Using Overseer in Water Management Planning

An overview guideline

Enfocus

Document reference:	Macintosh HD:Users:gerardwillis:Dropbox:Current project files:Overseer Ltd:Last drafts:Using Overseer in Water Management Planning_FINAL_12 October18.docx
Date of this version:	12/10/2018 11:34 AM
Report status:	Final Draft
This report was prepared for Overseer Ltd under contract by:	Gerard Willis, Enfocus Ltd
Client Review by:	Caroline Read, Overseer Ltd

Enfocus Limited
36B Jellicoe Road
Pukekohe

021 630 349

Contents

1 Introduction

This guide has been prepared to provide an overview of Overseer, its actual and potential use in water quality management and the issues and limitations of its use in that context. It is aimed at making policy-makers and regulators operating in accordance with their Resource Management Act (RMA) responsibilities, aware of best use of the tool and of how to avoid inappropriate use.

The guide provides an overview of key generic issues. More detailed guides may be prepared for specific policy and regulatory applications.

2 What Overseer is designed to do

Overseer is a software application that allows a user to model nutrient flows within a farm system. It is designed as a decision-support tool for farmers and their advisers to provide scientific rigour to nutrient budgeting and assist in making decisions about nutrient management. Importantly, it allows comparisons to be made between different management scenarios. It is capable of testing a wide range of "what if" scenarios – allowing users to estimate production and environmental risks from various farm system changes and management interventions.

Overseer contains seven nutrient sub models that consider the inputs and outputs of the seven main farm nutrients – Nitrogen (N), Phosphorus (P), Potassium (K), Sulphur (S), Calcium (Ca), Magnesium (Mg), and Sodium (Na). In addition to this is also models greenhouse gas emissions on a per hectare and per product basis.

From a water management perspective, a key output of the N sub model is the amount of N lost from the whole farm including, in particular, N leached below the root zone.

It is important to note, however, that although Overseer is a flexible and powerful science-based decision-support tool it was not designed to measure absolute leaching values at a specific point in time. This means that it is necessary to understand the model's innate characteristics and limitations to ensure it is not used outside of its modelling design boundaries when used for compliance purposes. When used in regional plans, planning provisions should be written in a way that is consistent with the model's assumptions and limitations.

3 What Overseer models

3.1 Nutrients

As noted above, although typically associated with N leaching, Overseer actually models flows of all seven key nutrients (as well as greenhouse gas emissions).

3.2 Farm types/Farm systems

In terms of the farm types that Overseer is capable of modelling, a comprehensive list is kept on the Overseer Ltd website at https://www.overseer.org.nz/what-is-overseer. This is updated from time to time.

While the list of crops and systems able to be modelled has grown significantly in recent years, some less common or newly introduced systems, animal species horticultural, arable or fodder crops may not currently be able to be modelled within Overseer.

The current list of systems able to be modelled is as follows.

Pastoral enterprises including dairy, beef, sheep, dairy goats and deer. These include farms using fodder crops (fodder beets, kale, rape, swedes, turnips), forage crops (ryegrass, barley, oats, maize, rye corn, triticale) and animal housing or grazing off.

Permanent fruit crop enterprises including avocado, kiwifruit, apples, peaches and grapes.

Horticulture enterprises[1] including:

Green vegetables – broccoli, brussel sprouts, cabbage, cauliflower, lettuce, spinach
Legume vegetables – beans, lentils, peas
Root vegetables – kumara, potatoes, carrots, beets, parsnips
Other vegetables – onions, sweetcorn, squash tomatoes.

Arable crop enterprises including:

Grain crops – barley, maize, oats, wheat
Seed crops – ryegrass, clover.

3.3 Main Output: Nutrient budget

The main output from Overseer is a nutrient budget showing:

- **nutrients added** (from the atmosphere, supplementary feed, animal transfer, fertiliser, irrigation water);
- **nutrients removed** (in product, in animals leaving the property, in supplementary feed transferred off farm, to the atmosphere, through leaching and surface run-off); and
- **changes in the pool** or "stock" of nutrients held within the farm system.

4 Why Overseer is being used by regulators

Overseer was first used in regulatory (RMA) context more than a decade ago. Over time that usage has gradually increased in line with increasing expectations on regional councils to manage diffuse discharges and achieve freshwater objectives and limits under, in particular, the National Policy Statement for Freshwater Management (NPS-FM). For reasons outlined in this guide, various concerns have been raised from the beginning about the use of Overseer in a

[1] Because Overseer models and represents a long term annual average for nutrient cycling it presents additional challenges in representing vegetable and arable cropping where there is a new crop or multiple crops each year on rotation. Overseer provides for nutrient cycling within these land-uses on a seasonal basis for year-end nutrient budgets, using long term average climate data and an estimate for residual nitrogen from a previous land use cycle.

For reporting and accountability for regulatory purposes, typical long term crop rotations should be provided for.

regulatory context. Despite those concerns, its use has grown to the point of becoming commonplace throughout much of New Zealand. That is largely because there has been a good level of support from both regulators (regional councils) and by the regulated (farmers -supported by much of the broader agricultural sector).

Both parties have been clear that if you are going to regulate farming, regulate for *outcomes* desired (i.e. environmental performance) rather that telling farmers what, and how, to farm by regulating *inputs* (stocking rates, fertiliser use etc) or by controlling more detailed management practices than have traditionally been controlled (beyond, for example, farm dairy effluent discharges).

Performance, or "effects-based", control is generally considered preferable because it allows for flexibility and innovation on farm. It also has the advantage that regional councils do not need to retain large numbers of staff with extensive farm management/farm systems expertise or be involved in typically unwelcome micro intervention in farm management decisions.

Overseer is seen as a way to enable performance-based regulation - at least of diffuse nitrogen discharges (leaching). Diffuse discharges of other key agricultural contaminants (P, sediment and *E.coli*) remain regulated by rules and property-specific farm environment plans (FEPs) that directly control farm practices in many regions.

Despite the predilection for performance-based regulation, N remains the only contaminant commonly regulated in that way. That is because Overseer has been seen as allowing for that approach to be taken for nitrogen whereas similar readily applied quantitative estimates are not available for other diffuse contaminants.

As is discussed in this guideline, however, care should be taken in the way Overseer is used in effects-based planning and, despite the preference for outcomes based-regulation of agriculture, Overseer should not be considered as a substitute for a broad, multipronged approach to water management more generally.

Box 1 - Overseer in the broader context of water quality management

Managing freshwater quality requires a matrix of regulatory and non regulatory interventions to be brought to bear to address point and non point source discharges of a wide range of contaminants as well as a range of physical and biological threats to waterways.

