Australian Government

This program is funded by the Australian, State and Territory Governments' National Action Plan for Salinity and Water Quality.

Metrics for MBIs: A summary of the literature

Patrick O'Connor

Project team

Patrick O'Connor

Acknowledgements

This document was produced as part of the national Market-based Instruments Capacity Building Programme, a project of National Market-based Instruments Pilot Program. Funding was provided by the Australian, State and Territory governments through the National Action Plan for Salinity and Water Quality. The National Action Plan for Salinity and Water Quality (NAPSWQ) is a joint Australian and State and Territory Government initiative that encourages governments and regional communities to work together to address salinity and water quality issues in priority catchments.

Contact details

For further information or copies, please contact:

Social and Economic Unit, Community Partnerships

Queensland Department of Natural Resources and Water

GPO Box 2454, Brisbane, Qld 4001, Australia

Telephone: (07) 3239 3875

Fax: (07) 3224 2363

Email: Social.Economic@nrw.qld.gov.au

www.marketbasedinstruments.gov.au

This report may be cited as:

Collins, D. and M. Scoccimarro. 2008. *Market-based Instruments Decision Support Tool*. Market-based Instruments Capacity Building Programme, Market-based Instruments Pilot Programme and the National Action Plan for Salinity and Water Quality. Department of Natural Resources and Water, Brisbane.

9 311662 177308

Disclaimer

Table of contents

Introduction

Metric design and use in market-based instruments (MBIs) in natural resource management (NRM) is a relatively new enterprise and relies heavily on the science, economics, policy development and implementation experience of previous NRM programs and recent MBI programs and trials. The collective experience in design and use of metrics in MBIs is only partially documented, is incomplete for some MBI types and NRM issues and can be inaccessible to policy and implementation teams interested developing and using an MBI. This literature summary has been prepared to provide a précis of work on issues fundamental to good metric design and use. The summary is arranged under headings related to criteria for good metric design and the process for design.

This document is written to support *Metric Essentials* and to direct readers to relevant studies and program reports for further reading. It is not an exhaustive bibliography of metric or MBI-related reports but aims to provide readers with evidence for key concepts and a starting point for further investigation.

What does a metric do in supporting the implementation of MBIs?

In simple terms, a metric defines what is bought and sold in the market. Many existing markets and institutions do not adequately value environmental goods and services and the resulting market failure results in misaligned incentives between private providers and society (Stoneham et al., 2003). Market failure can occur because of asymmetries in information between landholders and environmental agencies. In simple terms, the purchaser (i.e. environmental agency) knows best what they want to purchase, the seller (i.e. landholder) knows best what the value of the goods and services or property rights are, the MBI operates to efficiently reveal the information held be each party to the other to achieve cost-effective management of the natural resource. The metric is the quantitative expression of what is being bought and sold.

Eigenraam et al. (2007) suggest that an environmental agency, not landholders, knows its preferences and priorities regarding different environmental assets and an agency may be in a better position than landholders to predict how management actions will enhance environmental assets. The environmental agency needs a ‘metric’ to describe the good or service that would come from improved environmental management by the landholder. The metric then becomes a tool for discriminating between alternative investments in the market of landholder offers.

Other benefits of metrics

The main use of a metric is the quantification and combination of multiple attributes into a single weighted score. This score can then be used as the basis for a range of decisions, including decisions about allocation of funding through an MBI. The metric score can also be used for evidence-based prioritisation of different sites or projects for purposes other than implementing an MBI. Connor et al. (in press) have shown that much of the efficiency gained by using a conservation tender instead of a uniform payment incentive scheme (in the Catchment Care program) came from improved measurement of the relative values of different actions (or similar actions in different parts of the landscape) and the subsequent ability of the program to prioritise and fund the highest value for money projects. Using the metric to rank priorities for investment has benefits which may persist regardless of the mechanism of incentive distribution.

A well-designed metric which provides an evidence base for prioritisation could be used to prioritise other intervention programs, not only incentive payments. Depending on the cost of developing and collecting data for the metric, the metric may be an efficient tool for determining the priority sites for preparing management plans or even targeting priority landholders for capacity building. Conner et al. (in press) were able to use the metric developed for the Catchment Care program to evaluate the cost-effectiveness of the pre-existing fixed-price incentive scheme. Having collected data for the metric, decisions could also be made about priority sites for inclusion in a program monitoring the environmental assets and threats.

What are essential elements of good metric design to effectively support MBIs?

A number of authors (Parkes et al., 2003; Oliver et al., 2005; Whitten, 2005b; and Rolfe, 2007; Whitten et al., 2007) have described frameworks for metric design which combined incorporate the principles:

- focus on program objectives and priorities
- combine quality and quantity assessments where appropriate
- provide an objective, reliable and repeatable measure of goods and services
- be simple to understand and explain to applicants and transparent enough to allow fairness
- enable manageable calculation of net change or outcomes
- ensure design and implementation are cost-effective
- allow comparisons and discrimination between a range of realistic scenarios or groups of goods and services
- be defensible if data is limited or uncertainty high

- consider the reversibility of impacts, side effects, market interactions and chance of success
- take into account trigger thresholds that would have a major impact on the desired outcomes
- consider time lags for outcomes to be realised.

An additional consideration for metrics is to be clear about what is not measured. Once investment has been made in the framework, data collection, storage and analysis for a metric there can be temptation to use the metric for other purposes, some of which may stretch the logic used to develop the metric in the first place. McCarthy et al. (2004) suggest a number of qualifications to the use of the 'habitat hectares' metric (Parkes et al., 2003) for purposes outside the original use. These include questions about the use of the metric as a tool for setting region-wide or statewide objectives, limitations for assessing sites with natural disturbance regimes, usefulness of comparing quality between different vegetation types, and use of the metric to evaluate conservation values in intensively cleared landscapes. If a metric is designed to support the MBI in achieving the defined objectives of the program, it may not be appropriate for other uses, or may require modification for additional use.

A common misuse of metrics for MBIs is the use of metrics designed for decision making (e.g. as the basis for discriminating between bids in a conservation auction) as the basis of monitoring programs. Monitoring may be best achieved using a subset of attributes of the metric, a combination of attributes not included in the metric or other indicators which are more sensitive or more cost-effectively measured and reported than the metric. The power of a monitoring design is critical to detecting the change of interest (Field et al., 2007) and may be higher if alternate indicators are used to those used in the MBI. The monitoring design should also balance power and cost-effectiveness (Field et al., 2004) and the appropriateness of the metric for monitoring should be considered in light of this constraint.

