			REVIEW ARTICLES			COMBINED		
Articles (A)	Reference Items (R)	Ratio (R/A)	Articles (A)	Reference Items (R)	Ratio (R/A)	Articles (A)	Reference Items (R)	Ratio (R/A)
328	**12929**	**39.4**	**4**	**569**	**142.2**	**332**	**13498**	**40.7**

Figure 3. Some citation details for an individual journal (Forest Ecology and Management)



ISI JOURNAL CITATION reports® Powered by ISI Web of Knowledge

2000 JCR Science Edition

ISI JOURNAL CITATION reports® Powered by ISI Web of Knowledge

2000 JCR Science Edition

FOREST ECOLOGY AND MANAGEMENT

Mark	Journal Title	ISSN	Total Cites	Impact Factor	Immediacy Index	Articles	Cited Half-Life	Citing Half-Life	Source Data
	FOREST ECOL MANAG	0378-1127	3231	0.982	0.172	332	5.5	9.9	Data
	CITED JOURNAL CITING JOURNAL				*(To print: first click in bottom frame)*				

Full Journal Title:	FOREST ECOLOGY AND MANAGEMENT		
ISO Abbreviated Title:	For. Ecol. Manage.		
JCR Abbreviated Title (20 character):	FOREST ECOL MANAG		
ISSN:	0378-1127	**Issues/Year:**	24
Journal Country:	NETHERLANDS	**Language:**	ENGLISH,,
Publisher:	ELSEVIER SCIENCE BV		
Publisher Address:	PO BOX 211, 1000 AE AMSTERDAM, NETHERLANDS		
Subject Categories:	FORESTRY		

Impact Factor

Cites in 2000 to articles published in: 1999= 178 Number of articles published in: 1999 = 275
1998 = 411 1998 = 325
99 + 98 =589 99 + 98 =600

Calculation: Cites to recent articles / Number of recent articles = 589 / 600 = **0.982**

Immediacy Index

Cites in 2000 to articles published in 2000=57
Number of articles published in 2000 =332

Calculation: Cites to current articles / Number of current articles = 57 / 332 = **0.172**

Cited Half Life

Cited Half-Life: **5.5**
The cited half-life is the age range of 50% of the journal's cited articles.

Breakdown of the citations *to the journal* by the cumulative percent of 2000 cites to articles published in the following years:

2000	1999	1998	1997	1996
1.76	**7.27**	**19.99**	**33.02**	**44.94**

Figure 4. The basis of calculation of some key citation inidices

10. SOME JOURNALS TO TARGET FOR FORESTRY-RELATED ARTICLES

The following journals are some of the ones which would seem appropriate for forestry research, particularly with a systems or socio-economics orientation.

Agricultural Systems
Australian Forestry
Annals of Tropical Research
Australian Journal of Environmental Management
Australian Journal of Agricultural and Resource Economics
Ecological Economics
Economic Analysis and Policy
Economic Geography
Forest Ecology and Management
Forest Policy and Economics
Forest Science
Journal of Environmental Management
Journal of Forest Products
Journal of Forestry
Journal of Sustainable Forestry
Natural Resources Forum
Small-scale Forest Economics, Management and Policy
Society and Natural Resources

11. SOME REFLECTIONS AS A JOURNAL EDITOR

A number of observations may be made from experience as a journal editor. The following comments arise from four years experience as editor of *Economic Analysis and Policy* (for the Economic Society of Australia Inc. – Queensland Branch). This journal, which has been published for more than 30 years, is issued twice yearly, with occasional special issues, individual issue typically containing about seven papers plus several book reviews.

The paper processing system

When submitting papers to a journal for publication, it is desirable to have a clear understanding of the paper processing system adopted by the journal. The following is a typical example of how this process works.

When papers are received, the editor will usually have a quick look at them, and perhaps seek an opinion of a colleague or member of the editorial committee. An editor can often judge quite quickly the suitability of a paper and the probability of favourable reviews. If they judge the paper is not suitable for publication, even with some further work, then the editor is likely to make an executive decision to reject the paper and advise the author accordingly. This could occur for example if the research reported lacks substance, the analysis appears weak, or the literature review is absent of weak.