A common tool is to set receiving water quality standards across a spectrum of physico-chemical and biological attributes, many of which may form "freshwater objectives" under the NPS-FM. These standards reflect the outcomes sought for the water bodies required to provide for the values associated with each water body.

In order to ensure the achievement of those receiving water standards, various interventions will be required. These include discharge limits (limits on how much contaminant can leave a process or activity) and other forms of regulatory and non regulatory intervention (including non discharge limits – being limits on the nature and scale of "risk" activities).

Discharge limits

Discharge limits can take many forms.

- Where discharge limits apply to point source discharges they are generally expressed as contaminant concentrations measured in-stream at the edge of the zone of "reasonable mixing";

- Where the discharge limit applies to non point source discharges they can be catchment load limits (e.g. tonnes/year) and/or leaching limits (kg/ha/yr) - generally modelled rather than measured given the technical difficulties of measuring diffuse discharges.

Discharge limits are an important element to water quality management because they make decision-making on individual applications much simpler. In the absence of discharge limits, decision-makers must assess the impact of individual activities on the achievement of receiving water standards (i.e. "freshwater objectives" under the NPS-FM) – something that is technically difficult. Similarly, nutrient load limits can make managing cumulative effects both more certain and more transparent.

The difficulty is that setting diffuse discharge and catchment load limits relies on the ability to quantify the discharge at the level of an individual property. To estimate total diffuse loss of contaminants using direct measurements is prohibitively expensive. There is no accepted way to quantify property or activity-specific diffuse discharges of key agricultural contaminants such as pathogens (E.coli) or sediment. There is, however, a way to at least *model* another key contaminant - nitrogen. That is where Overseer fits in.

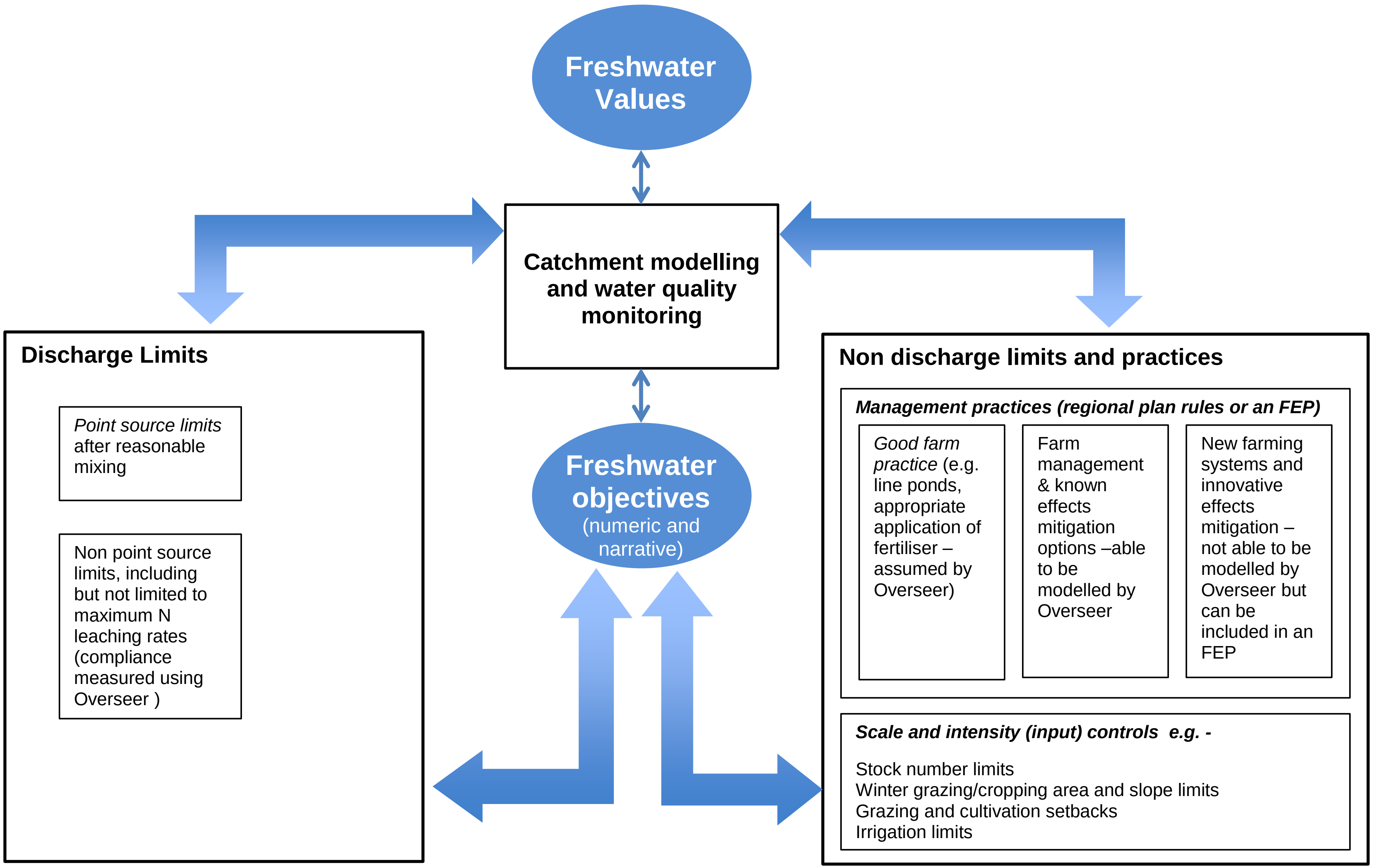
Non discharge limits and other regulatory and non regulatory interventions
Non discharge limits include restrictions on the practices and technologies that may and may not be used – including input standards. They are designed to improve environmental performance but not necessarily to a quantified amount of contaminant reduction (although likely reduction may be modelled or estimated to demonstrate effectiveness). In terms of managing agricultural activities, controls on the collection and storage of farm dairy effluent is an obvious example. Other examples will be limits on the amount of irrigation, winter grazing, stock exclusion and cultivation setbacks. The non discharge limits and practices will typically be specified in regional plan rules and/or in farm environment plans (FEPs) that allow limits and practices to be designed bespoke to individual properties. They will address the full range of diffuse contaminants.

The limitation of non discharge limits and prescribed good management practices (GMPs) is that the relationship between successful implementation of the limit/practice and the receiving water standard is not direct and the effectiveness uncertain. For that reason there is generally a preference for discharge limits to be imposed where it is possible to do so.

