Ethical issues facing future availability of forestry data: a discussion

Andrew Haywood[1,2,3]

[1]Department of Primary Resources, PO Box 100, Beecroft, NSW 2119, Australia
[2]Email: andrew.haywood@poyry.com
[3]Current address: Poÿry Forest Industry, 5/437 St Kilda Rd, Melbourne, Victoria 3000, Australia

Revised manuscript received 20 June 2006

Summary

The rapid growth of research and data management systems suggests a future where we will be able to manage forest systems with a greater understanding and accuracy than ever before. We are already developing forest management systems well beyond the traditional aspatial and textural formats. However, as data become more accessible, with associated increased resolution (both temporal and spatial), and easier access to various statistical tools, ethical issues that were not previously considered significant in the traditional management of forestry data are now emerging as very important. Data that previously may have been thought of as private or protected can now be easily copied or moved and combined with other data, resulting in new concerns about who uses the data and in what way. In addition, as data mining capabilities are increasingly built into supporting software, issues associated with privacy, accuracy and accessibility are becoming more important. It does not take great imagination to envisage a situation, in the not too distant future, where regulations and legal decisions will dictate certain security safeguards for the data management systems that handle such data. These developments are likely to restrict data availability and, as a consequence, hinder the evolution of forest information management systems. However, there are actions that forest data managers could take to alleviate potential future restrictions of data availability.

Keywords: data communication; resources; availability; management; ethics; decision making; models

Introduction

> Data is not information, information is not knowledge, knowledge is not understanding, understanding is not wisdom*.

The above quote clearly makes the point that just because we have acquired some data, it does not necessarily mean that we are informed. Similarly, being informed does not necessarily make us knowledgeable about a topic. Although this quotation is accurate, the essential process of converting forestry data into wisdom (management decisions) is greatly assisted if clear and explicit ethical standards are adopted in the collection and use of data. Without regard to ethics in the data management process, the current and future usefulness of the data will be put at risk, and this issue is the focus of this discussion paper.

Forest professionals need to be aware of ethical standards of conduct and how these may dictate both the development of data management systems and the use of data. Traditional approaches to forestry data management have focused more on technical issues, primarily the collection and efficient storage of data. Recent developments, including ease of data access and the manipulation of data, have increased focus on social and cultural factors, while increased emphasis on multi-disciplinary and multi-agency approaches has exposed the importance of adopting clear and explicit ethical standards in the collection, management and use of forestry data. It is essential that data management is approached in the context of individuals, organisational structures and the practices related to the use of data. In any data management system, there should be a focus not only on how individuals use information but also on how individuals misuse information.

The aim of this discussion paper is to highlight the growing importance of ethics in data collection, management and data use within the forestry profession. I provide a generic definition of ethics, and examine ethical issues pertaining to the specific issues of privacy, accuracy and accessibility. In conclusion, I suggest potential steps that forest data managers can take to mitigate some of the issues related to the future availability of forestry data.

Definition of ethics

Although it is probably impossible to provide a definition of 'ethics' that all people would agree to, a general working definition is required. I decided that the following would suffice for the purposes of this paper:

> a system of moral principles, by which human actions and proposals may be judged good or bad, or right or wrong. (Macquarie Dictionary 1998)

The ethical concerns associated with future data availability presented in this paper do not include the management of intellectual property rights, as this issue has been adequately discussed elsewhere. Nor do I consider intentionally malicious behaviour, including computer crime, software piracy, hacking, virus surveillance and deliberate invasions of privacy. This paper

*Internet sites attribute this quotation to Clifford Stoll and/or Gary Schubert, US-based IT academics. A forerunner is to be found in T.S. Eliot's *The Rock*, Part 1, 1934.

focuses on the more subtle and important impacts that data collection and use may have on individuals and organisations in the forestry sector and how these, if incorrectly handled, may affect the future availability of data.

Privacy

Improper access to personal information is the ethical issue that usually comes to mind when 'privacy' is mentioned. In this section, I examine privacy from the viewpoint of data amalgamation, location privacy and Internet technologies, and how each creates unique problems for data management and its long-term availability.

Data amalgamation

Any unauthorised access to information about an individual or their property may be thought of as an invasion of privacy. However, even authorised access may create privacy concerns when access to individual data sources is used to combine information (Mason 1986).