In other cases, the editor may correspond with the author and say the paper has potential but is not suitable to send to review in its current form, e.g. it may be too long, too unsatisfactory in format (no Abstract, inadequate Introduction), or containing too much algebra for the journal audience. If the editor is generous with their time, they may provide considerable feedback at this stage, even to the extent of providing various suggested text changes.

If these hurdles are passed, the editor or their committee will identify suitable reviewers, and ask them if they are prepared to review the paper for the journal, typically sending the full paper or the abstract to assist them to make a decision.

In due course, the reviewer's comments are received. These may be to reject the paper, or that the paper is suitable for publication provided specified changes are made. The review comments are usually forwarded unattributed to the authors. If the reviews are mixed, the editor must make a decision about whether to proceed with the paper. If the decision is to proceed, then the editor will provide an indication to the authors of what changes they think are needed for the paper to be satisfactory.

Once revisions are made to the paper, the editor may decide themself whether these changes are adequate, or may send the revised paper to the reviewers and ask them if they are satisfied with the changes. In the latter case, if a reviewer is still unhappy with the paper, the process may be prolonged.

Hopefully, the paper is eventually judged to be acceptable, at which stage the editor may require wording changes, more

complete details on references. They may also request camera ready copy of complex graphics or computer screen images, and may discourage coloured diagrams because they can lose definition in black and white. The paper eventually goes to the typesetter, after which page proofs are produced and forwarded to the authors. Once all authors have corrected their page proofs, the typesetter makes final corrections and the issue is printed.

As indicated by these notes, paper processing can be a long and tortuous process.

Submitting papers

Journals typically provide instructions to authors of how papers are to be formatted. Often about four copies are to be provided, which must follow quite detailed formatting instructions. Presumably, the copies are date stamped on receipt, one copy is retained by the editor, one is filed away, and two are posted to reviewers. This is a rather archaic system, and nowadays it is generally more convenient if a single copy is submitted as an email attachment, not necessarily is a precise format, and copies are emailed to reviewers. The formatting according to in-house style can be then be done if and when the article is accepted. It is worth asking the editor if this method of submission is acceptable.

Contacting the editor

Some authors are loath to contact the editor after a paper is submitted. (Others phone them repeatedly.) In general, if there has been no response from the editor (other than an acknowledgment of receipt of a paper) after a few months, it is a good idea to contact the editor and inquire about progress of the paper. This of course will depend on the journal. Some have long turnaround times, but the authors are rather precious and will not respond well to contact. But generally, a follow-up is a good idea, the check that the editor has not lost the paper, or (more likely) failed to chase up slack reviewers.

Responding to reviewer's comments

When the editor indicates a paper will be considered further subject to making specified revisions, it is critical for the authors not only to make these revisions or provide good reasons for not doing so (the general rule is that 'the reviewers are always correct'), but also to provide a list of the changes they have made.

Handling page proofs

Page proofs may be provided as hard copy or in electronic form. At this stage only minimal and essential changes are acceptable. Generally, the author should forward had copy with clearly marked changes, although it the changes are minor it may be acceptable to advise by email.

What not to do

There are a number of things authors can do which will upset editors. In that it is wize to develop a friendly relationship with an editor, with a view to submitting further papers, these need to be kept in mind.

* Simultaneous submission of papers to more than one journal is unacceptable behaviour. There is a lot of work for editors and reviewers in processing a paper, and if the paper is withdrawn from one journal to be published elsewhere or they become aware that it is being considered by another journal, then your name is likely to be remembered should you submit to that journal again.

* Submitting a paper that is substantially the same as has been published elsewhere (by the contributor or by someone else) is also unacceptable. A 'repeat' paper may be detected by the reviewers, in which case the editor will probably say cautiously that the paper cannot be accepted because it is too similar to something that is published elsewhere. An even worse outcome (for the contributor and journal) would be for the paper to be published and then for someone to submit a comment paper pointing out the lack of originality.