Table 1 illustrates how the basic principles of metric design were dealt with in a case study price-based MBI focused on reducing recharge to saline aquifers over a 10–30 year period.

Table 1. Principles of metric design and Wimmera Steep Hill Country (SHC) case study recommendations for addressing principles (modified from Whitten and Shelton, 2005).

Design principle	Description	Wimmera (SHC) metric case study recommendations
Quantity/Quality	A physical quantity or index of biophysical outcomes. There are usually a number of measures that deliver different messages to landholders and represent subtly different outcomes. For example, estimating salt discharge differs from estimating change to recharge volumes. Usually outputs are estimated using a proxy based on changes to inputs. For example, using models relating input changes (area and location of vegetation) to changes to salt outcomes.	Estimate salinity outcomes in tonnes of salt using input-based proxy measures based on changes to vegetation cover and other salinity-reducing management actions across the designated site
Relative change (or additionality)	Important if the goal of policy is to improve outcomes from a baseline, rather than to pay for some absolute maximum quantity or secure ongoing provision. The baseline is usually defined as the higher of what would otherwise happen (often termed business as usual), or a specified duty of care	Change should be measured relative to a uniform benchmark for value of business as usual. This reduces the difficulty of collecting baseline information from each site and creates an implied minimum duty of care. Bids in areas with scattered trees may complicate this baseline.
Location	The location where change occurs can generate different values to the community. Location is incorporated in the biophysical measure of quantity. This leads to three further related considerations: 1. Does the path to the point of estimation matter; 2. Are there any biophysical thresholds (or hotspots) that	The contribution of individual changes will be measured at the best downstream location for determining their relative values. The boundaries of those eligible to tender may be an alternative means of taking location into account.

	are likely to be created or impacted in different pathways? 3. Do any packages of management change generate synergistic outcomes?	
Timing	All things being equal, earlier outcomes are preferred over more distant outcomes	A steady-state estimate is favoured due to the relatively short time horizons and uncertainty in time to achieve change to outcomes
Risk or certainty of management change effectiveness	Some management changes may be more likely to be successful than others. The key factor in success may be the initial establishment or the on-going management. Likelihood of success can either be considered within the metric design or in the payment mechanism	Weight by assessed probability of successful implementation
Risk or certainty of outcome success	Even with successful establishment of the management change there may be uncertainty about the eventual impacts on outcomes. For example, this may be the case with management changes for which less is known about their impact on recharge.	Weight by assessed probability estimated outcome being achieved
Irreversibility	Irreversibility is related to risk. Where thresholds are anticipated, such extinction of species, there is a case for favouring less risky actions that achieve change sooner.	No irreversibility issues identified
Spillover impacts	Spillover impacts are consequences of the specific management change elsewhere in the system. For example, reducing recharge may affect base-flows in streams and rivers in the catchment.	No serious mechanism design issues identified. Wimmera SHC will also reduce base-flows in streams and rivers in the catchment. In some cases this can lead to a perverse outcome whereby the salt concentration in the remaining flow can be higher.

Focus on program objectives and priorities

The attributes in a metric, how they are combined and weighted and the data available for use in the metric determine what the metric measures. To be effective in achieving policy and program objectives, a metric needs to be constructed to measure the right natural resource properties, processes and management outcomes to meet stated objectives. Where the endpoint of the program is not well-defined it is too easy to develop a metric based on available data or theory at the expense of targeting the key outcomes required (Failing and Gregory, 2003).

Oliver et al. (2005) describe the policy and program objectives which drove the development of a biodiversity metric for the NSW Environmental Services Scheme (ESS). The policy objective was to promote land-use or land-management change that improved the provision of a range of environmental services (Oliver et al., 2005). The types of environmental services that would be examined were determined first, and then the best ways to measure a range of environmental services were developed (Grieve and Uebel, 2003). Metrics also needed to provide a measure for use at the property, catchment and regional scales.

The BushTender trial (Stoneham et al., 2003) had a policy objective of establishing contracts for nature conservation with private landholders. The metric used in BushTender (including the Habitat Hectares component (Parkes et al., 2003) focussed only on measurement of the biodiversity significance, management improvements and cost of individual bids in a conservation auction. The same metric has subsequently been combined with metrics for other environmental services to achieve a broader set of policy and program objectives from the EcoTender trial (Eigenraam et al., 2008). EcoTender has the extended policy objective of cost-effectively establishing contracts for provision of several jointly supplied environmental services (for improvements in terrestrial biodiversity, aquatic function, saline land area and carbon sequestration). The policy driver for EcoTender extends that of BushTender to capture a greater range of environmental services but also to avoid perverse outcomes where improvement of provision of one environmental service has a negative impact on the provision of other services, especially where these services are undervalued by existing markets. While there are many advantages from designing and implementing metrics and MBIs covering multiple environmental services is not always easy and programs and policies need to be guided by these challenges (e.g. Gole et al., 2005).

Combine quality and quantity assessments where appropriate

Metrics for MBIs need to provide an objective, reliable and repeatable measure of the goods or services. Different NRM issues require different metrics because the goods and services differ and may be simple quantities (e.g. megalitres of water) or complex qualities (e.g. condition of native vegetation) of the natural resource. Measures of quality often combine a number of attributes which represent the composition or function of the natural resource.

While quantity metrics may be simple, they will not always be simple to calculate. Whitten et al. (2005a) used complex biophysical modelling to measure recharge (megalitres) at different scales (from the paddock scale to the region scale). Measurement of recharge required a number of estimations and judgements including the estimation of diffuse source net recharge to shared irrigation aquifers and the identification of net recharge impact boundaries (Whitten et al., 2005a).