For example, one government agency may record landowners' names and land ownerships, while another may be empowered to store land records and forest valuations for tax reporting purposes. Separately, the databases are likely to be properly managed, but if a third party combines the records it may be possible for unauthorised parties to gain commercially-sensitive financial data about individual landowners. In the current technological climate, natural resource databases are rapidly increasing in size, complexity and connectivity. Consequently, when multi-disciplinary and multi-agency research projects involve adding data fields or combining data and information from a variety of sources, the ethical implications of those activities should be carefully considered.

Location privacy

In the last decade a relatively new privacy issue has emerged in the field of geographic information systems (GISs) and relates to location protection. For example, many indigenous Aboriginal cultural sites on public land in Australia are protected by law, policy or regulation. However, entering the site locations into an organisation's GIS may disclose locations for an unethical use. A similar situation can occur with respect to biodiversity and the protection of rare and endangered species within forestland. Consequently, innovative approaches are required by forest data managers to permit resource monitoring and protection while simultaneously ensuring no loss of privacy resulting from unauthorised disclosure of location. For example, the New South Wales National Parks and Wildlife Service have an approach whereby they supply consultants and other external parties their fauna and flora records only to the nearest kilometre (R. Kavanagh, Forests New South Wales, *pers. comm.* 2004).

Remote sensing is strongly associated with the development of GIS technologies and is becoming an increasingly accessible technology within the forestry profession. Current remote sensing capabilities in Australia allow almost anyone to 'peek into' someone else's property — potentially evaluating, without the owners knowledge or consent, the health or value of a forest. Even when the data are obtained legitimately, there is currently no guarantee in Australia that sufficient security safeguards have been installed to protect unauthorised access and use. It is envisaged that continuing research will create even greater opportunities for such unauthorised remote intrusions to occur.

Internet technologies

As more and more data become available through the Internet, how they are being used is becoming an important issue. Surfers of the web can easily 'obtain' data from a variety of sources and misuse or misinterpret them. This can cause significant problems for the original providers of the data, and can threaten the future availability of data. Similar concerns in relation to re-packaging of information from Internet technologies may arise when public funding of research (at both the federal and state level) places research information in the public domain. Enterprising commercial organisations have the opportunity to turn that public-domain information into company revenue. While not technically illegal, it may be unethical if there is little or no value added to the publicly available and funded information. As a consequence, the following incorrect view may be adopted — why should government be involved in providing public funding for such data? If adopted, this view would restrict the potential future availability of such data.

Accuracy

Accuracy is a broad topic and has many associated ethical issues. System construction, system inputs, error quantification, language and culture, and information presentation and filtering can all affect accuracy, each creating problems for management and long-term availability of data.

System construction

A system developer's ability to know, understand and predict all states (especially error states) is limited for the complex biological systems often found in forest management information systems. This leads to several ethical issues related to system accuracy. Well-known approaches are to document assumptions, and to develop appropriate test conditions and performance through system validation and verification. It would seem logical that an ethical developer would identify and correct all system errors. Unfortunately, for a variety of reasons, including time and/or resource constraints, developers often ignore or gloss over these ethical issues. This brings into question the long-term usefulness of the data and their storage, as often it becomes impossible to retrospectively correct errors in information. Disclaimers, system updates and patches, as well as a lack of substance in system warranties, arise from the developer's implicit acknowledgement of this issue. Similarly, ethical conflicts arise from the development of decision support tools, which end-users may believe to have a higher degree of accuracy than is actually the case. In many areas of forestry research and management, most model results represent only broad approximations. Consequently, the onus is on the developers to adhere to ethical standards in the collection and development of models, and to effectively communicate the limitations to the user.

System inputs

There are accuracy issues associated with deciding which specific information to use in a system. For example, it is often difficult to select appropriate socioeconomic and biological indicators or to decide which predictive model should be adopted. An indicator in this context is defined as:

> something that points to an outcome or condition and shows how the system is working in relation to that outcome or condition (Holling 1978).

For example, in a forest plantation stand evaluation model such as STANDPAK (Whiteside *et al.* 1997), stand top height is a key indicator of site quality. However, there may be a large set of potential equations available to predict stand top height (Vanclay 1994). One equation may simply predict stand top height from basal area and age; others may employ a different combination of variables. The equation selected will have different properties in terms of accuracy, precision, data costs and suitability for understanding and forecasting. The selection of an equation relates, in turn, to precision and bias in the estimators used. The requirements of the intended user should ultimately determine what specific information to use in forest research and management. Often indicators are chosen either because they have been used historically or are the only ones available, despite their limited value. Ethically, the researcher or practitioner should be concerned with ensuring the most appropriate indicators are chosen. In the author's experience, this seldom occurs in practice.