* Submitting a paper which contains many views or subjective statements but lacks references can come across as presenting the author's personal philosophy of the world. Unless the author is quite brilliant, it is probable that the ideas have already

been thought of by others, and that a literature exists on them.

* Other things which will upset editors include: failure to provide a rejoinder explaining what changes have been made in response to reviewers' comments; failure to shorten a paper when the direction is that it must be cut down in length; failure to provide promptly camera-ready graphics when this is requested; long turn-around time with page proofs.

Length of papers

How long should the paper be? This will vary between disciplines. In the physical sciences, papers as short as about 2000 words are often published. Longer papers are acceptable in the social sciences, though manuscripts longer than about 7000 words approximately are generally not welcomed by editors. They take up too much space in the journal, such that one long paper may displace two shorter ones. They test the endurance of readers. They may invoke the response that the editor would be prepared to consider a shortened version. A length of 5000-6000 words is probably more optimal. Often, journal instructions will indicate, after a statement about journal articles, that 'shorter contributions are welcome'. Sometimes there is a separate section in journals for shorter comments.

Some further observations

A few further observations may be made concerning dealing with journals:

* Bear in mind that the editor is always looking for suitable papers to publish. With increased advertising of journals, including advertising on web sites, the number of papers submitted is increasing (increasing demand for publication), but there have been various new journals launched in the last few years, so the supply of journal outlets is increasing.

* In general, papers should not contain a large amount of graphics. These make a paper look too much like a slide show. Tables provide more information than graphs, because with tables numbers are known exactly.

* Don't wait for the editor to get back to you, if the time delay is too great. They may have lost the paper, or not got around to chasing up a slack reviewer.

* Choice of reviewers can make or break a paper. The author is unlikely to have any say over choice of reviewers, but in some circumstances can make suggestions. In the case of invited papers, or converting workshop papers into journal articles, sometimes a deliberate choice if reviewers is made such that they will not be 'unfriendly'.

* The focus or editorial policy of the journal must be kept in mind. Some papers no matter how good simply are not suitable for the journal. 'Swiggle papers' (highly mathematical, with lots of Greek characters) should be sent to 'swiggle' journals. Sometimes an editor will say it is necessary to 'Move the maths to an appendix, and expand the policy section'.

* It can be difficult to obtain acceptances for 'review' articles, as distinct from results of original research.

* In general, when a paper presents results of a survey, some hypotheses and statistical analysis would be expected, together with some test statistics and diagnostics to justify assumptions.

* Journals are generally short of book reviews. Unsolicited book reviews are usually welcome.

12. CONCLUDING COMMENTS

The publication record is invariably regarded as a key criterion of ability of researchers, including those working within universities. Different publication strategies will work best different people, and it is necessary to think strategically about having research findings published. To some extent, success is a matter of persistence, and skills can only be developed by experience. The observations and suggestions presented here are designed to stimulate further thought about publication strategies.

20. Experiences Gained and Lessons Learnt from the Training Workshop

Steve Harrison, Ed Mangaoang and John Herbohn

This review module attempts to draw together some of the experiences gained and lessons learned from conducting the training workshop on socio-economic research methods in forestry research. A workshop of this nature had been an ambition of The University of Queensland forestry research group for more than two years, so it was gratifying to see this ambition fulfilled. As a first run by the research group at such a training workshop, this was something of an experiment, and impressions gained here may be of assistance should the 'training package' be employed in other situations. This paper reflects on the workshop experience and on the feedback obtained from delegates and presenters. No formal survey was undertaken to gain impressions of participants, but considerable informal feedback was obtained.

---

1. OBJECTIVES OF THE WORKSHOP

In designing the workshop, a number of objectives were in mind:

- The workshop would fulful part of the capacity-building objective of ACIAR project ASEM/2000/088 – Redevelopment of a Timber Industry Following Extensive Land Clearing. In particular, there would be development of new skills by participants in socio-economic analysis and in modelling forestry systems.