It is important to note, though, that there is a wide range of concepts, tools, and models used in water management. Overseer is just one. It does not sit in isolation and will not be effective as a management tool if used without complementary regulatory and non regulatory tools. In short, if used, it should be used in combination with other mechanisms to ensure good water quality management. For the reasons set out in this guideline, a balance needs to be struck between placing emphasis and effort on using N-loss estimates as modelled by Overseer and on applying other tools and techniques including FEPs, GMPs and on supplementing modelling with actual measurement of environmental effects.

The relationship between Overseer, land use management and good water quality management and the components necessary to ensure good farming is illustrated in Figure 1.

Figure 1 – Overseer, Good Farming and Good Water Quality Management



5 How Overseer is being used by RMA policy makers and regulators

Regional councils that use Overseer tend to use it in a number of different ways[2]. These broadly fall into two categories: compliance measurement and limit setting.

5.1 Overseer as the compliance measurement tool

Most obviously, Overseer is specified in regional plans as the modelling tool to be used to determine compliance with N limits specified in terms of kilograms of N leached by hectare per year (kg N/Ha/yr). Such a limit may be imposed as a condition of a permitted activity (or other consent category) thereby acting as a pre-condition for qualification for one category of consent rather than another.

Where a resource consent is required, that consent will also specify the N leaching limit to be complied with (or require observance with an FEP that contains the N leaching limits). The consent will typically require compliance with that leaching limit to be demonstrated using Overseer to model N leaching.

No region currently imposes quantified P loss limits at the farm scale and hence Overseer modelling of P loss is not currently undertaken for compliance purposes (see section 7.1).

5.2 Overseer used in limit setting

Overseer's role in regulation is often more pervasive than simply being used as a compliance measurement tool. Overseer is commonly used in the methodology for *setting* limits.

Overseer modelling is used, for example, to set limits:

- said to reflect the *natural capital* of land (as with Horizons One Plan - where the limits were set for each land use (LUC) class to reflect leaching from that LUC class by a hypothetical farm modelled by Overseer on the basis of the dry matter estimated to be produced from that LUC class); or
- the "baseline" leaching for each individual farm during a benchmark period. For example a certain past year, or averaged over multiple past years (as in the Canterbury Land and Water Regional Plan (CLWRP) and Waikato's Proposed Plan Change 1).

Furthermore, Overseer has also been used (in the CLWRP[3]) in conjunction with what is known as the Farm Portal to set limits based on individual farms operating a *good management practice* (GMP) - by modelling farms using certain proxy input parameters (rather than real farm data) to represent what each farm *would* leach if it was operated at GMP. That theoretical GMP leaching rate is then imposed as the leaching limit.

Another variation in the use of Overseer is to use aggregated farm scale Overseer results to feed into the modelling of potential nitrogen *loads* at a catchment scale (from which in-stream concentrations can also be predicted and limits set accordingly). This is necessary in catchments where land use change is

---

[2] Acknowledging that at the time this guide was prepared not all regional councils were using Overseer.
[3] As inserted by Plan Change 5.

"locked in"[4] (or proposed) and accordingly needs to be provided for in regional plans. In those cases, it is critical to know the additional N (and P) loads and test different land use scenarios so that freshwater objectives and limits can be set at the appropriate level (i.e. to accommodate a future load that may already be consented or is otherwise to be provided for)[5].

6 Some basics about Overseer

Models are complex and the issues associated with their use in regulatory contexts can be difficult to articulate and to comprehend at first blush. This section begins the explanation by setting out the very basic design features of Overseer.

As noted in section 2, Overseer was designed to be used as a decision-support tool for farmers. That application has influenced the design and scope of the model. Overseer is most accurately described as a *strategic tool* for N loss assessment rather than a *tactical tool*. That is, it allows the user to understand the long-term impacts of system-wide changes to a farm, rather that day-to-day changes in N loss.

This is explained further in the following two sections.

6.1 Overseer models rather than measures

It may go without saying, but the first critical point to understand is that Overseer does not measure or even "calculate" actual N leaching. 'Actual' N loss will never be known because it cannot be reliably measured for a whole farm. Overseer *models* that loss by a series of algorithms designed to mimic, as much as possible, the nitrogen cycle as it relates to individual management areas ("blocks") making up a farm property.

As such, Overseer has many of the same limitations as other sorts of models. Models inevitably simplify very complex processes and tend to standardise highly localised variability. Overseer is no different.

This is most graphically illustrated by considering the way Overseer accommodates climatic variables. Climatic variables (rainfall, evapotranspiration and temperature) can have a significant influence on both modelled and actual leaching rates. While Overseer allows users to specify their own annual climate data, it is recommended best practice to use the "climate station tool" provided within Overseer to determine those values. The climate station tool draws on a NIWA-supplied database of average annual values over the period 1980-2010 (which are then broken down into monthly data). Values are based on data collected at specific weather stations (often at the closest town) and the data selected relates to the closest station.

Because Overseer provides for long-term annual-average estimates of nutrient cycling, it is appropriate that the climate data used is the long-term annual-average climate data. This is not a flaw in the model but it does mean that, to the extent that leaching rates are influenced by short term climatic variables,

[4] Such as, for example, where irrigation consents have been granted allowing for land use change.
[5] Catchment loads are not always calculated in that way. The alternative approach is to calculate target loads but multiplying concentration by flow.

modelled outputs need to be understood to be long-term average leaching rates rather than actual representations of leaching in any particular month or year. This is discussed further in section 7.4.

6.2 Overseer only models loss below the root zone - not to water

The second important point is that Overseer does not model N loss to groundwater or surface water. It models N leached *below the pasture/crop root zone*. How much of the N leached below a farm ends up in ground or surface water depends on the extent of denitrification in the vadose zone (being the conversion of nitrate and nitrite by soil bacteria in anoxic environments to N's gaseous form, N_2), with the rate sometimes referred to as an attenuation rate. Denitrification occurs at different rates depending on various biophysical conditions (particularly the amount of carbon in soil and temperature) and hence rates of denitrification can vary significantly across a landscape. Denitrification rates are highest where there is high carbon content in soils – in areas of peat or in production landscapes developed through wetland drainage where decaying vegetative matters persists in soils.

Whether or how much of the total N *load* lost to groundwater (after any denitrification in the vadose zone) ends up in a particular surface water body depends on local hydrogeology (and, in particular, on the extent of groundwater discharges into the marine area). What the *concentration* of N in a ground or surface water body might be depends on other inflows and the diluting effect of those inflows.