In a situation where a social or cultural indicator is being considered, ethical issues are likely to be even more important. If the indicator misrepresents a value set then it cannot be considered to be accurate. Indicators have a long history of use in forecasting systems for resource management (Holling 1978). Holling considered that such indicators must be relevant, understandable, dependable, reliable and timely. More recently, where there is increased emphasis on sustainability in natural resource disciplines, indicators should have additional properties. Indicators appropriate for sustainable systems should (i) include community carrying capacity, (ii) identify the links between economic, social and environmental *wellbeing*, (iii) be able to be used in the community, and (iv) focus on the long-range view. They must also measure local sustainability that is not at the expense of global sustainability (Hart 1999). It must be acknowledged that the usefulness of such indicators is strongly related to scale. It is easy to envisage that some indicators, useful at the community scale, may be difficult to measure at the regional scale and some regional indicators may have little meaning at the community scale. Due to these indicators compressing so much ecological, economic or social information into a single variable or sets of variables, it is vital that they are chosen, measured and interpreted carefully. If not evaluated carefully, the indicators may bias results in a particular direction and cause incorrect management decisions with disastrous consequences which bring the future usefulness of the data and its collection into question.

When an indicator is based on subjective probabilities, accuracy can be affected by a whole array of biases that require a system developer to make ethical decisions. Saveland *et al.* (1988) discuss examples of some of these biases in relation to forest fire management in the United States, including hindsight bias, preference for certainty and loss versus benefit. In hindsight bias, people consistently exaggerate what could have been anticipated in foresight. Also, people tend to favour certainty over uncertainty, so if a fire is controlled immediately, there is no need for a forest manager to worry about any future development. Finally, potential losses have a greater influence on fire decision than potential benefits (Saveland *et al.* 1988) due to the risk-adverse behaviour of fire management personnel.

Error quantification

In theory, error limits of predictions within a system should always be quantified. However, while error limits of individual equations can be easily obtained through the statistical fitting process, it is unusual in practice for system models to actually compute the consequences of error (uncertainty) when combining multiple equations. Also, many statistical processes rely on algorithms that approximate a mathematical process, and this in itself can produce unknown errors.

Mowrer (2000) examines error propagation in natural resource simulation models and presents several approaches (Monte Carlo simulation and Taylor series expansion) to project errors. In Australia, this is becoming increasingly important in the area of carbon accounting, as several models are typically used in combination to predict future levels of carbon sequestration. Both Standards Australia (2002) and IPCC (2000) provide a useful framework for practitioners to quantify error propagation in natural resource simulation models.

Language and culture

How language and terminology are used to define a question can greatly influence the collection and applicability of data or information. This is true for any information system in which users are forced to communicate using concepts unfamiliar to them. Cultural mismatch is of special significance in studies of indigenous peoples where the participants may have a different concept of forest than the researcher or practitioner. Often indigenous peoples have a holistic view of the land that includes trees, plants, animals and people (Buchy and Hoverman 1999). Once basic cultural differences are identified, the important challenge becomes one of understanding the ramifications of these differences and how they affect data needs and data use.

Presentation of information

Certain ethics apply to the accuracy with which results are portrayed by data management system outputs. Many of the biases, both unintentional and intentional, affect both information collection and the interpretation of results. While many issues relating to presentation arise inadvertently, presentation style can be deliberately selected to influence perception as illustrated in the classic book *How to Lie with Statistics* (Huff 1954).

As described in the book *How to Lie with Maps* (Monmonier 1996), when data appear in the form of maps the influence on perception can at times be extremely subtle:

> As a scale model, the map must use symbols that almost always are proportionally bigger or thicker than the features they

represent. To avoid hiding critical information in a fog of detail, the map must offer a selective, incomplete view of reality.

When features are generated automatically in a GIS, the potential to mislead, although often unintentionally, is high. For example, map generalisation (Buttenfield and McMaster 1991) is a common GIS function, but key features of interest may disappear when map scale changes. Monmonier (1996) discusses the role of this process in attempts to hide adverse results in environmental impact assessments. As with statistical measures, one should not provide a single value (or map) as completely representative of the current situation. A variety of maps at different scales should be provided to offer a more complete picture of reality.

The ability to influence perception by selectively, and often unethically, filtering information when it is passed from one level to another within an organisation can result in invalid decision-making at the managerial level (Bella 1992). Statistics, images, graphs and maps are all methods of summarising, presenting or filtering information for forest research and management decision-making. Ethical decisions should provide a framework for the selection and presentation of information and consequently may significantly affect the accuracy with which the recipients may perceive a situation.

