Australian Agricultural and Resource Economics Society Conference

Adelaide, February 2010

**Paper Title:** Preliminary principles to guide best practice water quality regulation from an economic perspective

**Paper presented at a mini-symposium “Informing Water Quality Policy in the Great Barrier Reef”**

**Authors:** Anthea Coggan, Stuart Whitten, Tertius Greyling
CSIRO Sustainable Ecosystems
Contact author: Anthea.coggan@csiro.au
Abstract:
Regulatory regimes intended to enforce changes to land use or management impose costs on landholders and governments. Landholder costs comprise changes to capital equipment, changes to crop or enterprise management including direct compliance costs, opportunity costs of lost production, and transaction costs from informing themselves about regulatory requirements, potential compliance strategies and administration associated with implementation of these strategies. Governments must design and implement the regulatory framework along with an appropriate compliance structure and other associated costs. In this paper we apply economic theory, in particular relating to institutional economics and transaction costs, and the degree of heterogeneity landholder net cost to propose a set of principles to guide selection of appropriate regulatory targets and design of implementation frameworks.

Key words: diffuse source pollutants, regulations, economic efficiency, transaction costs, water quality

Acknowledgements:
The research in this paper has been funded by CSIRO through its Water for a Healthy Country Flagship.
1. Introduction

Recent environmental economic contributions to the policy debate of how best to manage diffuse source pollutants, such as those generated by agricultural land uses, have tended to focus on incentive based approaches. However such approaches may not deliver the desired policy outcome or may be regarded as ethically or morally inappropriate. In such settings regulatory approaches may be required either as the primary approach or as part of a package of measures to deliver desired outcomes. Historically environmental economists have tended to concern themselves primarily with issues of efficiency within a rational actor framework and thus focused on a cost benefit analysis of the regulatory approach or the relative payoffs from compliance and therefore the likelihood of compliance. In this paper we extend the traditional approach to include institutional economic considerations of transaction costs and the likely implications of a diffuse source agricultural setting. Our objective is to develop a set of preliminary principles that should inform the design and delivery of regulatory approaches in diffuse source settings.

We define a regulatory approach as involving at least some element of coercion with respect to landholder behaviour. That is, a binding requirement on landholders usually backed by formal legislation which results in a set of regulations. The requirement may involve actions or results. An action oriented regulation will require landholders either undertake or abstain from specific actions. A result oriented regulation may require specific performance standards are achieved.

Our interest is in the nature of the lessons from economics for the design of regulatory approaches to diffuse source pollutant settings in agricultural land use contexts. In particular we are interested in identifying a set of principles that could be employed to promote economic efficiency within a regulatory approach. We are also interested in the interaction between regulatory design and the characteristics of diffuse source settings. In particular we seek to identify whether there are specific characteristics of diffuse source agricultural settings that should be considered. We consider the implications of our preliminary set of principles for water quality regulations with emphasis on sugar cane production in the Great Barrier Reef (GBR) catchments.
The paper is structured as follows. In the next section we present some background information about when and where regulatory approaches are considered with an emphasis on why one may consider a regulatory approach. In the third section we briefly summarise several aspects of regulatory design from an economic perspective and propose a conceptual model for identifying these factors in the field. In section 4 we qualitatively apply the model to water quality in the GBR for cane. We conclude the paper with a short section summarising conclusions and describing the research/information needed to formally consider the approach.

2. Background
Across Australia, 63% of land is held under private ownership or management (ABS 2002). Governments are often required to consider ways in which changes to private land management or use can be achieved in order to deliver on public objectives. In most cases the sources of either the damage to environmental assets, or from which beneficial contributions to public environmental objectives can be sourced, can be described using aspects of a diffuse source pollutant problem. A diffuse source pollutant is defined by D’Arcy and Frost (2001) as “non-point source contamination, such as sheet runoff from fields or seepage into groundwater, as well as a multiplicity of dispersed, often individually minor point sources, such as surface water drains in urban areas, field drains, and ferruginous springs associated with abandoned mine drainage”.

The key characteristics of a diffuse source for the purposes of policy design in this paper are threefold:

- There are usually a large number of heterogeneous sources;
- The diffuse nature of the source makes it difficult to measure or monitor; and
- There is often (but not always) a pre-existing de facto right to existing activities and practices.
Examples include water quality, salinity, biodiversity, and carbon implications of land use decisions amongst others.

There are two broad ways that a change to landholder behaviour with respect to the diffuse source can be motivated: 1) through the application of voluntary instruments;
or 2) through the implementation of required actions or regulation. Each of these instruments is described briefly below.

Voluntary instruments are those instruments that generate change by private landholders through a voluntary shift in land management or use. Voluntary instruments include non financial instruments such as moral suasion and the provision of information, advice and support or financial based incentive mechanisms. Moral suasion essentially relies on a landholder’s sense of moral and ethical standards and is enforced primarily through social or peer pressures. Suasive approaches encourage landholders to behave as a “good citizen” and conform to changes to the accepted norms of behaviour. A related non-financial instrument is the provision of information advice and support. This may include the provision of specific information about management and production techniques to landholders, or the facilitation of knowledge transfer through participation in workshops and knowledge networks. The voluntary adoption of best management practises (BMP) to deliver land management change which is popular with industry groups and government is a form of information, support and advice. Non-financial approaches are unlikely to be effective in isolation where practice change incurs a large private financial cost. Incentive based instruments supporting voluntary change may be used in such settings to encourage adoption. These are often also termed financial-incentives and include cost-share arrangements, grants, auctions and tenders as well as less direct methods such as taxes, subsidies and rate rebates.