Gibbons and Freudenberger (2006) reviewed indices of vegetation condition and reported that there is no standard definition of the 'quality' of vegetation but that quality had been defined in different ways according to use. The assessment of quality is often a matter of reliance on theory, expert knowledge and the measurement of surrogates. The River Murray Forest Project in South Australia derived a metric for a trial single-sealed bid reverse auction for revegetation with both carbon sequestration and biodiversity conservation values. This simple metric converted continuous data to categories for the diversity of species to be planted, the distance to and size of nearby remnant vegetation, the total and individual patch size of plantings, and measures of the carbon sequestration potential and security (O'Connor and Collard, 2007). More detailed datasets and more sophisticated models of the system could include additional attributes and combine them in different ways. However, the cost-effective solution of using simple attributes for surrogates of biodiversity and carbon-sequestration values enabled the market for revegetation to be tested ahead of more detailed metric availability.

The assessment of quantity and quality may rely on surrogates, estimates or models and the assumptions underlying these should be documented along with the metric. An example of this is the use of spatial data either for quantity (e.g. hectares of native vegetation) or quality (e.g. percent remnant native vegetation—a potential component of a landscape context metric). The requirements for standard but sophisticated spatial analyses of attributes such as landscape context should not be underestimated (Gole et al., 2005). The use of models such as those underlying geographic information systems (GIS) should also be understood and the accuracy of models should be tested and reported.

Provide an objective, reliable and repeatable measure of goods and services

Failing and Gregory (2003) highlight the problem of developing lists instead of indicators when trying to describe the condition of ecosystems. There is a tendency to use indicators for which there is already data or well-developed understanding of a process rather than high representation of the core attributes of the system which are important in the specific management context (Failing and Gregory, 2003). Indicators (or attributes in a metric) will be most useful when they are selected as appropriate measures or surrogates for key properties of the system, after the program objectives have been defined.

A large number of metrics measuring environmental assets, processes and outcomes have been developed for different purposes. However, not all of these metrics have been tested for reliability and where they have been tested some inconsistency has been found. Cushman et al. (2008) examined a large set of metrics to determine if there is a parsimonious set of metrics which reliably measure landscape structure. The study showed that few metrics were universally appropriate for all use in all locations (only eight out of 54 metrics measured attributes present in all the locations tested), and only seven out of 49 components of landscape structure explained more than 10% of the variance of the metrics tested (i.e. many of the metric attributes were not very important in many of the locations tested). This study demonstrates that metrics may be robust in one location or for one set of examples but should be tested for new use in new applications and locations. Metrics need to be tested for attribute redundancy to ensure they are reliable and consistent (Cushman et al., 2008).

Andreasen et al. (2001) provide an example of the issues which need to be addressed in developing an objective, reliable and repeatable metric. They describe considerations for developing a metric to measure ecological integrity in terrestrial systems. Key considerations include the need for the metric to be meaningful at different spatial scales, grounded in natural history, relevant and helpful, flexible, measurable and comprehensive enough to measure composition, structure and function. Andreasen et al. (2001) also suggest the information useful for evaluating a candidate metric, combining attributes into a metric, and testing a metric.

Attributes in a metric also need to be weighted to reflect their relative importance. The weightings need to take account of the relationship of attributes to each other and to the end-point or objective of the metric/MBI. Weighting is a complex task and must be done using reliable and defensible methods. Some of these methods are discussed later in this summary (e.g. see the section on dealing with limited data and uncertainty below) and others can be reviewed in texts such as Clemens (1996) and von Winterfeldt and Edwards (1996).

The methods used to combine attributes in a metric are also important and need to be considered to avoid redundancy (where different attributes measure part or all of that measured by another attribute) and compensation (increase in one attribute masks the decrease in a related attribute). Caution is recommended when adding attributes together in a quality metric due to the risk of compensation of one attribute for another. The example given from McCarthy et al. (2004) illustrates that the loss of trees may be compensated by the increase in coarse woody debris in a metric which combines the two (the example discussed is from the Habitat Hectares metric (Parkes et al., 2003). These two attributes are not equivalent or interchangeable but an additive metric may not distinguish between two cases where the decreasing value of one attribute is compensated by the increase in the other. Weighting attributes or multiplying attributes together can help overcome the problem of compensation between attributes. The problem of arriving at a zero value for the metric, if only one attribute has a zero value but the other attributes do have meaningful values, is avoided if the minimum value of each attribute is set above zero (McCarthy et al., 2004). In practice, a mixture of addition and multiplication is usually appropriate (Gibbons et al, 2005).

Be simple to understand and explain to applicants and transparent enough to allow fairness

MBIs aim to establish contracts between parties which optimise the environmental benefit from investment. A number of authors have highlighted that market efficiency is affected by information asymmetry (Laffront, 1990; Latacz-Lohmann and Van der Hamsvoort, 1997) and that land managers may not have all the relevant information about government priorities in an MBI or how information about priorities would influence contracts established through the MBI. One mechanism for defining and quantifying priorities is the MBI metric. Communication about the program objectives and priorities must go beyond explanation of the metric; however, a metric which is easily explained will assist the communication. Parkes et al. (2003) and Oliver et al. (2005) both highlight the need for the metrics developed in their programs to be simple to understand. This does not mean that all attributes and their scoring will necessarily be understood by all participants, but that land managers can be provided with a clear message about the important components of the environmental services which are sought through the MBI (and metric).

Overly simple metrics can be criticised as not representing all possible contributing attributes of the system or of not being ‘scientifically’ defensible. Scientifically defensible in its simplest form means that the logic and data used to construct the metric come from a systematic and falsifiable approach to decision making. Where judgements need to be made to accommodate uncertainty and knowledge gaps, the judgements should be made on the best available evidence and programs implemented to examine outcomes and improve the metric (Failing and Gregory, 2003).

While providing clear information about metric components can improve the efficiency of the market, it is not necessary to reveal all the weighting and scoring of a metric. Cason et al.

(2003) have shown that landholders may misrepresent their costs if they know they are offering high-quality benefits and the total environmental benefit from a fixed MBI budget may be lower. Metrics and MBIs can still be designed to be simple to understand and explain while absolute scoring and weighting processes may remain hidden if it is though market efficiency will be improved (Stoneham et al., 2003).

Gole et al. (2005) report that landholders involved in the Auction for Landscape Recovery (WA) valued contact with community support officers who were able to explain the MBI program to them. They also report that some landholders experienced initial difficulty working with the design of the program which may have reflected difficulties in defining and communicating the objectives of the program and how multiple benefits would be assessed.