In short, Overseer only models the nitrogen cycle to the bottom of the root zone and to the farm gate and is only part of the answer to the question of what effect land uses might be having on receiving environments. To understand the effect of N lost below the root zone or in surface run off beyond the farm boundary catchment models are needed[6].

6.3 Data inputs

Overseer requires users to enter input data. Some data is compulsory and some is not. Some data entered may be specific to the farm but in many cases default values are available if farm specific data is not. In some cases the Overseer best practice data input standards (OBPDISs)[6] require the use of default Overseer-provided data rather than farm specific data.

What data is used obviously has the potential to affect modelling results. A standardised and consistent approach to data entry is encouraged, and it is important that those using the model for regulatory purposes have good understanding of the model and its modelling assumptions and the different approaches to modelling a farm that are possible within the software.

6.4 Spatial framework for modelling

Although we refer a lot to Overseer modelling farm N leaching, in fact Overseer models largely at the scale of a user defined management "block". Those using Overseer will divide farms into multiple blocks that reflect their use. These blocks will also be differentiated by bio-physical attributes (e.g. soil, topography).

[6] See Box 3 for explanation

Some information is entered into Overseer at the farm scale but much is entered at the block scale. How the farm file is set up (how many blocks are defined and where the boundaries are drawn) is therefore a critical factor. Defining those blocks accurately is key to generating reliable and meaningful outputs. This is another reason why nutrient budgets developed for regulatory purposes should be produced by an experienced, certified adviser with a good understanding of farm systems and the Overseer model and its assumptions.

7 If you are going to use Overseer in regulation what do you need to know?

7.1 Modelling P loss

Aside from N, the other important nutrient from a water management perspective, is P. Overseer's P sub-model assumes run-off as the main discharge pathway (as opposed drainage which is the main transport pathway for N).

Understanding Overseer's approach to modelling P loss is important because although no regional council currently sets property-specific P loss limits (with compliance to be measured by Overseer), there are sometimes calls for such an approach to be taken.

Overseer's P sub-model uses a "risk" approach. That means it considers sources of P[7] and predicts losses taking into account well-accepted risk factors (e.g. rainfall, topography, soils). Expected P loss is calibrated against catchment studies.

Hence the P loss model will take into account critical source areas (CSAs) in the sense that losses recorded in catchment studies include losses from CSAs, but it will not necessarily model the losses that are particular to the modelled farm. For the same reason, it will not take into account any CSA-specific mitigations that may be in place on the modelled farm.

It is also important to note that Overseer makes assumptions about loss of P in sediment due to topography, mass flow, stream bank or stream bed erosion but does not model short term changes to these critical source areas.

Box 2 - Implications for use: Overseer and P limits

Overseer models farm system P losses at block scale, based on topography, land use, soil type and climate. Apart from the use of wetlands it cannot be used to estimate losses based on mitigations implemented at the block or whole farm scale, and therefore is probably of limited value to the regulatory context.

Focusing regulation only on those sources of P able to be modelled by Overseer may lead to inefficient regulation because it may result in emphasis on P sources with high cost mitigation when P sources with lower cost mitigation are ignored.

[7] Sources include background and incidental run off (i.e. P in soil, effluent and fertiliser), P associated with stock and drains accessing streams. Point source discharges of effluent to streams, border dyke irrigation and septic tank drainage.

7.2 Overseer assumes best practice farm management and quasi-equilibrium

Overseer assumes that practices implemented on farm follow recognised good practice. For example, Overseer will assume that when fertiliser or effluent is applied it is applied evenly at the time stated and not poorly (e.g. applied unevenly or at times of heavy rainfall etc).

This can mean that when changes are made to improve practices, or to ensure that GMPs are undertaken more consistently than they were, actual N loss will decrease but that will not show up in the N loss estimates. Another example is where a farmer invests in lining an effluent pond. Such investment will likely have a material effect on the amount of N leached from the pond, but the Overseer modelled farm leaching rate will remain unchanged from the pre-investment result. That's because Overseer already assumes the pond does not leak.

Conversely, if management practices regress on a farm (because, for example, of change in personnel), or if farm infrastructure (such as a pond liner) fails, the effect on N loss will be invisible to Overseer. This reinforces the fact that Overseer models a system based on long term annual averages but not necessarily all the detailed practices that, on a day-to-day basis, can make a difference to short-term N loss.

Similarly, Overseer assumes that inputs and farm management practices are in *quasi-equilibrium* with the farm productivity. Quasi-equilibrium means that the model assumes that the inputs and outputs are in equilibrium with the farm productivity.

In other words, the model assumes that the input data "makes sense" in that the inputs match the outputs. What this means in practice is that the model does not validate, "auto correct" or send an error message when the two sides of the ledger (i.e. inputs and outputs) do not add up[8]. This means that it is possible for unrealistic systems to be modelled producing N loss results that will not be reliable.

It is important to understand that this characteristic is not a flaw in Overseer but a deliberate design characteristic that enables Overseer to model less conventional systems such as organic and bio-active farms.

Accordingly, those using, or auditing, Overseer modelling (or any report on farm system impacts) need to understand what a realistic and an unrealistic farm system looks like.

7.3 Two types of Overseer Budgets

Although reference is typically made to an "Overseer budget" or to "Overseer modelling results", in fact, Overseer can be used in two different modes to generate two different budgets – an *actual* (referred to as *Year-End*) budget and a *predictive* (referred to as a Predictive or Scenario) budget.

The year-end budget uses data that describes the current annual farm system operating to model losses. Year-end budgets are commonly created using data

[8] The only exception to this is the Overseer will generate error messages when the system being described by the input data would not produce (or would produce too much) grass for the stock said to be carried.

from one year's activities, but the current advice[9] is that five years' worth of actual activity data is required to estimate an annual average leaching number. Further work is needed to refine recommendations on the averaging of activity data for monitoring purposes. Implications of using annual data with a long-term annual average model to produce a Year-end nutrient budget are discussed in Section 7.4.

The predictive budget uses data based on what is planned to happen on the farm over a future time period.

In other words, Overseer can be used to model what happened, or what is planned.

It is important that regulators are clear about which budget is the relevant one for the purpose of compliance. This is discussed in more detail in section 7.4 below.

It is also important to note that the approach to auditing budgets will vary depending on what type of budget is required. Verification of the accuracy of year-end budgets is relatively straightforward with paperwork associated with financial transactions (invoices and receipts) and stock movements required to be kept for tax and other purposes (e.g. NAIT in respect of cattle). Given predictive budgets represent future scenarios, evidence to demonstrate management approaches were used would need to be collected after any systems changes are implemented. As discussed later, where no change is proposed for the farm system, predictive budgets can be based on long-term average management (e.g. irrigation) and production data. Where changes to production systems are proposed that approach is not possible.