An alternative approach is a non-voluntary approach which requires landholders to comply by changing their behaviour: usually through use of government’s coercive powers but sometimes through other options such as industry supply requirements. For required actions, compliance is supported by legal or financial penalties. The focus in this paper is on approaches involving use of government’s coercive powers. Examples include provisions under the Environment Protection and Biodiversity Conservation Act 1999, native vegetation clearing laws (Queensland Vegetation Management Act 1999) water pollution regulations (Queensland Great Barrier Reef Protection Amendment Bill 2009) and so on.
There is an extensive literature which outlines when you would use one approach over another to generate desired land management or practice change (see Young et al. 1996; Stoneham et al. 2000; ABARE 2001; Comerford 2004; Pannell 2008). This is summarised and extended in Coggan and Whitten (2008) from which several themes emerge. First, the overarching consideration from a perspective of economic efficiency should be that the benefits of the instrument outweigh the costs.\textsuperscript{1} Next is an emphasis on understanding the complexity of the setting in which the approach is considered and tailoring the approach to deliver in that setting. Third, supporting the overall economic efficiency consideration is attention to the transaction or hidden costs involved in the approach. Finally, an appreciation that pragmatic decisions will need to be made in applied settings and that instrument design is far from a perfect science: or in economic parlance, first best solutions are rarely possible.

In considering the economic efficiency aspects of the approach there are three key elements to the calculation of instrument cost that may require particular attention when considering regulatory approaches:

1) The cost employing a particular instrument such that it achieves the specified objectives including transaction costs;

2) The inclusion of landholder costs associated with practice change in any efficiency analysis; and

3) The nature of the environmental problem and the likelihood that the instrument will deliver the desired outcome compared to the cost to the government and society of not achieving the objectives.

Government actions through voluntary or non-voluntary instruments which are intended to deliver behavioural change by private landholders are not costless to create or use. Instrument development and use incur significant transaction costs to both the developer and administrator of the instrument and to the parties (in this case private landholders) who later engage with the instrument. They may also incorporate significant monetary transfers from one part of society to another in the case of financial incentives. Transaction costs are the cost of resources used to define,

\textsuperscript{1} Community/political acceptability of different instrument types is also important to instrument selection. For some impacts the community may not feel that it is acceptable to pay for a change and instead may demand that those that may have a potential impact on public good be taxed for this impact. This is often discussed in terms of beneficiary versus polluter pay principles.
establish, maintain and transfer property rights (McCann et al. 2005). Studies of environmental policies have found that transaction costs to the public agency creating and implementing the policy can be up to 40% of the total policy cost (Howitt 1994; McCann and Easter 2000; Falconer et al. 2001). There is also some data available on the transaction costs of public policy to private parties. Such transaction costs range from 15 to 50% of total costs to private parties (Falconer 2000; Falconer and Saunders 2002; Vatn et al. 2002; Rorstad 2007; Kuperan et al. 2008; Mettepenningen et al. 2009).²

When transaction costs are taken into account it is clear that some policies will incur lower design, implementation and administration costs than others. In some cases a well designed and targeted regulation may have lower direct and indirect costs and equal effectiveness credentials to other (more flexible) policies. Similarly there may be cases where failure to consider ongoing transaction costs, such as those associated with an effective monitoring and enforcement regime, may lead to a false conclusion of the economic efficiency of a regulatory approach.

A false economic argument often employed to justify regulation is the relatively low cost to government because it avoids financial payment to landholders to deliver practice change. Failure to consider the potential that there are real costs to landholders of changing practices may lead to landholders, local communities, or the nation being worse off in terms of total economic welfare, even though the desired environmental outcome has been achieved (or even worse, imposing substantial costs while failing to deliver the desired environmental outcome). Economic efficiency requires attention to the full range of costs and benefits from the mechanism considered, not just those to government.

The likely effectiveness of an instrument should also be taken into account in its selection. This is particularly the case when the system of concern is approaching hard and irreversible thresholds or when the only way to achieve the policy objective is to

² The caveat on analyzing costs of policies is that the costs will change depending on the age and stage of the policy. Transaction costs will usually be higher at the start of a policy when there is much information collection and learning about the policy. For example, Falconer et al. (2001) found that transaction costs of Agri-Environmental Schemes in Europe began at 112% and eventually settled at 18% of total cost. This is why many analysts (McCann et al. 2005; Buitelaar 2007) concentrate on the drivers of transaction costs when thinking about policy effectiveness and refinement.
generate change from all stakeholders. Examples of this may include management of invasive pest species and diseases as well as environmental pollution. In such settings higher cost instruments which reduce the likelihood that such thresholds would be breached may be preferred. The caveat is that economic benefits would need to remain for the instrument to be economically efficient.

3. Identifying a preliminary set of principles to guide regulatory approaches to diffuse source pollutants

Our primary aim in this paper is to identify a set of preliminary principles to guide the design of regulatory approaches to diffuse source water quality issues. In this section we combine the key findings across a range of literature to inform regulatory design for diffuse source water quality problems. Our discussion is broadly divided into four overlapping parts. In the first we relate the water quality problem as a diffuse source problem to the concepts employed in risk based environmental regulatory approaches. Next we describe the implications of the heterogeneous nature of diffuse source agricultural settings for an economically efficient regulatory approach. We then broaden our consideration to include some of the lessons from behavioural and institutional economics: first with respect to the likely behavioural response of landholders to a new regulatory environment; and second in terms of relating transaction costs and institutional constraints to diffuse source issues. We conclude by setting out a preliminary set of principles to guide regulatory design.