Enable manageable calculation of net change or outcomes

Stoneham et al. (2003) and Oliver et al. (2005) describe metrics designed to incorporate the current status of environmental services and the predicted change following management change. Both these metrics require some estimation of a transformation function—the estimated change from current value to future value following intervention (this could include prevention or slowing of decline as well as improvement in the condition of the natural resource of interest). Reliable scientific information on the nature of the relationships between land-use change and ecosystem impacts is critical for the functioning of environmental markets (Whitten et al., 2005a). Stoneham et al. (2003) recognise that the transformation function for the BushTender trials was not known with certainty and that extreme or unexpected events are difficult to include in estimated functions.

The EcoTender program has taken the calculation of change another step by implementing a catchment modelling framework (CMF) to estimate multiple environmental benefits from revegetation and remnant native vegetation maintenance and improvement (Eigenraam et al., 2007). The CMF was able to measure a change in the level of ‘service’ provided by the landholder for the multiple environmental benefits sought by the program. The program recognised that some on-farm management actions could be incorporated in the future but further research is required to determine appropriate monitoring and enforcement strategies.

A key impediment for the Auction for Landscape Recovery (ALR) was the inability to develop or use effective estimates of future management benefit of tendered projects and threat/risk analysis (Gole et al., 2005). Methods already used in other projects were considered of questionable value to the ALR, because they did not transfer well to the new environment. Gole et al. (2005) recommended a dedicated research program to develop transformation functions and provide workable and meaningful methodologies.

Most work in this area has relied on biophysical models (e.g. Whitten et al., 2005a; Connor et al., 2006) or expert knowledge, with or without the use of MCA (e.g. Stoneham et al., 2003; Oliver et al., 2005). Efforts to document changes resulting from large agri-environment schemes which have spent billions of dollars in Europe and North America have been patchy and produced mixed results (e.g. Kleijn et al., 2006). MBIs with *a priori* assessment of natural resource condition, description of changes sought through management (i.e. management plans and contracts) and adequate monitoring programs (see Field et al., 2007, for a discussion of key issues for designing meaningful monitoring) offer an opportunity to evaluate and improve transformation functions and metrics.

Because outcomes may not be realised for a significant amount of time and can be difficult to monitor or observe (e.g. status and resilience of plants and animals) many MBIs use metrics which determine expected gain from intervention but establish contacts on the basis of outputs. Stoneham et al. (2003) and Windle and Rolfe (in press) discuss the issue of contracting on the basis of outputs but estimate outcomes in the metric to discriminate between bids in the respective conservation tenders. Another example is the land management changes required by the Wimmera Steep Hill Country (SHC) case study to reduce recharge to a saline aquifer over 10–30 years require expensive upfront investment in order to produce ecosystem services in the longer term (Whitten and Shelton, 2005). However, the success of these actions can only be measured at a much later date, thus compliance is difficult to measure. The relatively large upfront capital investment required is likely to be beyond the capacity of most land managers and so a payment schedule based on significant upfront support and ongoing performance-based payments based on management inputs is likely to be needed. In this example the metric may be accurate in predicting the change expected from land-management change but the program proponent may still bear the risk of failure and it will be difficult and potentially costly to monitor compliance.

Ensure design and implementation are cost-effective

There are few reports of the total or itemised costs of metric design and implementation. Costs can also be difficult to calculate when metric design and implementation costs are not easily separated from other costs associated with MBI design and implementation and when institutional or in-kind support is provided but not accounted for in metric development. However, metric design and MBIs are not unique in this regard as cost accounting for non-MBI approaches to NRM can also be limited by similar problems. Windle and Rolfe (in press) demonstrated that while the initial design and development costs of a conservation tender in the Australian rangelands were higher than a grant scheme, there was no evidence of any difference in operating costs between the grant and tender-based schemes in their case study.

Both the design and implementation of metrics need to be taken into account when estimating cost-effectiveness as metrics with low data collection and analysis costs will be more cost-effective. Administrative costs can account for 30%–80% of the total cost of an agri-environmental scheme (Falconer and Whitby, 1999) and metric costs are can be a significant component of the administrative cost. Data collection is often expensive and some programs have tried to limit these costs by designing metrics which require lower levels of expertise and less time to collect the data (e.g. Oliver et al., 2007). Rapid approaches can be very cost-effective if the methods remain objective, reliable and repeatable. Beverly et al. (2005) compared detailed and rapid assessment methods for predicting salinity impacts from land-use change and found that only the complex models could identify the changes of interest at the paddock–farm scale required by the program.

Lowell et al. (2007) report that model-calibration costs for an MBI focussed on the establishment of forest plantations were approximately nine times usage costs, though subsequent calibration costs in the same catchments would be lower as a consequence of the initial work. These authors also report that the model calibration and usage costs constituted around 30% of the total operational cost (i.e. administration, modelling, on-ground activities) of the MBI and 20% of the total cost of the program. In the case of the ALR (Gole et al., 2005), the estimate of administrative costs, including all research and operational costs, was around $500,000 and with an expenditure on transfer payments to landholders of around $200,000 administrative costs account for 70% of the total (although this project was a fixed-budget MBI pilot project).

Metric design and implementation will be more cost-effective if subsequent programs can use the metric and operate more cost-effectively as a consequence of the earlier work. There will usually be additional benefits for planning from the availability of additional data and this will improve the overall cost-effectiveness of metric implementation.

Allow comparisons and discrimination between a range of realistic scenarios or groups of goods and services

A number of MBI projects have conducted pilot programs or case studies as part of the testing the MBI design and metric. For example, Eisner et al. (2007) conducted a pilot program of using 10 applications in a competitive-tender MBI aimed at changing agricultural management practices in the Great Barrier Reef catchments to reduce sediment and nutrient loads impacting on the reefs. The metric was able to distinguish between different management-practice changes within a single agriculture type and between different sites within a sub-catchment.

The Catchment Care competitive auction for biodiversity conservation and water quality gains tested the metric under the full range of auction conditions before applying it in the trial, and made adjustments to the metric to ensure it was selecting bids consistent with the goals of the program (Bryan et al., 2005). The tests used created 1000 bids with a complete mix of plausible scores for all metric attributes (and groups of attributes). Stepwise multiple linear regression was used to identify attributes that most strongly influenced the metric score and the metric was refined after simulations to better reflect the goals and priorities of the program (Bryan et al., 2005).