Audits will require supporting data from the past 3 or 5 years, where it is retained and available. Some regulatory approaches require this data is retained and made available to council on request. In the event of a future scenario analysis, the nutrient budgets must be audited against the Farm Environment Plan to ensure it is consistent with the proposed operation of the farm.

Predictive budgets will be required if there is a need for a consent applicant (for example) to demonstrate compliance with out-year limits (that is, an obligation to reduce N leaching by a required amount by a specified future date) in order to be granted consent. In such instances compliance with predictive budgets cannot be verified at the time a consent is issued and must be taken on good faith. However, year-end budgets can be prepared at those future dates to demonstrate that average leaching performance corresponds with earlier predictive budgets.

7.4 Long term averaging

As noted earlier, Overseer is a long-term averaging model. That is, the nitrogen leaching estimate generated represents the long-term annual average leaching from the farm if the management system described remains in place. It is not a one-off prediction for a single year (despite often being presented in those terms).

---

[9] David Wheeler, M Shepherd, M Freeman and D Selbie, AgResearch, Hamilton, OVERSEER® NUTRIENT BUDGETS: SELECTING APPROPRIATE TIMESCALES FOR INPUTTING FARM MANAGEMENT AND CLIMATE INFORMATION, In: *Nutrient management for the farm, catchment and community*. (Eds L.D. Currie and C L. Christensen). http://flrc.massey.ac.nz/publications.html. Occasional Report No. 27. Fertilizer and Lime Research Centre, Massey University, Palmerston North, New Zealand.

As discussed in section 6.1, one of the ways Overseer takes a long-term averaging approach to modelling, is in the way it uses climate data. Rather than using short term or "real time" climate/weather data, Overseer is designed to use average monthly data calculated over a 30-year period. Again, this is a deliberate design feature rather than a "flaw". It allows assessment of the impact of a system change or management intervention without the "noise" of weather variability.

However, it will mean that the extent of actual drainage (and hence N leaching) in any one year could depart significantly from that assumed by Overseer if the rainfall, or rainfall pattern, in that year differs significantly from the long-term average.

Despite that, it is important to understand that much input data can be entered as either annual (actual) data or long-term average data.

The current Overseer best practice advice is that when Overseer is used to produce a predictive nutrient budget, long-term average management and production data should be used. While advice is that five years' worth of actual data is needed to produce a long-term annual average (that is comparable to a predictive estimate), more analysis is needed to understand the uncertainty associated with modelling one year's worth of activity data.

The ability to use year-end annual data was added to Overseer because regulators wanted to be able to model actual annual management (being a more easily verifiable concept in terms of having the "paperwork" to more easily substantiate claimed input data).

Given the use of annual actual data is not necessarily appropriate to use based on the modelling assumptions, a note was added to the Best Practice Data Input Standards that recommends: "*If annual data inputs are used, it is also recommended that a rolling average or trend analysis of outputs is used to reduce the impact of year-to-year variability when monitoring the degree of compliance with any target or critical value. In addition, the uncertainty of Overseer predictions can be reduced if the focus is on a percentage change over time (rather than an absolute change*)."

Issues relevant to regulation and compliance can arise, however, if certain annual data are used in the year-end budget. That is for two reasons.

- First, a year-end budget may not give a good representation of actual leaching when annual data is used against a background of long-term average weather data.

  To illustrate the point, irrigation rates vary from year to year depending on weather conditions. In a dry year more water is applied. Because climate/weather data used in the modelling is long term average (rather than actual), Overseer will model the N loss to be higher than it should be since it will not "see" that the irrigation reflects a dry year and assume additional drainage (and hence leaching) that will not have occurred in reality.

- Second, if the target is calculated on the basis of expected long-term average performance then measuring a one-off year based on real data will not be comparing "apples with apples". As discussed above, there will be greater variability in the year end nutrient budgets than will occur when using long term averages for input data used in predictive nutrient budgets

Accordingly, regulatory regimes that require annual budgets to be prepared using year-end input data from the past year may need to provide for annual variation from targets in recognition that the target is based on average production and average management settings. Alternatively, regulation will need to permit year-end budgets to be prepared using long-term (5+ years) average data (rather than actual data from 1 year's activities). Another similar option is to allow for the use of multiple year rolling averages (the current recommendation is for 5-yearly) for compliance purposes.

7.5 Overseer modelling to assess relative change in leaching

For the various reasons set out in this guide, Overseer in a regulatory context is probably best regarded as a tool for assessing the *relative change* in nitrogen leaching between different points in time rather than a model that attempts to estimate N leaching in absolute terms.

This means Overseer will be very good at assessing whether a (say) 10% reduction in N leaching has occurred on a particular property (given a series of practices) over a prescribed (say) five-year period. It can be used in that way with considerable confidence. In regulating N leaching the absolute number may be much less important than knowing whether a prescribed level of reduction has been achieved. In that way Overseer is highly suited to estimating relative reduction from a baseline leaching rate, but much less suited to estimating whether a specific numeric limit has been complied with (unless, as indicated above, that numeric limit is calculated as a percentage of an Overseer modelled "baseline" rate).

7.6 Uncertainty and sources of potential inaccuracy

One of the first questions asked about the use of Overseer in a regulatory context is "how accurate is it?"

As with any model, Overseer is never going to be 100% accurate in the sense that the estimated N loss will always equal the measured N loss at the block or farm scale (assuming it could be measured perfectly).

In consideration of error and uncertainty, it is important to understand that direct measurement of leaching loss, using a lysimeter for example, will include measurement error, attributed to the measuring tool itself, and sampling error, attributable to the soil and site characteristics where the tool is set down. To have a representative direct measurement of the farm leaching losses would require many lysimeters to be installed on each farm to account for this variability in measurement and sampling. This would be prohibitively expensive to operate for each farm, and without this high level of replication in measurement, the error and variability introduced by direct measurement would likely be greater than the error and uncertainty associated with modelling.

There are a number of sources of potential uncertainty that are important to understand if you are going to place regulatory weight on a farm's modelled leaching rate in any particular year.

The fact that Overseer can use long-term averaged data rather than real time data is one reason the model will likely yield a leaching rate that differs from the actual rate at any particular point in time. But there are three other sources of uncertainty as discussed below.