Risk-based approaches to regulation

Water quality issues are commonly caused by a range of pollutants and a range of sources. Even where single pollutants are involved there are usually a range of differential impacts caused by the pollutant with different likelihoods and degrees of damage. That is, the potential consequences of pollutants are likely to vary, especially where a range of environmental pollutants are emitted. Regulatory bodies in different jurisdictions, such as the Environment Agency in the UK and the Environmental Protection Agency in the USA, have turned to risk-based approaches (Hornstein 1992; Pollard et al. 2002). These approaches have traditionally been used in industries facing high levels of risk such as the petroleum and nuclear energy industries. In order to address the need for differentiation between the risks posed
across a range of sources and environmental impacts, risk-based regulatory approaches assign the limited regulatory resources to the highest risks and worst performers. The intention is that prioritisation of regulatory action should deliver public sector efficiency improvements and reduce regulatory burdens.

The methods employed by the Environment Agency in the UK include the use of a tiered approach in risk assessment, risk rating systems for risk management resource-planning, and the use of risk scenarios (Pollard et al. 2002). This approach seeks to release ‘regulatory attention’ from low-risk activities to be re-assigned to high-risk activities. Figure 1 illustrates the savings in terms of resources (time and/or human resources) or reduction in risk that can be attained via risk based resource planning which effectively lowers the overall risk profile.

![Figure 1: Risk profiles for conventional resource planning versus risk-based resource planning](image)

Gouldson et al. (2009) highlights, however, that risk-based approaches cannot be applied uniformly across the board. Rather they must account for the different nature
of diverse environmental problems (climate change, flooding, radioactive waste disposal and so forth), the different types of risk dealt with at different levels within organisations, and the types of risk encountered across different sectors or geographical areas. Some types of risk are better suited to a formalised, consistent approach while others are difficult to measure in a consistent and formal way.

In practice, risk is usually determined by the consideration of the *likelihood* of an event occurring coupled with the *severity* of the outcome should it occur. For example, in the context of risk-based safety criteria for nuclear industry, this translates to the probability of exposure to a given dose and the probability that the dose will result in “death or a serious health effect” (Rothstein et al. 2006).

The combination of thinking from Pollard et al. (2002), Gouldson et al. (2009) and Rothstein et al. (2006) can be combined to suggest that there are two components to environmental risk for diffuse source water pollution: 1) the risk to the environment posed by the pollutant; and 2) the risk that the pollutant will be delivered from a specific agricultural activity. For example, there might be high risk of environmental damage associated with accidental release of a chemical, but if the chemical is stored and used appropriately the risk of release actually occurring may be low. The combination of these two factors of risk may result in consequential risk varying according to the likelihood of the pollutant being delivered from agricultural land uses and the likely damage that would result from release. As an example risk could be quantified as described in Table 1 with resultant combinations of risk being assigned different regulatory requirements.

<table>
<thead>
<tr>
<th>Severity of damage when potential pollutant is present</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>L</strong></td>
</tr>
<tr>
<td><strong>L</strong></td>
</tr>
<tr>
<td><strong>M</strong></td>
</tr>
<tr>
<td><strong>H</strong></td>
</tr>
</tbody>
</table>

Table 1: Factors to be considered in an environmental risk matrix
Implications of landholder heterogeneity

The heterogeneous nature of diffuse source pollutants for environmental regulation has already been alluded to above but is likely to be much broader than can be encompassed by a single dimension assessment of the likelihood of the pollutant being delivered. A principal characteristic of diffuse source pollution is the heterogeneous nature of the sources; particularly across agricultural land uses. By definition this class of pollution is spread over large areas and composed of numerous small sources, each of which discharge variable quantities of different pollutants dependent on a wide range of factors. Enterprise heterogeneity is not limited to biophysical factors such as land use, soil type, weather conditions and so on, but extends to the size and business structure of the farming operation, landholder goals and other social and economic factors.

The factors generating landholder heterogeneity will have direct and indirect impacts on both the risk of pollutants exiting the farm boundary and also on the relative costs (or benefits) and practicalities of changing management. Evidence suggests that even within relatively uniform industries such as cane production, substantial variation will remain (Van Grieken 2009, Emtage 2009). For example, Carpentier, Bosch and Batie (1998) developed a conceptual framework to estimate cost saving via the application of spatial information in a case study for dairy farms in the USA. They concluded that the heterogeneous nature of diffuse source pollution meant that achieving a 40% reduction in Nitrogen runoff would be 75% lower using a targeted performance standard over a uniform performance approach. In both cases the farming practices to reduce runoff were similar (strip-cropping, manure storage, manure incorporation and shifting timing of manure application). However, under the targeted performance standard fewer farms (but posing a higher risk of runoff) were targeted in comparison to the uniform performance standard.

There is agreement within the literature that a single inflexible regulatory instrument is unlikely to satisfactorily address diffuse source pollution (Gunningham & Grabosky 1998; Goulder & Parry 2008). The nature of nonpoint source pollution, described by Dowd, Press & Los Huertos (2008) as “diffuse, prone to discharge in impulses and difficult to pin on any single pollution event or source”, makes the uniform use of a single regulatory instrument unlikely to succeed. Russell and Powell (1999) note the
benefits of a flexible approach in such environments which would allow for a heterogeneous response on the part of landholders taking into account their own situation. Flexibility also facilitates dynamic efficiency and innovation amongst those that are being regulated.