O'Connor et al. (2007) and O'Connor and Collard (2007) also conducted sensitivity analysis of metrics designed to discriminate between bids in a conservation tender and a tender for revegetation contracts, respectively. These sensitivity analyses allowed examination of the relative importance of attributes in the different metrics and assessment of how different potential bids would rank at the end of a competitive tender. Both projects used simulated datasets to examine the ability of the metrics to compare different groups of goods and services.

Be defensible if data is limited or uncertainty high

Data for decision making is often lacking and decisions often need to be made in an environment of uncertainty. Uncertainty in metric construction and use can arise from many sources but commonly arises from uncertainty about measurement error, model and model input parameter errors, spatial variability, errors in spatial data, the effects of aggregation of spatial data, temporal variability, the inherent variability in natural processes, model and model parameter uncertainty (Nguyen et al., 2006).

There is an increasing body of literature addressing decision making in uncertainty which is relevant to metric design for MBIs. A number of tools including decision tables and decision trees have been proposed for formal decision making that involve identifying three main components: acts, states, and outcomes (Resnik, 1987). The acts refer to the decision alternatives, the states refer to the relevant possible states of the system, and the outcomes refer to what will occur if an act is implemented in a given state. Ben-Haim (2001) advises that key to decision making in uncertainty is judicious use of available information, calculation of the consequences of incorrect decisions (both undesirable outcomes occurring or desirable outcomes being missed), the incorporation of value judgements and the preparedness of the decision-maker for risk-taking.

Eisner et al. (2007) note uncertainty in their metric measuring benefits from changing agricultural management practices in the Great Barrier Reef catchments to reduce sediment and nutrient loads impacting on the reefs. They note the uncertainty comes from both limitations and assumptions in the underpinning conceptual framework of the metric and from shortage of appropriate data to populate the metric. Eisner et al. (2007) used an expert panel to estimate the nutrient and sediment load reductions from different combinations of sugarcane management practice. They recognised that the disadvantages of using the expert panel instead of a model were potential lack of transparency of decisions and bias in the expertise available. However, benefits included some reduction in uncertainty by producing a highly flexible assessment system in which any combination of practices used by a farmer could be assessed (this capability was not available from existing models). Components of the system which had uncertainty that was considered too high were omitted from consideration in the metric (i.e. estimation of loss of nitrogen to the atmosphere). Eisner et al. (2007) also dealt with uncertainty in some metric attributes by expressing values for those attributes as ranges rather than as means.

Another approach to dealing with uncertainty is preparedness for trial and error and communication and discussion of findings. Oliver et al. (2005) recognised that data collection for their metric estimating biodiversity benefits from changed land use and management introduced some uncertainty through inter-operator variability (observer bias). These authors have designed their metric to enable scrutiny of the expert opinions on which they are based and the authors invite feedback on the method.

Multi-criteria analysis (MCA) has been used in a number of projects to combine and weigh components of a metric. Oliver (2002) and Oliver et al. (2007) facilitated expert panels to develop a set of indicators of vegetation condition based on their importance and feasibility using an analytic hierarchy process (AHP), a form of MCA. Hajkowicz (2006) also presents a framework suitable for using expert and other input to produce metrics for a range of purposes. These examples deal with uncertainty by allocating weighting to attributes using available data, knowledge and opinion in a transparent and defensible way. When this approach is combined with thorough uncertainty analysis (see section above that discusses discrimination between a range of realistic scenarios), uncertainty can be better understood and accounted for. Reckhow (1994) suggest the use of simple models with thorough uncertainty analysis where uncertainty is high, because complex models may be difficult to scrutinise (the Information Paradox of Rowe (1977).

Not all reports on metric design and MBI implementation report enough information about the construction of the metric, weighting of attributes, and approaches to dealing with uncertainty to permit scrutiny and learning by others. This is an issue for future learning, and MBI practitioners should be encouraged to document and report the assumptions and limitations of the metrics they design and use to allow further enquiry and development around problems of uncertainty.

Consider the reversibility of impacts, side effects, market interactions and chance of success

Where there is a risk of irreversibility there is a case for favouring less risky actions that achieve the change sooner (Whitten and Shelton, 2005). Metrics can be designed to take account of irreversible impacts, though this is often better dealt with through the rules established in the MBI design. Oliver et al. (2005) showed how land-management changes which led to difficult-to-reverse impacts on derived native grassland condition produced very low (negative) metric scores (Land Use Change Impact Score). Depending on the regulatory framework underpinning the MBI, the same irreversible impacts could have dealt with by not allowing the proposed negative land management change at all (regulation), or ruling land managers undertaking that management change out of the MBI. Metric design should ensure that unwanted irreversible impacts are not promoted or rewarded in an MBI.

Environmental benefits from natural systems are usually linked and cannot always be managed separately. Wu and Skelton-Groth (2002) demonstrated that the failure to recognise multiple benefits from funding for conservation programs could result in the loss of some of the conservation benefits. Strappazzon et al. (2003) argued that where multiple environmental goods were produced from an action, funded programs should concentrate on those goods for which there are not existing trading systems. This has been demonstrated in practice by Eigenraam et al. (2007), who conducted the EcoTender in Victoria to achieve multiple benefits for terrestrial biodiversity, aquatic function, saline land and carbon sequestration from revegetation and remnant vegetation management. The tender-based MBI calculated environmental benefits for the first three of these outcomes, with carbon sequestration benefits treated as though there were an existing market for that service. EcoTender considered how the multiple benefits from supported actions could be measured, but took the approach that where the program interacted with an existing (or likely) market for carbon sequestration, it may best to allow that market to access the carbon sequestered through the program (landholder were also able to sell their carbon sequestration units to the program at a fixed price, as though that was the ‘market price’ in an established market). EcoTender demonstrates the importance of purchasing multiple environmental outcomes where a focus on only one of the outcomes could produce perverse results with respect to another; for example, purchasing only the saline land benefits through revegetation may result in declines in aquatic functions if water tables fell too far.