7.6.1 Errors inherent in the use of measured data

Overseer relies on inputting data (both user supplied and default data) that is gathered through field measurements. That data will contain errors because it is only a representation (sample) of conditions at the time and place of measurement that may not be perfectly representative of what is actually occurring on the farm being modelled. Furthermore, instruments for gathering data (such as flow meters or soil moisture probes) have a degree of inaccuracy. In short, all scientific data has a margin of error.

7.6.2 Data input errors and inaccuracies

As can happen when entering any data, errors and discrepancies can occur. These can be:

- *inadvertent* errors as happens when someone enters a different value than they intended.
- entry errors resulting from misunderstanding or lack of knowledge (where someone enters a value they believe to be correct but which isn't).
- *Interpretation differences* occur when one user judges that the certain input parameter applies but another user would adjudge that the different value should be entered. This occurs when there is professional discretion/expert judgement to be applied and subtle differences in input parameter options. In short, that means that different Overseer users can generate different results in respect of the same farm.
- errors or variation associated with a user entering data in an attempt to "work around" a problem with Overseer not being able to model a particular farm system.
- deliberate attempts at manipulation introducing inaccuracies where the user enters a value they know to be incorrect (that is it does not reflect the farm system being modelled) but which they believe will contribute to a modelling result that is in their, or their client's, interest. (These are not always easy to detect and subtle differences in parameters such as soil type or rainfall, for example, can significantly affect the modelling result).

A particular source of potential inaccuracy is associated with the way the farm is set up and described in Overseer and its relationship with the actual farm system.

7.6.3 Model design – imperfection in the model's representation of natural processes

Overseer models very complex systems and relationships. It inevitably simplifies these complex systems. That means there is uncertainty in the validity of principles underpinning the model design, some of which relate to matters know we do not fully understand (known unknowns) and some of which will be unknown unknowns.

Three particular characteristics of the model that need to be born in mind are as follows:

(a) Overseer (rightly) takes into account only what is known with sufficient certainty to be able to model. Not all systems and practices employed on farms are able to be modelled within Overseer. Similarly, not all mitigations that a farmer may be using will be able to be fully accounted for by Overseer (at the very least there will be a delay between practices and technologies being used and Overseer being updated to account for them. That is because Overseer relies on approaches to have an acceptable level of evidence to update and then to validate the model. The time to develop the evidential database can lag behind the new innovations being adopted by farmers);

(b) There is micro variability across farms and within management blocks (such as changes in soil types and the effects of landscape features) that may have a significant effect in the short term but which Overseer cannot take into account;

(c) As noted in section 7.2, Overseer assumes that certain good management practices are adopted on farm when they may not be. Hence adoption of those practices or regression from those practices will not be reflected in Overseer modelling results.

7.6.4 Quantification of a margin of error

The only quantification of the margin of error in Overseer modelling was undertaken when Overseer was first released[10]. An assessment by Ledgard and Waller (2001) compared measured drainage nitrate nitrogen with modelled farm block drainage nitrate nitrogen estimates. They concluded that the "error" (or imprecision) in the long-term estimate of average nitrate leaching for pastoral systems is around +/- 20%. The estimated uncertainty in nitrate-N drainage was given as at +/- 25-30%. That figure has since been widely quoted.

No updated qualification of an error margin has been produced. We do know that the 2001 estimate did not take into account errors associated with measurements or uncertainty from data input (i.e. sources outlined in 7.5.1 and 7.5.2 above). On the other hand numerous improvements have been made to Overseer that will have had implications (generally positive) for uncertainty and error margins.

Overall, it is generally thought that more recent versions of Overseer will have a similar level of accuracy relative to measured leaching to that estimated in 2001.

Although 30% may seem high in terms of regulatory test, it is considered good in the context of modelling of complex biophysical systems, and as discussed above, direct measure (for example using lysimeters) also include significant measurement and sampling error.

[10] Although there have been several papers addressing it; e.g. OVERSEER®: ACCURACY, PRECISION, ERROR AND UNCERTAINTY Mark Shepherd, David Wheeler, Diana Selbie, Laura Buckthought & Mike Freeman (2013); and Shepherd, M. A., Wheeler, D. M., Freeman, M. F. & D. R. Selbie. (2015). Rationale for Overseer® Nutrient Budgets evaluation and recalibration. A client report submitted to Overseer Management Services, May 2015.

Box 3 - Implications for use – minimising potential for user inaccuracy

Not all of the above uncertainty associated with Overseer can be removed entirely. Three measures are, however, already extensively used to reduce the potential for the inaccuracies or to minimise the significance of the effect of inaccuracies on the modelling results.

Overseer Best Practice Data Input Standards

Overseer Best Practice Data Input Standards (BPDISs) have been developed by Overseer Ltd. These standards provide guidance on the correct data input to use. They aim to reduce inconsistencies between different users when operating Overseer to model individual farm systems.

Planning provisions that specify the use of Overseer should always include reference to the fact that Overseer should be used in accordance with the Overseer Best Practice Data Input Standards.

Nutrient management qualifications

Massey University offers certificate grade qualifications in Intermediate and (separately) Advanced Sustainable Nutrient Management. Achieving proficiency in the use of Overseer is a core competency in gaining the certificate. Users can also become *certified* under the *Nutrient Management Advisor Certification Programme* (NMACP). NMACP is operated by the Fertiliser Association of NZ (FANZ)[11]. The standards and criteria for certification include holding the Massey University certificates referred to above.

Overseer is an expert user system. Planning provisions that specify the use of Overseer should always include reference to the fact that Overseer modelling should be undertaken by a person with the Advanced Massey certificate as the minimum requirement when the results are to be used for compliance assessment.

Audit

Regional plans generally require farmers to model N losses using Overseer. Although this should be done in accordance with the BPDISs and by a suitably qualified person (being a person with the Massey University Advanced Sustainable Nitrogen Management Certificate), the fact remains that the person undertaking that work is working for the farmer client and therefore, issues of independence will arise. One way of addressing that issue is for the Overseer modelling analysis to be subject to separate audit.

Requiring audit by a suitably qualified person that was not involved in the initial Overseer modelling reduces the potential manipulation of input data and may be appropriate in some circumstances.

7.6.5 Residual uncertainty and the relevance of model assessment (calibration and validation)

Although the measures outlined in Box 3 above will help to reduce some of the opportunities for error and uncertainty discussed earlier, they will not do so

[11] Although now at arms length through the subsidiary company, *Nutrient Management Advisor Certification Programme Ltd*

entirely. Furthermore, issues associated with model design and use of measured data will not be improved by those measures (which relate solely to user/data input errors).