**Incorporating landholder behaviour into an effective regulatory approach**

Environmental policy design has increasingly moved beyond a model predicated on optimising agents with perfect information. Recent developments have emphasised the potential for institutional and behavioural economic theories to contribute to policy design. Both theories are relevant to the likely landholder response to regulation. The institutional approach identifies the importance of both the perceived and actual property rights held by landholders (prior to and after regulation) and the institutional characteristics of each of the parties involved. Behavioural economics identifies the importance of considering landholders' behavioural response to the changed setting imposed through regulations.

Though informed by social and psychological theory rather than behavioural economics as such, Ayres and Braithwaite’s (1992) concept of regulatory pyramids (Figure 2) provides a useful framework for the consideration of regulatory approaches. Their approach is intended to enhance the efficiency and effectiveness of regulatory approaches where there will be a large number of participants. In the pyramid in Figure 2 the majority of pollutant sources are located at the base of the pyramid. These regulated entities are assumed to generally comply with regulations and are willing to work with the regulator. At the other extreme at the top of the pyramid one finds entities who are wilfully non-compliant, in other words those sources that pose the greatest risk of non-compliance (but not necessarily of environmental risk). Enforcement costs are highest at the top of the pyramid and lowest at the base.

The use of ‘strong’ regulatory compliance enforcement which apply the greatest pressure, such as revocation of licences, large financial penalties or criminal prosecution, are reserved for those sources located at the top of the pyramid which pose the greatest risk of non-compliance (Ayres & Braithwaite 1992, pp. 35-39). The eventual option of a ‘strong’ approach also creates downward pressure on the more
compliant sources lower down in the pyramid without the need to actually use these ‘forceful’ instruments at these lower levels. The mere threat posed by their possible use is usually sufficient to persuade most sources at the lower levels to comply with regulatory targets. Instead less intrusive instruments are used at this level and ultimately through the encouragement of dialogue and the provision of information the aim is to promote self-regulation at the base of the pyramid.

![Regulatory compliance pyramid](image)

**Figure 2: Regulatory compliance pyramid**
*Source: Ayres and Braithwaite (1992 p. 35); Braithwaite (2006).*

The model presented by Ayres and Braithwaite (1992) provides some lessons for regulatory approaches to managing diffuse source water quality issues. Primarily that to be cost effective a gradated approach to policy should be sought. While Ayres and Braithwaite focus on compliance, based on the discussion provided by Pollard et al. (2002) (and others) on environmental risk, we believe that for water quality the focus of gradation can be expanded to directly relate to the degree of environmental risk. In this case the gradation occurs around the strength of the regulation and the compliance enforcement to these regulations (such as that for Ayres and Braithwaite).

A generic example of a gradated environmental risk approach is described in Table 2. A risk evaluation such as that summarised in Table 1 would determine if a landholder
had a low, medium or high environmental risk. The strength of the regulatory structure applied would depend on the risk category. Low risk landholders may only be subject to suasive approaches, low to medium landholders may be required to demonstrate that risks are managed and so on. Compliance approaches should also be informed by Ayres and Braithwaite (1992); even at high levels of environmental risk the full force of the law should only be used as a last resort on those who are wilfully non-compliant.

A gradated approach will also prove useful where new regulatory approaches require substantial change to accepted activities and processes. In agricultural communities there tends to be cultural resistance to regulation, especially if producers are organised into political lobbies (Gunningham & Grabosky 1998, p. 393). Therefore the design, implementation, monitoring and control of regulation can be quite demanding in terms of monitoring and policing environmental targets, building networks for the diffusion of information and capacity building, and so forth. To ease implementation, initial implementation may focus on suasive approaches targeting landholders with high environmental risk and then progressively extended to other risk categories. Compliance requirements could then be progressively raised to the desired level for each risk class. For example, high risk landholders would be the first targeted in information campaigns, to demonstrate risks are managed, the first subject to risk auditing or direct sanctions and so on. Given that information is costly to attain and will take time to spread through various local networks, penalties for lack of compliance may be moderated in the initial stages in order to support suasive approaches but ramped up after a certain time.
Table 2: Pollution risk focussed gradation of regulation and compliance enforcement*

<table>
<thead>
<tr>
<th>Gradated compliance regime</th>
<th>Risk associated with activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
</tr>
<tr>
<td>Primarily suasive approaches</td>
<td></td>
</tr>
<tr>
<td>Demonstrate risks are managed</td>
<td></td>
</tr>
<tr>
<td>Make-good provisions and formal directives</td>
<td></td>
</tr>
<tr>
<td>Fines and non-criminal sanctions</td>
<td></td>
</tr>
<tr>
<td>Criminal prosecution</td>
<td></td>
</tr>
</tbody>
</table>

Transaction costs of regulatory approaches

The cost of creating (McCann et al. 2005; Buitelaar 2007) and then enacting, implementing and enforcing a policy (Thompson 1999) is often significant for environmental regulation. Significant transaction costs may extend to those that are regulated (Thompson 1999; Vatn et al. 2002; Rorstad 2007; Kuperan et al. 2008; and Mettepenningen et al. 2009). While there are multiple sources of transaction costs, three sources are likely to be particularly important when considering regulatory approaches to diffuse source pollutants:

1. The costs of gathering sufficient knowledge to identify environmental risk per Table 1 given both complexity of environmental impacts and the diffuse source setting;
2. The cost of administering a regulation across a large number of heterogeneous diffuse sources. In particular the cost of monitoring both water quality (see for example Collins & McGonigle 2008) and landholders will be high;
3. The potential cost of compliance particularly where there is an emphasis on demonstrating compliance or a requirement for costly changes to management.

Underpinning knowledge is critical to reducing administration and compliance costs. Without a sound understanding of the sources and consequences of environmental risk there is little ability to target or apply a gradated approach to regulation. Knowledge itself is costly however, so there will always be a trade-off between the costs of gathering the knowledge necessary to refine regulatory (or indeed any
approach) and the potential benefits from a more efficient or effective regulatory design.