Accessibility

Accessibility is essential to ensure the future availability of data. Data managers must ensure that users have physical access, an appropriate skill set and a correct decision-making environment, and also that they use a compatible system.

Physical access

Appropriate access to data and associated systems is dependent on both technical and intellectual processes. To correctly utilise a system, users must have access to the appropriate hardware and software technology. In addition, they must also be able to provide the required inputs and have the intellect to understand the information presented. For example, in a web-based system, users must have reliable connections to the Internet and sufficient bandwidth. Each end-user must also have a browser compatible with the material sent, and the necessary helper applications or browser plug-ins for viewing and hearing content. If an intended user is located in a remote area (as is often the case with forestry applications in Australia) technological issues can be of great importance. For these reasons, when a forest management information system is developed, its implementation should be part of an integrated process that includes full consultation with all users. Implementation may include specifying duties for dedicated support staff such as technology transfer officers or field personnel.

Potential users

Accessibility is also limited if the outputs or results from a forest management information system are presented inappropriately. For example, data may be aggregated at a fixed scale that may have limited value for many potential users. In other situations, the use of language and concepts that are beyond the end-user's understanding or working vocabulary might render a decision support system useless for a large fraction of the intended users. Even though it is neither practical nor possible to accommodate all who might 'stumble upon' data or a system, potential users need to be well defined and understood by developers to minimise potential misuse of data.

Decision-making environments and skill level

A potential ethical dilemma exists when providing access to sophisticated data management systems. Forest management information systems may considerably increase the ability of users to make or influence decisions that were previously beyond the limits of their knowledge and experience (Garg 2002). For example, senior forest management may gain direct access to operational-level data and information summaries. This gives them the ability to bypass intervening distortions, resulting in potentially more accurate perceptions. However, greater accuracy is dependent on a forest management information system that has been developed with appropriate ethical considerations. In addition, senior-level managers must be both willing to use and capable of using the system to achieve distortion-free information sharing. It should be noted that this type of situation has been the bane of statisticians for years! The development of powerful, commercial statistical software packages, with point and click approaches, has allowed forest researchers and technical foresters to perform all manner of inappropriate statistical tests on data without a full understanding of what they are doing. While it is acknowledged that current statistical software manuals and online help contain a great deal of information regarding model specification and assumptions, they should not and cannot replace a well-founded understanding of basic statistical concepts by the user.

System compatibility

Digital networks increase our capability to connect systems, databases and information-rich environments, but they also raise many ethical problems. The goal of seamless, transparent and 'user friendly' information access makes *compatibility* an essential property of databases and vocabularies. This essential property requires both technical and intellectual processes to improve compatibility within local, regional, state and federal forest information systems. Compatibility ensures that systems, procedures and cultures of organisations are developed in such a way as to maximise opportunities for exchange and re-use of data, whether sourced internally or externally. For example, 'forest health scores' may exist in separate state forest service databases and may be merged into a national-level database, but the definition of forest health score may be fundamentally different in each state. This will lead to incompatibility in the merged national database. If databases are, at times, incompatible in definitions and measurement, their use is likely to result in inconsistent research outcomes and management decisions. These inconsistencies may lead to individuals and organisations restricting access to their data. Recently, there has been a national initiative to harmonise forest health scores for young eucalypt stands (Stone *et al.* 2003).

Discussion

Data management and its role in decision-making creates numerous ethical issues — right and wrong, honesty, reliability, loyalty, responsibility, truth, accountability and fairness. Privacy,

accuracy and accessibility are also issues that must be considered in the development and delivery of data management systems. Choosing a particular approach to data management may either hinder or facilitate dealing with these issues. Historically, developers of forestry data management systems have focused on the technical aspects of systems, whereas today numerous social and cultural factors have resulted in a greater focus on ethical standards. A hybrid methodology that simultaneously addresses both technical and ethical requirements may be the solution.

Privacy has long been considered an inherent right of individuals in a democratic society. Initially, this involved protecting individuals from unwanted or unwarranted invasion of their physical space. More recently, the definition of privacy has extended to include an individual's information space. In the natural resource sector and the forest sector, threats regarding future availability of data information systems currently under development are more likely to arise from unintentional and unforeseen information breaches than from intentionally malicious behaviour. These breaches can occur when information sources are combined or used in unintended ways. As long as information about individuals exists and is accessible by others, individual privacy can potentially be compromised. During development, system designers need to be aware of users, co-developers, the public, cultures, special interest groups, commercial enterprises, state and federal governments and other groups that may be affected directly or indirectly by use of their data. Designers must also consider the information that a system uses or generates, and the decision-making landscape that it affects or creates.