- This would be an opportunity to communicate the forestry research mindset of the Australian team to the Filipino team, and obtain feedback, and hence to explore interactively the paradigms, methods of analysis and assumptions upon which each group is working.

- The planning of the ACIAR project would be advanced, particularly through the research 'case studies' which were tightly tied to objectives of the project. This would include development of a prototype model for simulation of community forestry systems.

- The workshop would establish the basis for improved communication and closer understanding between the researchers throughout the remainder of the ACIAR project.

- A further objective in planning the workshop was to provide a stimulating and enjoyable experience for participants.

- A planned tangible output was this training manual, to be supplied subsequently to all participants.

2. PLANNED WORKSHOP ARRANGEMENTS

Formal presentations were designed on a number of specific topics, but with a view to group participation and discussion. More specifically, the workshop meetings were designed to include the following components:

1. Presentations by authors of the notes
2. Presentations by discussants
3. Small-group development and presentation of case studies
4. A fieldwork component
5. 'Hands-on' computing sessions.

Ideally, the instruction papers would have been prepared and distributed to delegates in advance of the workshop. This would have allowed formal sessions to concentrate on specific issues, and provided a basis for contributions by discussants. Nominated discussants could have been invited to prepare brief

presentations on the basis of materials provided to them in advance. In particular, they could have lead discussion on the relevance of the particular concepts and methods in forestry research, from their organizational perspective. However, in general this was not possible, and a folder containing some of the printed modules was provided to participants at the time of registration, with other modules provided during the workshop.

An objective of the workshop was to have each participant prepare a small simulation model using the Simile package, and most were able to achieve this.

The first three days of the workshop covered the main presentation sessions, with the emphasis for the next three days on case studies. The field trip was scheduled to provide a change from the formal sessions, and to collect data for subsequent used in development of the prototype community forestry model.

3. LOCAL ARRANGEMENTS

The organisation on the part of staff of Leyte State University was excellent, in terms of meeting facilities, copying of materials, transport and meals. This kind of meeting always imposes a heavy burden on the local organisers, and their efforts were greatly appreciated. A number of social events were organised, which made the occasion memorable and built morale and cohesion in the group.

4. REVIEW OF TOPICS COVERED

Five major topic areas were covered, namely:

1. Basic computing skills, using spreadsheets, discounted cash flow analysis and statistical analysis software.

2. Specific research techniques (sample surveys, cost-benefit analysis, non-market valuation, linear programming).

3. Modelling and simulation of community forestry systems with Simile.

4. Grant seeking, research administration and publication strategy.

5. Practical exercise in developing a research project.

5. SUITABILITY OF THE CONTENT AND ENGAGEMENT OF PARTICIPANTS IN DISCUSSIONS

As a first run, it was to be expected that there would be some glitches. An ambitious agenda was planned for the workshop, and though much material for presentations had been drafted before the meeting, continued module development took place between workshop sessions.

The choice of topics of necessity depended on the interests and skills of the group of presenters. In this regard, it was fortunate that the visiting group covered a wide range of interests, and that senior staff from Leyte State University were also able to make presentations.

The program proved to be highly ambitious, in scope, complexity and amount of material presented. However, flexibility was provided in the schedule, between presentations of new materials, discussion of case studies and computing sessions. In general, days ran from 8 am to 6.30 pm. Despite this pace, by all impressions, delegates found the workshop absorbing, sometimes continuing computing through coffee breaks.

It was hoped prior to the meeting that the presentations would simply be lead-ins to topics, and that there would be considerable interaction and debate at workshop sessions. To some extent, this ambition was achieved. However, it is to be noted that the material presented was rather foreign to most of the participants, and this at times constrained the discussion. Although the participants had a background in natural resource management, and were engaged to some degree in policy and extension work, most had little background in economics or social research.

No doubt, at times some participants felt rather lost with the topics, and it will require

further reflection to gain a working understanding of the techniques presented. Were a similar training workshop run in the future, with the materials which are now available, it would be possible to provide notes for reading in advance and hence make greater use of meeting time.