Transaction costs highlight that the regulatory approach used (within the graded response) must match with the capacity of the regulator. Keeping track of agricultural diffuse emissions is very costly (Goulder & Parry 2008; Brady 2003). The scope and scale of monitoring trace diffuse source pollution to the individual parcels of land would be prohibitive as the costs of monitoring could well exceed the budgets of many local regulatory agencies (Dowd, Press & Los Huertos 2008; Young & Karkoski 2000). As a result various alternative approaches to the traditional command and control regulation have been considered to circumvent the obstacle posed by the cost of diffuse source pollution monitoring. For example, Romstad (2003) suggests holding landholders accountable collectively as a group or team via ambient environmental monitoring. His model sees the team enter into a contract with the regulator which is favourable or unfavourable depending on their performance relative to a targeted level. In effect it is assumed that monitoring of ambient quality is possible locally but that distinction between individual farmer’s emissions is impossible (or at least not cost-effective). The model relies on the informational structure of agricultural communities where the different agents may possess information on the actions of other agents such as knowledge of farming practises, efforts to reduce diffuse source pollution, and accidents such as spillages that could increase nutrient loads. This approach, which seeks to return to direct regulatory measures but in a way that is more feasible, makes it easier to achieve precision in diffuse source pollution regulation and leverage local information in order to reduce the costs of compliance for individual landholders.

Ayers and Braithwaite (1992, p. 102) propose ‘enforced self-regulation’ as an alternative that would offer the regulator greater security whilst shifting much of the administrative costs to the producers. The producers generate their own rules which are subsequently ratified by the government and a breach of these rules is considered an offense. Internal compliance groups are established to monitor and recommend disciplinary action. Also, the compliance groups are statutorily compelled to report any breaches to the relevant government agency. The regulatory agency’s role is to determine that the rules comply with government policy, to ensure that the
compliance group acts independently from individual landholders, to audit the performance of the compliance group, to perform occasional spot inspections and to launch prosecutions. However it should be noted that to the extent that administrative costs are merely shifted off the regulatory agencies books rather than reduced, self-regulation is cost shifting not cost reducing).

In the case of diffuse source pollution in agricultural settings strong producer groups offer a potential mode of implementing a self-regulation approach. In a case study for California’s Central Coast farmers, Dowd, Press & Los Huertos (2008) present an example of how this type of model can actually be implemented. The regulator (The Central Coast Regional Water Quality Control Board) used a cooperative approach based on a combination of voluntary actions and regulation. The farming community was presented with an opportunity to independently address diffuse source pollution through monitoring and adoption of BMPs, but backed up by the threat of regulatory action in the event of inaction. This meant that the regulator had to shoulder a smaller administrative and cost burden because environmental monitoring was left to the farmers. Following strong objection from the farming community to the monitoring costs, a cooperative monitoring programme was introduced as a more cost-effective means to monitor water quality on the main tributaries of the Central Coast catchments. This approach gave the farmers the flexibility of choice to make farm management changes in a way that is best suited to their particular circumstances, reduces the environmental risk through the threat of regulatory action, and offers both the regulator and the farmer a more cost-effective way to address the need for better environmental management.

The form of compliance required will have important consequences for landholder costs. We have previously noted the potential for targeted approaches to be more cost effective than uniform approaches. We have also noted the potential for a gradated approach to reduce the costs of compliances. Russell and Powell (1999) identify two further interactions that should be considered with respect to the costs of complying with regulations according to whether approaches specify ‘what outcome is to be achieved’ or ‘how the outcome is to be achieved’ or both. Specifying a particular outcome requires monitoring of achievement (e.g. pollutant exports from farms).
Specifying how it is to be achieved requires a focus on process instead of outcome. Focusing on process also runs the risk of slowing or precluding innovation.

**Capacity and constraints in implementing regulatory approaches**

Any policy instrument will require a specific set of capacities, skills and institutional structures for success. There are three primary areas where there may be significant impediments to regulatory approaches: limited capacity, misalignment or absent capacity; and inadequate or misaligned institutional support.

In diffuse source agricultural settings where environmental management is a relatively new need, there are often a range of skill shortages, both within existing entities (such as natural resource management authorities and government agencies) and for any new entrant. The required set may differ substantively for regulatory approaches compared to alternative instruments and may not be present where regulatory approaches have not previously been applied. Furthermore, the implementation of information hungry risk-based approaches can place large cost pressure on already resource-constrained agencies and regulators. Complex and costly monitoring will also need to be considered per Goulder and Parry (2008) as discussed above. Similarly, Gunningham and Grabosky (1998) notes that agricultural industries are typically geographically dispersed and there may not be strong formal or informal networks to aid in information distribution. This may be exacerbated by local government or regulatory departments not having the expertise needed to deal with risk-based approaches to the same extent as the lead regulator.

Furthermore, regulatory approaches usually derive their power from the potential use of the coercive powers of government which require a very specific set of enforcement skills. Failure to adequately align enforcement with agency powers or to resource enforcement activities may lead to landholders ignoring regulations or to patchy and inequitable enforcement of laws.

There may also be obstacles to the implementation of risk-based approaches in existing regulatory frameworks (for example some legislation may preclude the use of risk-based methods) (Rothstein et al. 2006). Similarly, consideration will need to be
given to the mode of enforcement and the permissions and other legal requirements that will be required – such as who is authorised to exercise any right of entry for inspection purposes and so on. Institutional misalignment can be particularly costly if there are no agencies active in the region to which powers can be appropriately delegated; often the case in remote areas or where non-government organisations are the primary mode of delivery for existing approaches.