The use of inappropriate communication and language is often at the centre of problems associated with accuracy. Even if an information system does not estimate the accuracy of results explicitly, it is important to ensure that end-users are aware of the potential variability in outcomes and the implicit assumptions and trade-offs that have contributed to these outcomes. Similarly, non-textural rendering of system outputs should be designed to address accuracy concerns in the flow of knowledge. It is also essential to address the way in which knowledge flows through organisational hierarchies, and to ensure its appropriate use at different organisational levels through the creation of an explicit ethical framework.

As well as accuracy issues, communication is at the centre of many accessibility issues. Information delivery must be aligned to the concepts appropriate to intended users, and information overload avoided, as knowledge may be inaccessible if the recipient is swamped with information. Limitations in technical accessibility by some groups may entail the development of an integrated range of systems and processes to ensure access by all stakeholders engaged in decision-making.

Organisations, including the Institute of Foresters of Australia, have developed codes of ethics for their members. However, these codes are not always well explained, may contain few sanctions and are seldom applied. In addition, the codes tend to hold individuals at fault, not whole organisations, and as a result do not make people more ethical. While there will always be some ethical responsibility on the individual forest researcher or practitioner, most of the responsibility still resides with organisations to instil standards of ethical conduct that create an atmosphere of social morality for their employees and members. Self-regulation is more readily accepted and more effective than regulation by governmental institutions which may not always fully understand the issues involved. Thinking in advance, especially at the organisational level, about ethical issues that may eventually impinge on data availability in the forestry profession is likely to alleviate potential future restrictions and allow us to develop and manage forest systems with greater understanding and accuracy than before.

References

Bella, D.A. (1992) Ethics and the credibility of applied science. In: Reeves, G.H., Bottom, D.L., Brookes, M.H. (Technical coordinators) *Ethical Questions for Resource Managers*. USDA Forest Service General Technical Report PNW-GTR-288, pp. 19–32.

Buchy, M. and Hoverman, S. (1999) *Understanding Public Participation in Forest Planning in Australia: How Can We Learn from One Another*. Australian National University, Canberra, 81 pp.

Buttenfield, B.P. and McMaster, R.B. (1991) *Map Generalization: Making Rules for Knowledge Representation*. Longman, Essex, UK.

Garg, R.K. (2002) *Forest Management Information Systems*. Microdot Systems (P) Ltd, New Delhi, India.

Hart, M. (1999) *Guide to Sustainable Community Indicators*. Hart Environmental Data, North Andover, MA.

Holling, C.S. (1978) *Adaptive Environmental Assessment and Monitoring*. John Wiley and Sons, Chichester.

Huff, D. (1954) *How to Lie with Statistics*. W.W. Norton and Company, New York.

IPCC (Penman, J. *et al.* (eds))(2000) *Good Practice Guidance and Uncertainty Management for National Greenhouse Gas Inventories*. Intergovernmental Panel on Climate Change. Published for the IPCC by the Institute for Global Environmental Strategies, Japan. http://www.ipcc-nggip.iges.or.jp/public/gp/english/

Mason, R.O. (1986) Four ethical issues of the information age. *MS Quarterly* **10,** 5–12.

Monmonier, M. (1996) *How to Lie with Maps*. 2nd edn. University of Chicago Press, Chicago.

Mowrer, T. (2000) Uncertainty in natural resource decision support systems: sources, interpretations and importance. *Computers and Electronics in Agriculture* **27,** 139–164.

Saveland, J.M., Stock, M. and Cleaves, D.A. (1988) Decision-making bias in wilderness fire management: implications for expert system development. *Applications in Natural Resource Management* **2,** 17–30.

Standards Australia (2002) *Carbon Accounting for Greenhouse Sinks: Afforestation and Reforestation*. AS 4978.1(Int) — 2002. Standards Australia International, Sydney, 20 pp.

Stone, C., Wardlaw, T., Floyd, R., Carnegie, A., Wylie, R. and de Little, D. (2003) Harmonisation of methods for the assessment and reporting of forest health in Australia — a starting point. *Australian Forestry* **66,** 233–246.

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

Whiteside, I.D., West, G.G. and Knowles, R.L. (1997) *Use of Standpak for Evaluating Radiata Pine Management at the Stand Level*. New Zealand Forest Research Institute Bulletin No. 154, 14 pp.