Take into account trigger thresholds that would have a major impact on the desired outcomes

Trigger thresholds are quantity or quality states that, if reached, will result in non-linear improvement or decline in a defined value. In simple terms this could equate to ‘the straw that breaks the camels back’ or saving enough money for an overseas trip—all amounts less than the threshold do not trigger the outcome. In natural resource management it may be possible to change management on a number of properties but it will only be after a specific property joins the program that the results are really noticeable. An example of this would be conservation works on remnant vegetation patches that are widely separated in a fragmented landscape. While each patch may improve due to management, potential gains from species movements between sites (dispersal) may not occur until an important connecting site in the landscape is managed for conservation. An example of this is the MBI designed by Rolfe et al. (2005) to improve landscape linkage for biodiversity in the southern Desert Uplands of Queensland. The MBI design recognised the different values of investment in different parts of the landscape and introduced a ‘limited cooperation’ model to improve landholder involvement in determining locations for remnant vegetation corridors and so achieve higher order outcomes than would have been achieved if individual properties submitted bids in a conservation tender without awareness of the greater gains from conservation on contiguous patches of remnant vegetation.

The potential to benefits from optimising contract selection in MBIs has been taken further by Gole et al. (2005) in the Auction for Landscape Recovery in Western Australia. This project recognised that there could be strong interdependencies between the potential contracts for conservation work being funded. Bid interdependencies mean that the decision to fund any one contract changes the benefits score for other projects. For example, when the goal is regional biodiversity conservation, the biodiversity contribution of a contract reflects only those components of biodiversity not yet captured by other funded projects (the principle of ‘complementarity’, see Margules and Pressey, 2000). Hajkowicz et al. (2007a) re-analysed the data from the Auction for Landscape Recovery using different models of complementarity and showed that the optimal outcome for conservation was not selected by simple bid ranking but depended on the rules used to determine complementarity. Despite the additional benefits of more sophisticated methods for selections contracts to capitalise on complementarity, Gole et al. (2005) report that the challenges of communicating the process and its conceptual and computational complexity was a key limitation of the project.

Consider time lags for outcomes to be realised

Transition function must reflect the change within the time specified in the contract or the expected gains may be reduced if land management or use change after the contract period finishes and before the expected gains are realised. For some MBIs the time lag may be simple (e.g. the benefits from a cap and trade on water should be realised within a short period). For other MBIs the time lag may be difficult to determine or not occur for many years (e.g. carbon sequestration will increase until trees reach maturity). Published reports have not clearly specified the expected timing of outcome achievement or the types of contracts and whether contracts remain in force until the expected time of outcome achievement. For example, the BushTender metric includes a services score attribute for retaining dead trees and logs (not removing firewood) and contracts in the BushTender trial were established for actions over three years. It is not reported whether the outcomes from management services such as retaining dead trees and logs are expected to be realised in three years (Stoneham et al., 2003). The time lag to expected outcomes also means that services scores contracts of different length (within the same MBI) should be weighted according to the benefits likely after each time period. Benefits will rarely accrue linearly with time. Because the type, amount and timing of changes from management intervention are difficult to predict and assess, many MBIs have established contracts on the basis of outputs rather than outcomes (e.g. Stoneham et al., 2003; Rolfe et al., 2005; Connor et al., 2006).

How do metrics vary according to the type of MBI or NRM issue?

Metrics define and measure goods and services and a given metric should be able to be used in a range of different MBIs. Once environmental benefits can be defined and measured in a metric, the metric should be able to be used in the most appropriate MBI available. The Victorian BushBroker scheme (an offset credit scheme; Crowe, 2004) and EcoTender (a multiple-outcome tender-based scheme; Eigenraam et al,, 2007) use a consistent assessment method for scoring vegetation quality and gain through management actions designed for the Victorian BushTender metric (conservation tender; Stoneham et al., 2003). Tisdall (2007) experimented with the same metric in auctions (two kinds) and cap-and-trade MBIs. Using the same metric for different MBIs has multiple benefits for cost-effectiveness and learning. Different MBIs using the same metric can be compared for cost-effectiveness (i.e. environmental benefit per dollar), leading to improved understanding of the market and subsequent improvement of the MBIs used. The consistency of using common metrics may also have benefits deriving from consistency of data collection and management systems within a jurisdiction.

However, different metrics for the same goods and services may need to be designed for use in individual projects, locations and MBIs because each project has a different starting point, different requirements for metric design and cost-effectiveness, and different theories of how actions link to outcomes.

Appendix 1 in *Metric Essentials* provides examples of metrics used for different NRM issues and different MBI types. Some reviews of metrics available for different NRM issues have also been undertaken for:

- waterborne pollution reduction—Eisner et al. (2007)
- conservation of remnant vegetation—Gibbons and Ryan (2007)
- salt and water yield under different management—Beverly et al. (2005).

Using or adapting an existing metric

Gibbons and Ryan (2007) reviewed four existing metrics to determine the benefits and costs of using or modifying an available metric compared to developing a new metric. The reviewed the suitability of the Forest Conservation Fund conservation value index developed to support a program to protect and manage old growth forest in Tasmania (Eigenraam et al., 2006); the assessment methodology underpinning BushTender, which has Habitat Hectares at its core (Parkes et al., 2003); the metrics underpinning the incentive component of the Property Vegetation Plan Developer used in NSW, specifically the BioMetric tool (Gibbons et al., 2005); and the index underpinning NatureAssist, which is a tender-based incentive scheme based in Queensland (Hajkowicz et al., 2007b). They found many advantages in adapting two existing metrics (Habitat Hectares and BioMetric) including cost and time savings, requirement for minimal training, and appropriateness for use in a conservation auction in box gum–grassy woodlands (BGGW). Limitations included some restrictions on use for non-state-government employees and the need to alter some management actions, attributes and the method for predicting change for use in the BGGW project.

Beverly et al. (2005) tested four model-based metrics for predicting whole-of-catchment mean annual salt and water yield under different management. One model (requiring intensive input data and solution times) provided finer temporal and spatial scale information within the catchment (farm and paddock scales) than the others and was preferred for that reason by the authors. Eisner et al. (2007) reviewed 19 metrics for transferability of the metric or components to a program investing in reducing waterborne pollutants in the Great Barrier Reef (GBR) catchments. The review provided useful lessons though none of the metrics provided a complete solution for the GBR. Components of a number of metrics were used to construct a new metric for the project. The main limitations for transfer of existing metrics were: different NRM issues addressed; lack of data or data at the appropriate resolution for the GBR project to use the metric; high cost or infeasibility of collection of data necessary for the metric; poor choice of indicators; potentially too complex for use by regional NRM groups. Some of the benefits noted were contributions to metric design; contribution for some attributes and models; understanding of metric limitations.