A preliminary set of principles

From the preceding discussion we have identified ten principles to help guide regulatory approaches to water quality in diffuse agricultural settings. These principles are primarily informed by economic efficiency considerations augmented by the likely science information and political limitations of the diffuse source agricultural settings. Ignoring the attitudes of those being regulated or the equity consequences of a particular policy often results in strong resistance, lobbying and the ultimate failure of a proposed policy. Similarly if the policy cannot function within the scientific limitations posed by monitoring and control of diffuse source pollution it is unlikely to be effective. We divide the principles between four upper level guiding considerations and six supporting principles according to their priority for consideration.

Guiding principles:

1. The regulatory structure should be scaled in accordance with the severity of damage from a particular pollutant.
2. The regulatory structure should match the risk of pollutant being generated. That is the likelihood that the particular activity will actually release pollutants.
3. The greater the heterogeneity amongst landholders the greater the potential economic benefits of a flexible approach.
4. Consider all costs of the regulation including to regulatory agencies, landholders and others for any given regulatory approach.

Supporting principles:
5. Incorporate a graduated response to (from least forceful to most forceful) to minimise enforcement cost to landholders and agencies. This response should be aligned with the environmental and pollutant generation risks.

6. Incorporate a graduated implementation process through time where significant change to attitudes or actions is required.

7. Where possible foster dynamic efficiency and innovation within the regulatory approach.

8. Account for the cost and complexity of monitoring pollutants or behaviour in regulatory design.

9. The design should be consistent with the regulator’s capacity to engage in the necessary aspects of implementation and administration including communication, auditing or monitoring and enforcement.

10. Incorporate complementary approaches where possible and consistent with risk to build in local information, reduce costs and empower participants including industry oriented self regulatory approaches, industry performance standards, and peer monitoring.

4. **A case study: Water quality in the Great Barrier Reef**

The world heritage listed Great Barrier Reef (GBR) is situated adjacent to the Queensland coast, consisting of an archipelagic complex of over 3000 reefs covering an area of approximately 350 000 square kilometres, the GBR is the largest reef system in the world (Haynes and Michalek-Wagner, 2000) (Figure 3).

Water quality in the GBR lagoon has declined significantly over the last 140 years (Moss et al. 1992; Neil 1997). The primary driver has been an increase in diffuse source pollution activities from agricultural and urban activities in the GBR catchments with lesser contributions from mining (Brodie et al. 2003). For example, sediment loads due to soil erosion have increased 3-4 fold over the last 140 years (Moss et al. 1992; Neil 1997; Productivity Commission 2003).
Brodie et al. (2003) estimated that 15 million tonnes of suspended sediment is exported to the GBR each year. Total nutrient influx to reef waters (principally nitrogen and phosphorus) has also increased by 30% mainly since agricultural expansion in the 1970’s (Brodie 1997; Pulsford 1996). Brodie et al. (2003) estimate that 77,000 tonnes of Nitrogen and 11,000 tonnes of Phosphorous is discharged to the GBR from the land each year, this is 3-5 times the pre-European load. The 2008 scientific consensus (Qld Dept of Premier and Cabinet 2008) stated that there are known correlations between river discharged material and water quality in the GBR lagoon; specifically phytoplankton biomass and pesticide concentrations in the GBR lagoon are directly correlated with river nutrient and pesticide loads respectively.

On the first of January 2010 the Great Barrier Reef Protection Amendment Act 2009 (GBRPA Act 2009) was introduced to help manage the impact of diffuse source polluters to the GBR lagoon and to achieve the water quality reduction targets
specified by the Qld State Government under its Reef Rescue Policy.\textsuperscript{3,4} Sugar cane is regarded as a major source agricultural pollutants and the \textit{GBRPA Act 2009} is intended to reduced pollutants from cane in the wet tropics and Mackay Whitsundays along with grazing in the Burdekin dry tropics (BDT) catchment. The \textit{GBRMPA Act 2009} requires landholders to determine crop requirements for nitrogen and phosphorous before application as part of a requirement to compile and comply with an environmental risk management plan. The regulation initially applies to landholders who are growing sugar cane on an area greater than 70 ha in the wet tropics catchment; Cattle grazing on more than 2000 ha in the BDT; or where the person is given a direction – that is a direction may be given when an ERMP is considered necessary to improve water quality being released from an agricultural property or because an environmentally relevant activity (ERA) is causing unlawful environmental harm. The direction must contain obligations and these must be complied with.

\textit{Principle 1: The regulatory structure should be scaled in accordance with the severity of damage from a particular pollutant.}

The dominant pollutant exported from sugar cane production in the GBR results from application of nitrogen (N) to crops. There is increasing literature and evidence to suggest that application of nitrogen is hazardous to the GBR. For example, post-European increases in the terrestrial nutrient load of river discharges have been cited as a possible cause of reductions in coral growth and a shift in the relative abundance and composition of corals and algae, particularly in near shore fringing reefs (Udy et al. 1999; Van Woesik et al. 1999; and Fabricius et al., 2005). Further, Brodie et al. (2001) note that an increase in nitrogen and phosphorous into the GBR lagoon is resulting in “impaired health and productivity of marine organisms and loss of habitats and associated aquatic connectivity”. Brodie et al. (2001) note that the long-term effects of eutrophication on inshore coral reefs of the GBR are only just becoming evident after a decade of monitoring. In the most completely studied reef province, the Whitsundays, an eutrophication gradient from the Proserpine River has been

\textsuperscript{3} The Reef Rescue targets are as follows: reduce discharge of dissolved nutrients and chemicals from agricultural lands to the GBR lagoon by 25%; reduce the discharge of sediment and nutrients from ag lands to the GBR lagoon by 10%; and in reef catchments increase the adoption of improved land management practice by at least 30% of agricultural land managers.