As already noted there will always be a margin of error - that is, a difference between the leaching modelled by Overseer and the leaching that is actually occurring.

The degree of uncertainty will increase the further a farm system departs from the conditions applied in the model calibration and validation processes (collectively known as the “model assessment process”).

The pastoral N leaching sub model has been continually validated (where modelled results are compared to measured leaching) and calibrated (adjustments made to the model to better match modelled leaching with measured leaching). However, these validation and calibration processes have not (and cannot) be undertaken for all potential conditions or all farm systems due to limitations in information availability.

The closer the farm system being modelled is to the conditions and system used in model assessment, the greater the certainty in the results. In broad terms, this means that modelling simple farm systems in flat pastures on common soil types will produce results with greater certainty than modelling complex farms in other landscapes or extreme climate conditions.

7.7 Overseer version updates

Overseer is updated twice yearly. One of those updates is generally substantive in nature – meaning that it can include enhancements that make a material difference to modelling results. That poses a problem for regulatory systems relying on Overseer modelling to demonstrate compliance with a fixed N loss value.

7.7.1 Need to update N leaching limits derived using Overseer modelling

As noted earlier, N leaching limits can be calculated using Overseer either in the form of:

(a) a “baseline” leaching rate from the modelled farm (i.e. leaching as it occurred at a fixed date) which forms the basis of future limits that apply to that same farm; or

(b) some other numeric limit that relies on, or is derived from, Overseer modelling such as from some fictitious or “model” farm.

When such limits are set they are generally based on the principle that the extent of allowed N loss is to be equivalent to a particular level of performance (i.e. a “baseline” year, or baseline +/- a percentage, or baseline adjusted for GMP, or a theoretical farm represented by a theoretical input data). That level of performance is quantified using the specified version of Overseer at the time the limit is set.

Similarly, the costs and benefits of compliance with such limits are assessed as acceptable or not on the basis of estimates using Overseer. That can only be done on the basis of the version that existed at the time the assessment is undertaken.

However, when at some future point in time, Overseer is used to assess compliance against limits, the version of Overseer that exists at that time will be different to that used to set the limit and assess costs. Because it includes new science or has addressed identified errors, that later version will, in all likelihood, generate a different leaching rate than the earlier version of Overseer would have generated. (And, if that later Overseer version was used to define the limit, that limit would also be different if the updated version of Overseer was used).

What is important in setting limits is the principle or underlying rationale for setting the limit at a particular level, rather than what number Overseer models the N loss to be that will reflect that principle or rationale. The number is, after all, only a best estimate given current modelling capability. So where Overseer sets a benchmark, based on a stable farm system, a different version of Overseer will set a different benchmark, despite no change in actual farm system nutrient losses.

That means that to be fair and equitable and to apply the underlying principles and rationale consistently (to achieve the level of outcome initially sought), the version of Overseer used to set a limit needs to be the same as that used to assess compliance with that limit.

Theoretically, that can be achieved either by ensuring that:

(a) the version of Overseer used to set the limit is used in all future years to assess compliance with that limit; or
(b) the limit is updated upon the issuing of any new version of Overseer (and that new version is used for compliance).

In practice, updates of Overseer result in earlier versions becoming unavailable and hence option (a) above is generally not viable. In any event, that approach would constrain water management according to older science, and should not be regarded as good practice.

Exceptions for leaching rates calculated on a different basis?

An exception to the above will occur when Overseer is not used to derive the leaching limit. That may be the case, for example, where N leaching limits are based on an allocation of targeted N load derived from in-stream DIN concentration limits. However, even in those instances, care should taken if Overseer was used to assess the cost and benefits of the limit as part of the limit setting process. In such cases, the cost estimate may prove unreliable if Overseer updates result in modelled leaching rates different from those generated by the version used for the initial cost benefit assessment.

7.7.2 Need to update Overseer-derived catchment load limits

As discussed earlier, Overseer can be, and has been, used to set catchment N load limits (tonnes/yr) by modelling N leaching from current and/or predicted future land use scenarios as part of catchment limit setting processes.

In other words, catchment load limits can be set as the aggregation of landscape wide Overseer modelled N loss from a desired current or future land use pattern.

When an update of Overseer is released and is required to be used by existing and new farms, the load limits will need to be reviewed if the aim is to ensure continued provision for the desired land use pattern.

Similarly, Overseer has been used as part of wider catchment models to gain an understanding of the relationship between a current land use pattern (and associated aggregate diffuse N discharge) and a particular in-stream N concentration. From that, an understanding of the relationship between future land use patterns and in-stream N concentrations can be gained.

When an update of Overseer is released, the understanding about the relationship between the existing aggregate diffuse discharge and the in-stream N concentration will need to be revisited since the new version of Overseer may indicate that a higher or lower level of discharge (relative to that calculated by the earlier version) is resulting in the in-stream N concentration being observed. This may, accordingly, also require revisiting any predicted effects of changing land use and anticipated water quality outcomes.

Box 5: Implications for use: setting and updating limits and assessing compliance

When Overseer is updated and a new version issued, any limit derived from a methodology that uses Overseer (whether a leaching limit or load limit) should be updated using the latest version of Overseer and any assessment against that limit should be undertaken using the latest version of Overseer. It should be stressed that doing this does not mean that the receiving environment receives a greater or lesser load of nitrogen – only that the estimation of that load is more accurate.

This applies to limits set in regional plans and any limit imposed by way of condition of resource consent.

In terms of limits set in plans, that can be achieved by:

- expressing the leaching limit as a methodology (e.g. “the average annual N leaching rate occurring over the 2013-2015 period as modelled using the most recent version of Overseer” (Canterbury Land and Water Regional Plan); or
- allowing updated limits to apply through the consenting process - with the methodology for updating prescribed in appendices of plans (e.g. Canterbury Land and Water Regional Plan: Plan Change 3)
- having limits calculated using external calculators (incorporated by reference into a plan) that provide for the use of the latest version of Overseer (e.g. Canterbury Land and Water Regional Plan: Plan Change 5).

Leaching limits imposed by way of resource consents should be expressed as a formula (e.g. as the leaching rate determined by the most recent version of Overseer using a specified data input file).

8 Safe ways to use Overseer in a regulatory context

For all the above reasons there are ways in which Overseer can be "safely" used in regulation and ways in which it presents some risk.

As a general observation, the approach to be taken with Overseer is perhaps best summed up in the following quote from a study of the use of models in environmental regulation in the United States:

> *Models will always be constrained by computational limitations, assumptions, and knowledge gaps. They can best be viewed as tools to help inform decisions rather than as machines to generate truth or make decisions*[12]*.*

While it is not for Overseer Ltd to advise RMA practitioners on what provisions a regional land and water plan should contain to address diffuse discharges, the following may be of some interest.