Eisner et al. (2007) also note that information in reports about metrics from other programs do not always allow thorough evaluation, an issue for documentation and communication of metric design by their designers.

What approaches are most useful to facilitate links between implementers, scientists and economic design experts?

Setting priorities, selecting attributes and determining weightings requires both technical and value judgements. It is important to realise that different people may be required to make technical judgements and value judgements and the processes used to combine these expertise sets need to reflect the capabilities and contributions of the different people (Failing and Gregory, 2003).

Oliver (2002) and Oliver et al. (2007) used several processes to facilitate expert panels to develop a set of indicators of vegetation condition based on their importance and feasibility. The approach included use of a Delphi technique to generate potential indicators without initial evaluation. The process used electronic mail to contact and receive contributions from experts, and asked an initial set of experts to nominate three other experts to contribute to the program. This approach not only allowed the program to access a large expertise set, it facilitated access to experts who were not initially known or identified by project staff and minimised any time or practical constraints for experts to contribute (e.g. compared to face-to-face workshops). The second phase of the program used a version of MCA, the analytic hierarchy process (AHP). This tool enabled analysis of the knowledge and opinions of experts about attributes which are most important surrogates of biodiversity and those which could be most feasibly assessed. Results indicate that an ‘arithmetic consensus by contribution’ can be reached and attributes categorised according to importance and feasibility. The study also highlighted that experts have some biases, in this case a correlation between importance weights for attributes and the spatial scales at which the experts work. Gole et al. (2005) report that the use of an expert reference group to undertake the main work of determining outcomes from management actions was a key success factor for the program.

References

Andreasen J.K., O'Neill R.V., Noss R. and Slosser N.C. (2001). Considerations for the development of a terrestrial index of ecological integrity. *Ecological Indicators* 1: 21–35

Ben-Haim, Y. (2001). *Information-gap decision theory*. Academic Press, San Diego, California, USA

Bryan B.A., Gatti S., Connor J., Garrod M. and King D., (2005). *Catchment Care—Developing an auction process for biodiversity and water quality gains*. A NAP market-based instrument pilot project. CSIRO Land and Water and the Onkaparinga Catchment Water Management Board.

Beverly C., Bari M., Christy B., Hocking M. and Smettem K. (2005). Predicted salinity impacts from land use change: comparison between rapid assessment approaches and a detailed modelling framework. *Australian Journal of Experimental Agriculture* 45: 1453-1469

Clemen R. (1996). Making hard decisions: An introduction to decision analysis, Duxbury Press, Pacific Grove, California.

Connor J.D., Ward J. and Bryan B.A. (in press). How Cost Effective are Conservation Auctions? *The Australian Journal of Agricultural and Resource Economics*.

Connor J.D., Clifton C., Ward J., and Cornow P. (2006). *Commonwealth MBI Pilot Project: Dryland Salinity Credit Trade.* National Market-based instrument Pilot Project (accessed at http://www.napswq.gov.au/publications/books/mbi/pubs/round1-project57.pdf)

Crowe M. (2004) BushBroker: A broker for biodiversity credits. In *Market-based tools for environmental management Proceedings of the 6th annual AARES national symposium 2003*. Eds Whitten S., Carter M. and Stoneham G. Rural Industries Research and Development Corporation (accessed at http://www.mobot.org/plantscience/CCSD/RNC%20Symposium/Reading%20materials/Whitten,%20Carter%20&%20Stoneham,%202004.pdf)

Cushman S.A., McGaringal K. and Neel M.C. (2008). Parsimony in landscape metrics: Strength, universality, and consistency, *Ecological Indicators* (in press) doi:10.1016/j.ecolind.2007.12.002

Eigenraam, M., P. Barker, M. Brown, A. Knight, and S. Whitten. 2006. *Forest Conservation Fund Conservation Value Index Technical Report*. Assessment Methodology Advisory Panel, Tasmania

Eigenraam M., Strappazzon L., Lansdell N., Beverly C. and Stoneham G. (2007). Designing frameworks to deliver unknown information to support market-based instruments. *Agricultural Economics* 37:261-269

Eisner R., Le Grand J., and Norman P. (2007). *A water quality metric for the Great Barrier Reef catchments*. Department of Natural Resources and Water, Queensland

Failing L. and Gregory R. (2003). Ten common mistakes in designing biodiversity indicators for forest policy. *Journal of Environmental Management* 68: 121–132

Falconer, K. and Whitby, M. (1999). The invisible costs of scheme implementation and administration. In: *Countryside Stewardship: Farmers, Policies and Markets*, Van Huylenbroeck, G. and Whitby, M., eds., pp. 67–88. Pergamon, Amsterdam.

Field, S. A., Tyre, A. J., Rhodes, J. M., Jonzen, N. & Possingham, H. P. (2004) Minimizing the cost of environmental management decisions by optimizing statistical thresholds. *Ecology Letters* 7: 669–675.

Field S. A., P.J. O'Connor, A. J. Tyre, and H.P. Possingham (2007) Making monitoring meaningful. *Austral Ecology* 32: 485–491

Gibbons P. and Freudenberger D. (2006) An overview of methods used to assess vegetation condition at the scale of the site. *Ecological Management and Restoration*. 7(S1):S10-S17

Gibbons, P., Ayers, D., Seddon, J., Doyle, S. and Briggs, S. (2005) *BioMetric Operational Manual Version 1.8: A Terrestrial Biodiversity Assessment Tool for the NSW Property Vegetation Plan Developer*. NSW Department of Environment and Conservation

Gibbons and Ryan (2007) A conservation value index for Box Gum Grassy Woodland. (document not published)

Gole C, Burton M, Williams KJ, Clayton H, Faith DP, White B, Huggett A, and Margules C (2005) *Auction for Landscape Recovery Final Report*. WWF-Australia. (accessed at http://www.napswq.gov.au/publications/books/mbi/pubs/round1-project21.pdf)

Grieve A. and Uebel K. (2003). *Developing New Income Streams for Farmers: NSW Environmental Services Scheme*. (accessed at http://www.forest.nsw.gov.au/env_services/ess/files/essREPORT.pdf.