6. LEVEL OF COMPLEXITY AND BACKGROUND KNOWLEDGE AND SKILLS ASSUMED

One difficulty in holding a workshop for a group of people from varying academic, industry, local government, NGO and other backgrounds is that the level of familiarity with the presentation topics and the level of expertise with computers can be expected to vary widely. In that the training workshop was designed to broaden the skills base and interests of the participants, and was outside the area of professional experience for most, some of the concepts in economics and sociology were found to be new and strange.

The modelling package Simile takes considerable effort to comprehend. Becoming adept at a complex computer modelling package such as this by nature takes a good deal of time, for assimilation, experimentation and consolidation, to gain confidence as a user. The workshop provided an ideal introduction to Simile, but substantial further time will be required before participants can start to develop and run realistic models. An added complication here is the need to become familiar with the concepts and procedures of modelling and simulation, which can be a hurdle for people without a background in this area.

It was found during the hands-on sessions that most of the delegates did not have the level of computing experience expected of the presenters, particularly in regard to the use of the Microsoft Office package and the Excel spreadsheet. While a desktop computer and this software is regarded as a standard facility at The University of Queensland, the high cost of computers relative to university budgets can make access more limited in lower-income countries. In the College of Forestry at Leyte State University, the stage has not been reached where each faculty member can have exclusive access to a desktop or notebook computer, and so there is less familiarity with Excel, which was applied widely in the modules for financial modelling (including discounted cash flow analysis, cost-benefit analysis and linear programming). Hence it was necessary to provide some background coaching in use of Excel.

7. WORKSHOP OUTCOMES

Areas found to be of high interest

There was clear enthusiasm for the modeling activities and hands-on computing in relation to the simulation modeling package Simile, development of financial analysis spreadsheets with MicroSoft Excel and linear programming using the Solver facility. The modules on research administration and publication strategy also appeared to be of high interest.

Benefits derived by delegates

In general, the workshop objectives as listed above were achieved to a high degree. As well as the more formal training aspects of the workshop, some practical assistance in computing was provided to the delegates, e.g.

- providing practical tips in computing, particularly with regard to use of the Excel spreadsheet package
- providing access to the add-ins in the Excel spreadsheet package, such as statistical techniques and the linear programming package (Tools – Solver).
- providing copies of the Simile modelling and simulation package developed by the University of Edinburgh, which was downloaded from the developer’s Web site.

Development of the prototype community forestry model

An important objective of the workshop was to make some steps towards development of a model of community forestry, which could be used to simulate forest industry development and carry out simulation experiments with alternative reforestation

policies. It was recognized that only limited progress could be made during the short period of the workshop and with the competing objectives planned. In particular, a field visit – unless organised as a participatory rural appraisal meeting – could not be expected to provide a great deal of detailed information for modelling purposes, although it did provide an excellent chance to gain first hand observations of a community forestry site.

Suitability of Organisational Arrangements

Some difficulty was found in providing sufficient access to computers during the hands-on computing sessions. Three notebook computers were brought to the workshop, and another notebook computer and one desktop computer were available locally, which meant that groups of about five people were sharing computers. Groups of three would have probably been preferable.

It was noted that local delegates were unable to attend all sessions. This was to a large extent due to their involvement in the preparation of materials and other arrangements for the workshop, and partly because of other commitments. This suggests that there could be advantages in holding workshops of this kind away from the normal place of employment of delegates, such that they become a 'captive audience', though this would of course require an out-of-semester time window, and would tend to increase costs in terms of travel and accommodation.

8. CONCLUDING COMMENTS

In general, the workshop appears to have been highly successful. While some tentative recommendations might be made for future training programs of this type – e.g. make available more computers to allow smaller user groups, gain a better idea of the computer access and computing experience of participants – it would seem desirable to retain flexibility to meet the needs of the audience and adapt presentations as suits the progress of the group.

It was to be expected that some problems would arise, but overall the workshop appears to have been highly successful in achieving its objectives. Lasting benefits are expected for the ACIAR research project in terms of a shared research methodology and vision, and progress in planning of specific activities. This engenders confidence for conducting such a training program as a capacity-building exercise in other locations.