8.1.1 Risks of using Overseer as a strict pass/fail decision-making tool

Overseer should, and will continue to, be an important tool in managing diffuse discharges of N and water management generally. However, there needs to be acknowledgement that using Overseer to demonstrate compliance, or compliance failure, against highly specific N leaching limits and to insist on corrective measures on farm in the event that Overseer estimates small exceedances of limits may:

- Lead to inequities in the way some farmers are treated relative to others that is unrelated to actual nitrogen loss performance.
- Drive incentives for "creative use" of Overseer
- Be difficult to justify and to enforce when legal tests of proof are applied.

8.1.2 Overseer in "safe harbour" regulation

For the above reasons, Overseer-derived limits and Overseer compliance assessment can be most appropriately used in "safe harbour" regulation. Safe harbour regulation is regulation that is designed to provide an easy, uncomplicated and non-discretionary route for a person to demonstrate regulatory compliance with a performance benchmark that is inherently ambiguous or which otherwise requires discretion to be exercised by the regulator/decision-maker. However, in regulatory design a safe harbour is not the only route to authorisation or demonstration of compliance.

In the RMA context, safe harbour can be thought of as a permitted or controlled activity rule in a regional plan. An Overseer-derived, or compliance assessed, limit could be a condition of such rules. That would be appropriate if, and to the extent that, an alternative pathway to demonstrating compliance is available through the resource consent process.

[12] *Models in Environmental Regulatory Decision Making 2007*, Committee on Models in the Regulatory Decision Process, Board on Environmental Studies and Toxicology, Division on Earth and Life Studies, National Research Council of the National Academies (US)

8.1.3 Overseer in "unsafe harbour" regulation

As noted above, Overseer can be used to estimate farm performance against a threshold value where failure to meet the value triggers closer scrutiny of the farm operation (but not necessarily a decision not to authorise the farm leaching above the trigger value).

This implies that once a N leaching level is breached (as estimated by Overseer) consideration is given as to whether the breach is reasonable given circumstances of the farm and the control that may already be exercised over inputs and practices that appear to be contributing to N loss. The question of whether the farm is consistent with the conditions applied to Overseer's validation and calibration will similarly be relevant.

In other words, compliance with a specific N leaching number by reference to Overseer modelling is not the sole means by which a farm may be authorised. Other factors are considered in recognition that Overseer is a model with all the limitations discussed in this guide.

The downside of this approach is that it requires individualised and specialist consideration of farm systems in the context of resource consenting. The upside is that it allows individual circumstances relating to matters relevant to likely actual leaching and environmental effect, to be taken into account.

This approach contrasts with one that uses Overseer as part of a pass/fail test that sees a limit imposed and the activity unable to be authorised (even under a consenting regime) until such time as Overseer can demonstrate that the limit will not be exceeded.

8.1.4 Overseer as an estimator of relative change in N leaching

Overseer could also be appropriately used to measure farm performance against a benchmark leaching rate for the purpose of measuring reductions achieved.

Under this application, absolute leaching numbers are less important than the trajectory and scale of leaching change over time.

Where Overseer is used to measure performance against a baseline rate (applied to the same farm) it can provide a good measure of change from that baseline. Many of the input data (soils, rainfall, temperature etc) will be unchanged from the baseline Overseer modelling.

Similarly, modelling to determine whether a (say) 10% reduction is achieved from that baseline would be a reasonable use of Overseer because it is modelling system changes only.

That said, the matters discussed in section 7 above continue to apply.

9 Summary

Using models in regulation introduces additional challenges, however, Overseer can be extremely valuable both as a tool for compliance assessment and as a tool used in wider limit setting processes.

One of the key benefits of Overseer is that provides a tool that enables N leaching regulation to be effects-based, meaning that N leaching targets can be met in ways that best fit individual farms. This contrasts with the alternative approach of relying only on input controls that can lead to higher than necessary adjustment costs and stifle innovation.

When used by regulators and policy-makers outside of its design purpose, the following must be recognised:

- Overseer *models* rather than measures N loss below the root zone and has all the usual limitations of models.
- Overseer only models N (and all other nutrients) to the farm boundary, not to the groundwater or surface water. Hence, it provides limited information on the risk to the environment.
- Because Overseer is designed to be a strategic decision-support tool for farmers that uses a long term averaging model where certain factors (e.g. daily or seasonal variations) that might complicate or confuse signals for robust on farm decision-making are removed. This makes Overseer a tool suited to long-term N loss trend assessment and farm planning rather than tactical (short-term/real time) N loss performance assessment.
- Overseer contains a margin of error that reflects the potential for user error, model design error and errors associated with measured input data. The potential for user error/inconsistency and manipulation can be reduced by requiring adoption of Overseer Best Practice Data Input Standards, requiring budgets to be prepared by suitably qualified persons and by providing for a level of independent audit). However, other sources of uncertainty will persist. There are therefore significant risks in using Overseer as a pass/fail decision-making tool.
- Like all models, Overseer is constrained by computational limitations and knowledge gaps, and should not be used in isolation or as the sole measure of good farming practice. However, when used in combination with other tools and mechanisms (including FEPs and regulatory control over high risk activities), Overseer provides valuable farm-specific information to support good water quality management.
- Overseer should be supported by measurement and monitoring of effects on water quality.
- For all the above reasons Overseer is best used in regulation in the role of measuring compliance with “safe harbour” limits and/or to indicate directional and/or relative change in leaching (i.e. a proportional reduction or increase from a baseline) rather than in the role of demonstrating compliance with an absolute leaching rate limits that is unrelated to the particular farm being modelled.

Finally, it is important to reiterate that the issues and uncertainties highlighted in this guide are not indicative of flaws in Overseer. That is partly because all models have limitations, they will always be imperfect. But it is also because when Overseer is used for its primary intended purpose any uncertainties or inaccuracies that result from the model are largely immaterial with implications “at the margin” for certain on-farm management decisions. For example, a little too much or a little too less fertiliser might be applied than strictly optimal for the farm system. These decisions can be easily rectified the following year(s).

However, when Overseer is used in the regulatory context for setting in-stream or leaching limits, or for measuring compliance with limits, Overseer should be used with care, taking due consideration of the inaccuracies, uncertainties and limitations of the model. For that reason, Overseer’s role and value in water management must be kept in perspective. It constitutes a valuable *decision-support* rather than a decision-making tool.