\textsuperscript{4} This paper is in no way intended as a review of the \textit{GBRPA Act 2009} though we acknowledge our motivation is in part derived from the recent regulatory activity in this area.
identified with coral reef growth, recruitment and reproductive success being severely reduced (Van Woesik et al. 1999).

According to scientific consensus, while nitrogen is the pollutant of most concern, phosphorous, sediments and a variety of agricultural chemicals are also exported from agricultural land uses and related (with varying levels of analysis and confidence) to environmental damage on the GBR.

The GBR covers an enormous area and nitrogen (and other pollutant) concentrations are far from uniform. Hence, a major challenge in designing effective regulation is to distinguish the degree to which nitrogen (or other pollutants) is a risk in a particular location and whether there is sufficient variability to justify differential regulatory levels or approaches.

Principle 2: The regulatory structure should match the risk of a pollutant being generated.

There is strong evidence to suggest that the application of nitrogen to sugar crops will result in increases in nitrogen in GBR waters which will cause environmental harm. For example, Brodie et al. (2001) highlight that rivers running through catchments dominated by agriculture have concentrations of dissolved inorganic nitrogen in flood flow of 30 times that of rivers with undeveloped catchments. Concentrations of pollutants such as DIN in river plumes during flood events are typically 10-50 times higher than ambient concentrations in non-flood periods and exceed the effects levels for biological actions on corals seagrasses and algae (Brodie et al. 2001).

It is clear that nitrogen application on crops can result in water pollution in the GBR and that this water pollution can have damaging impacts on the reef environs but it should not be assumed that all cane farmers in all parts of the reef will export pollutants. Contributions of nitrogen to the reef vary across catchments due to catchment size, volume, intensity and timing of runoff, and the land uses, soil types, vegetation cover and terrain producing the pollutant (JCU 2003). Variation in potential for pollutant delivery also varies between individual landholders (and possible between individual paddocks). The variation of risk between landholders is largely due to variation in the quantity of nitrogen applied and the application method.
Van Grieken (2009) uses 4 different nitrogen application methods in cane water quality modelling. The worst application is a set kg/ha while the best is based on an understanding of the exact nitrogen replacement required by individual paddocks at that point in time (see Thorburn et al. 2009a and b).

The potential variability in source risk suggests that a targeted approach is warranted whereby the variability in landholder risk of pollutant export is taken into account in regulatory design. Furthermore, there is likely to be considerable heterogeneity across a variety of landholder attributes, including ability to moderate the risk or their activities and the costs (or benefits) of doing so amongst other factors. This variability brings us to principles 5 and 7 – which suggest a gradated regulatory response according to aggregate environmental risk and to incorporate the potential for dynamic efficiency and innovation where possible. Use of a gradated may involve the use of a range of complementary approaches to build on local information, reduce costs and empower participants per principle 10. The gradated approach allows a clear focus on reducing the risk of pollutant export posed in high environmental risk settings rather than dispersing resources across lesser priorities.

The pollution risk variance can also be used to inform the order in which implementation of a regulation is considered and its compliance enforcement (principle 6). That is, a gradated approach through time would initially target high risk landholders, land uses or locations followed by those identified to pose lesser risks. Similarly high risk sources may eventually face more severe consequences for lack of compliance. Compliance enforcement would also need to be gradated through time to allow time for learning and adjustment and to reduce the potential for honest mistakes to be harshly penalised.

While this paper is not intended as a review of the GBRPA ACT 2009, we note that the act does employ a risk based approach. The 2010 Reef Protection Regulation targets sugar cane growers in the catchments of Wet Tropics and Mackay Whitsundays and graziers of the Burdekin dry tropics catchment; all of whom are identified as high priority pollutant sources. The regulation requires these targeted landholders to develop and comply with an environmental risk management plan.
which includes specific requirements covering determining crop requirements for nitrogen and phosphorus before application.

**Principle 3: The greater the heterogeneity amongst landholders the greater the benefit of a flexible approach**

Heterogeneity refers to the differences amongst landholders usually in relation to preferences, resources and production opportunities (Whitten et al. 2009a). Within a landscape there are three potential sources of heterogeneity: 1) biophysical; 2) management options and technologies; and 3) economic and social factors (Reeson et al. 2009). Biophysical heterogeneity reflects the different soil, rainfall, slope and other characteristics which will affect the level of risk that they pose for any given land use. There are also a variety of management options to generate an outcome (management heterogeneity) which interact with the biophysical heterogeneity to deliver different pollutant export outcomes and different degrees of effectiveness in avoiding impacts. Finally, landholders differ in terms of the cost that actions hold for them and in their financial resources or other capabilities to change management – this is often due to their enterprise viability and history, and the values that the landholders hold for their property.

There is abundant heterogeneity exhibited in the GBR catchments, even for particular land uses in particular regions. Even relatively small areas exhibit variation in soils and landholder management. Van Grieken (2009) has identified significant variation in farm enterprise structure within a single region by classifying landholders between:

1. Large sugar farms structured to maximise profits and a focus on improving management to that end;
2. Medium sugar cane farms with a tendency to face significant tradeoffs between labour and leisure time which was seen as a constraint to changing management;
3. Small cane farmers which are both capital and labour constrained but with farmers getting most income from off-farm sources;
4. Mixed crop farmers with multiple agricultural enterprises and hence management considerations complicating engagement and adjustment.