Hajkowicz S. (2006) Multi-attributed environmental index construction. *Ecological Economics* 57: 122–139

Hajkowicz S., Higgins A., Williams K., Faith D. and Burton M. (2007a) Optimisation and the selection of conservation contracts. *Australian Journal of Agricultural and Resource Economics*. 51: 39–56

Hajkowicz, S., Miller C., Marinoni O., Higgins A. and Williams K. (2007b). *Quantifying conservation benefits: A metric to inform purchasing decisions under the Queensland Environmental Protection Agency's Nature Assist Program*, Environmental Protection Agency, Brisbane.

Kleijn, D., Baquero R. A., Clough Y., Diaz M., De Esteban J., Fernandez F., Gabriel D., Herzog F., Holzschuh A., Johl R., Knop E., Kruess A., Marshall E. J. P., Steffan-Dewenter I., Tscharntke T., Verhulst J., West T. M. and Yela J. L.. (2006). Mixed biodiversity benefits of agri-environment schemes in five European countries. *Ecology Letters* 9: 243–254.

Laffont, J. (1990). *The Economics of Uncertainty and Information*, MIT Press, Cambridge.

Latacz-Lohmann, U. and Van der Hamsvoort C. (1997). Auctioning conservation contracts, a theoretical analysis and an application. *American Journal of Agricultural Economics*. 79: 407–418

Lowell K., Drohan J., Hajek C., Beverly C. and Lee M. (2007). A science-driven market-based instrument for determining the cost of environmental services: A comparison of two catchments in Australia. *Ecological Economics* 64: 61–69

Margules C.R. and Pressey R.L. (2003) Systematic conservation planning. *Nature* 405: 243–253

McCarthy M.A., Parris K.M., van der Ree R., McDonnell M.J., Burgman M.A., Williams N.S.G., McLean N., Harper M.J., Meyer R., Hahs A. and Coates T. (2004). The habitat hectares approach to vegetation assessment: An evaluation and suggestions for improvement. *Ecological Management and Restoration* 5**:**1, 24–27

Nguyen N., Woodward R.T., Matlock M.D., Denzer A. and Selman M. (2006) *A guide to market-based approaches to water quality*. World Resources Institute.

O'Connor P.J and Collard S. (2007). *The River Murray Forest Score: a metric for determining value in bids for revegetation contracts under the River Murray Forest Project*. Department of Water, Land and Biodiversity Conservation, South Australia

O'Connor P.J, Morgan A., and Bond A. (2007) *BushBids: Eastern Mount Lofty Ranges Biodiversity Stewardship Initiative*. South Australian Murray-Darling Basin Natural Resource Management Board

Oliver I. (2002) An expert panel-based approach to the assessment of vegetation condition within the context of biodiversity conservation: Stage 1: the identification of condition indicators *Ecological Indicators* 1: 223–237

Oliver I., Jones H., Schmoldt D.L (2007). Expert panel assessment of attributes for natural variability benchmarks for biodiversity. *Austral Ecology* 32: 453–475

Oliver I, Ede A, Hawes W and Grieve A (2005) The NSW Environmental Services Scheme: Results of the biodiversity benefits index, lessons learned, and the way forward. *Ecological Management and Restoration* 6: 197–205

Parkes D., Newell G. and Cheal D. (2003). Assessing the quality of native vegetation: The 'habitat hectares' approach. *Ecological Management and Restoration* 4 (s1): S29–S38

Reckhow, K.H. (1994). Water quality simulation modeling and uncertainty analysis for risk assessment and decision making. *Ecological Modeling*, *72*, 1–20.

Resnik M.D. (1987). *Choices: an introduction to decision theory*. University of Minnesota Press, Minneapolis, Minnesota, USA.

Rolfe J., McCosker J. and Windle J. (2005) *Establishing east–west landscape linkage in the Southern Desert Uplands Research Reports (Research Report No. 6)*. National Market-based instrument Pilot Project (accessed at http://www.napswq.gov.au/publications/books/mbi/pubs/round1-project18.pdf)

Rowe, W. D. (1977). *The Anatomy of Risk*. New York: John Wiley and Sons.

Stoneham G., Chaudhri V., Ha A., and Strappazzon L. (2003). Auctions for Conservation Contracts: An Empirical Examination of Victoria's Bush Tender Trial. *Australian Journal of Agricultural and Resource Economics* 47: 477–500

Strappazzon L., Ha A., Eigenraam M., Duke C. and Stoneham G.(2003) Efficiency of alternative property right allocations when farmers produce multiple environmental goods under the condition of economies of scope *The Australian Journal of Agricultural and Resource Economics*, Volume 47: 1–27

Tisdall J. (2007). Bringing biophysical models into the economic laboratory: An experimental analysis of sediment trading in Australia. *Ecological Economics*. 60: 584–595

Whitten SM, Coggan A, Reeson A, Gorddard R. (2007). *Putting theory into practice: market failure and market based instruments (MBIs).* Working Paper 2 in the Socio-Economics and the Environment in Discussion CSIRO Working Paper Series Number 2007–02. May 2007. ISSN 1834-5638. 46 pp.(accessed at http://www.csiro.au/resources/SEEDPaper2.html)

Whitten S.M., Khan S., Collins D., Robinson D., Ward J. and Rana T. (2005a). *Tradeable recharge credits in Coleambally Irrigation Area: Report 7 Experiences, lessons and findings*. CSIRO & BDA Group (accessed at http://www.napswq.gov.au/publications/books/mbi/pubs/round1-project33.pdf)

Whitten S.M. and Shelton D. (2005b). *Market for Ecosystem Services in Australia: practical design and case studies*. CSIRO Sustainable Ecosystems. (accessed at http://www.cifor.cgiar.org/pes/publications/pdf_files/Whitten-Australia.pdf)

Windle J. and Rolfe J. (in press). Exploring the efficiencies of using competitive tenders over fixed price grants to protect biodiversity in Australian rangelands. *Land Use Policy* (in press) doi:10.1016/j.landusepol.2007.09.005

Windle J. and Rolfe J. (2007). *Competitive Tenders for Conservation Contracts: A practical guide for Catchment Management Authorities and regional NRM groups*. Central Queensland University

von Winterfeldt, D. and Edwards, W. (1986). *Decision Analysis and Behavioral Research*, Cambridge University Press, New York.

Wu J-J. and Skelton-Groth K. (2002). Targeting conservation efforts in the presence of threshold effects and ecosystem linkages *Ecological Economics* 42: 313–331