Emtage (2009) also assesses landholders in the wet tropics grouping landholders according to their values rather than scale or type of operation describing five
landholder groups: 1) well connected and progressive; 2) production orientated; 3) multiple objective; 4) concerned but not engaged; and 5) disconnected and conservative. Emtage focused on landholder interest and engagement in NRM (this spans across agriculturally motivated landholders and those landholders with lifestyle and hobby farm objectives) as a vehicle to changing management. Three of the five groups are primarily focussed on agricultural production (groups 1, 2 and 3) but much diversity remained between production focussed groups when it came to interest in natural resource management and adoption of recommended land management and production practices. Emtage notes that approximately half of the members in the two remaining groups – concerned but not engaged (20% of sample) and disconnected and conservative (40% of sample) – did not conduct commercial agriculture.

The degree of diversity in across enterprises emphasises the potential benefits from incorporating flexibility in the way in which landholders are able to respond to regulatory approaches. Understanding the nature of landholder differences can also assist in incorporating effective strategies to encourage dynamic efficiency and innovation (principle 7) and to understand when the employment of complementary approaches my be practical and effective (sub principle 10). For example, Emtage (2009) suggests that a mixed communication strategy including both industry and environmental groups and describing the economic and environmental benefits of desired changes to management.

**Principle 4: Consider all costs of the regulation including to regulatory agencies, landholders and others for any given regulatory approach**

Economically efficient regulations should be designed with the objective of maximising the net benefits to society. This will generally require that the transaction costs to all parties are minimised without compromising on policy effectiveness (McCann et al. 2005). The degree of heterogeneity amongst landholders, both in terms of pollutant delivery risk and in the biophysical and enterprise diversity suggests that there a focus on matching regulatory response to environmental risk while incorporating flexible response is critical to economic efficiency (principle 5). Further opportunities to minimise the costs of new regulatory approaches may result from employing a staged implementation approach, particularly where management
changes may be costly, difficult to communicate or complex to implement (principle 6) and which are supported by complementary approaches (principle 10).

A particular cost and difficulty in implementing water quality regulations in diffuse settings is the cost and difficulty of effectively monitoring pollutants or behaviour and particular attention should be given to the tradeoffs in this space (principle 8). Where possible complementary approaches (principle 11) should be used to either augment accuracy of monitoring where resources are scarce or as alternative approaches. The cost of such approaches should always be included in any analysis of economic efficiency.

It is also important to ensure that the regulation chosen and its implementation pathway are consistent with capacity (principle 9) and where they are inadequate the costs of investing in them are included. While initial costs of enhancing capacity may be high, the resultant enhancement in regulatory effectiveness could result in lower costs in the long term. Capacity can almost always be enhanced by collaborating with other organisations in the region. In the GBR there are quite a number of regional catchment groups and industry groups who have extensive experience working in the region, working with landholders and possibly also in monitoring and enforcement. Linking in with these groups may be an ideal way to enhance capacity at minimal regulator transaction costs. Regulation design is not a clean science, there is always likely to be conflicts and tradeoffs in design and implementation. Accordingly, collaboration may not always result in enhanced regulatory capacity. The ability to do this will depend on factors such as the broader objectives of those who you wish to collaborate with.

5. Conclusions
Our objective in this paper has been to develop a set of preliminary principles from an economic perspective that should inform the design and delivery of regulatory approaches in diffuse source settings. We define a regulatory approach as involving at least some element of coercion with respect to landholder behaviour. That is, a binding requirement on landholders usually backed by formal legislation which results in a set of regulations. The requirement may involve actions or results. Our particular emphasis is on regulatory design for water quality in diffuse source settings.
We take an expanded view incorporating elements of behavioural and institutional economics which in turn suggests greater focus on the transaction costs, institutional structures supporting regulation and the broader behavioural response of landholders to regulation.

Ten principles that could be employed to promote economic efficiency within a regulatory approach were identified, divided between four guiding principles and six supporting principles. In brief these principles were:

1. Match regulatory response to severity of damage from a particular pollutant.
2. Match regulatory response to the likelihood of a pollutant being generated.
3. The greater the heterogeneity amongst landholders the greater the potential economic benefits of a flexible approach.
4. Consider all costs of the regulation including to regulatory agencies, landholders and others for any given regulatory approach.
5. Incorporate a graduated response aligned with the environmental and pollutant generation risks.
6. Incorporate a graduated implementation process through time where significant change to attitudes or actions is required.
7. Foster dynamic efficiency and innovation within the regulatory approach.
8. Account for the cost and complexity of monitoring pollutants or behaviour.
9. Ensure design matches institutional and agency capacities.
10. Incorporate complementary approaches to build in local information, reduce costs and empower participants.

We illustrated these principles via a brief discussion of the complexities and issues that would need to be considered in a regulatory approach to water quality in the GBR. There is abundant evidence of a high degree of variability in that matches the first four principles: severity of impact, likelihood of pollutant export, heterogeneity of landholders, and the importance of considering costs to landholders, regulatory agencies and others in design and implementation. Similarly there is evidence that most supporting principles are also important to effective regulatory design in the GBR setting.
References


Queensland Department of Premier and Cabinet 2008, Scientific consensus statement on water quality in the Great Barrier Reef. Queensland Department of Premier and Cabinet, Brisbane, Australia.


Van Grieken, M. 2009. Review of socio economic constraints to and incentives for the adoption of land use and management options for water quality improvement in the Tully Murray catchment (wet tropics). Marine and Tropical Science Research Facility Project 3.7.5 Milestone report.


