Manaaki Whenua
Landcare Research

Collation of nutrient, sediment, and *E. coli* losses from land uses to freshwater, and an initial analysis of some factors contributing to nitrogen loss

Prepared for: Ministry for the Environment

August 2022

Collation of nutrient, sediment, and *E. coli* losses from land uses to freshwater, and an initial analysis of some factors contributing to nitrogen loss

Contract Report: LC4185

John J. Drewry

Manaaki Whenua – Landcare Research

Richard McDowell, Chandra Ghimire, Alasdair Noble

AgResearch

Reviewed by:	*Approved for release by:*
Andrew Manderson Team Leader Manaaki Whenua – Landcare Research	John Triantafilis Portfolio Leader – Managing Land & Water Manaaki Whenua – Landcare Research

Contents

Summary

Project and Client

- The Ministry for the Environment (MfE) is investigating diffuse nutrient, sediment, and *E. coli* losses from farmland and the risk to freshwater. Some concerns have been raised about a commonly used model (OVERSEER) for regulation purposes.
- The Ministry's Expert Advisory Group provided advice on future options. These include a recommendation to develop a Nitrogen (N) Risk Index Tool, as an additional or complementary tool for use.
- There is therefore a need to obtain published literature data that reflect the losses from farm management and environmental characteristics, by collating a range of loss values and site characteristics such as climate conditions.
- There is also a need to determine how these site characteristics could influence total nitrogen and nitrate losses for land uses.

Objectives

- To collate New Zealand published literature values of P, sediment, and *E. coli* annual losses for selected land uses, with site characteristics that may influence their loss, into a database.
- To collate predominantly New Zealand published literature values of total nitrogen and nitrate annual losses for selected land uses, with site characteristics that may influence their loss, into a database.
- To conduct, in a parsimonious way, a multivariate analysis of characteristics in the database to determine those site characteristics that have the strongest statistical relationship with total nitrogen losses for a range of land uses.
- To develop equations for total nitrogen loss for seven land uses for runoff and leaching.

Methods

- A search of New Zealand-based studies from the published science literature and grey literature was conducted.
- A list of criteria were developed for inclusion of the N, P, sediment and *E. coli* losses and site characteristics for a range of land uses (dairy, sheep, beef, deer, forestry, horticulture, cropping, vegetables, native bush) into an excel database.
- Losses were expressed as annual yields e.g. kg N/ha/yr, for any one or more of nitrate-N, TN, DRP, TP, sediment, *E. coli*. Losses spanned a year or multiple years.
- Modelled estimates from literature studies using versions of OVERSEER were excluded due to recent concerns about the model.
- A range of site and study characteristics were recorded including land use, location, whether measured or modelled, mean annual temperature, slope, soil order, soil infiltration, anion storage capacity, bulk density, rainfall, irrigation, stocking rate, annual N and P fertiliser applied, annual soil nitrogen surplus, tillage, use of forage crops, soil Olsen P, and spatial scale (plot, catchment). Soil order, soil infiltration, anion

storage capacity, and bulk density were estimated from S-map if insufficient data were published for the site.

- A statistical analysis of site characteristics in the database was undertaken to determine those variables that have the greatest statistical relationship with N losses.
- There were several main steps in this analysis:
  - Initially, all available explanatory variables and two response variables were considered individually. The stepwise process removed explanatory variables that were not significant, not required, or had insufficient data.
  - Those explanatory variables that were significant individually were then combined, with their interactions, in a combined model. Two response variables were combined into one.
  - This combined model was reduced stepwise by removing explanatory variables to simplify the model further. Explanatory variables were combined to increase their utility.
  - The process continued until a parsimonious model (i.e. with a minimal number of assumptions, or explanatory variables needed) was found (termed the final model). The residuals were inspected, and no serious breaches of the model assumptions (e.g. distribution, homogeneity of variance) were found.
  - The final model was then disaggregated into individual equations for nitrogen loss for seven land uses for runoff and leaching.
- Initially, total nitrogen (TN) without nitrate-N was considered as the response variable; however, this resulted in a considerable loss of data. The nitrate-N was then investigated. Where the response variable TN was available it was used; otherwise, nitrate-N was used as the new response variable 'Nitrogen'.
- Methods of the stepwise process are described in some detail in the report – methods for initial analyses, followed by the combining and removing variables, then by the final model.
- The process is also summarised by presenting four equations of the fitted model terms (explanatory variables). The equation developed for the final model is then described.

Results

- The spreadsheet database supplied in this project had 155 rows of data, from 56 studies. In the spreadsheet there were 114 rows of measured data and 41 rows of modelled data; 58 rows of data for TN, 124 rows of data for nitrate-N. For land use the spreadsheet rows were: beef 3, cropping 31, dairy 47, deer 8, exotic forestry 13, gorse 4, horticulture 21, native forest 7, sheep 10, and vegetables 11. There were 46 rows with stocking rate data, and 97 rows for annual N fertiliser application.
- For the statistical analysis of N, the database had 142 rows of either TN, nitrate-N, or both observations available.
- An overview of results from the initial analysis are presented. Results and analysis of variance tables when combining explanatory variables are presented with levels of significance.

- Results and the analysis of variance table for the final model explanatory variables are presented with levels of significance. Explanatory variables that were significant in the final model were 'landuse', 'process', 'landuse' 'process' interaction ($P < 0.001$), 'rain_plus_irrig' ('total water'), 'fertiliser N', 'landuse' 'rain_plus_irrig' interaction ($P < 0.01$), and 'rain_plus_irrig' 'process' interaction ($P < 0.05$).
- The final model had 88 observations. This is made up of 22 observations for TN and 66 observations for nitrate-N. Of these, 12 had both TN and nitrate-N values (TN used in model), 10 had TN but no nitrate-N, and 66 had nitrate-N only.
- In the final model, the observations for explanatory variables 'land use' and 'process' (leaching and runoff), dairy had the greatest number of observations for leaching, followed by cropping and horticulture. There were very low numbers of observations (3 or less) for beef, deer, and sheep, with no observations for exotic forestry, gorse, and native forest.
- These results indicate a lack of data in the final model for the combinations of land uses and explanatory variables, hence the results are considered to have low confidence.
- The final model performance had an r^2 of 80%, (with 88 observations). Note that for this type of analysis, the error component is smoothed over all data, and was supported by the residual analysis. The number of observations used in the analysis changes as explanatory variables are removed, so statistics such as the AICs (Akaike information criterion) or adjusted r^2 are not comparable, hence are not presented.
- The final whole model was disaggregated into the equations for the seven land uses (for runoff and leaching) for simplification. This whole final model approach was used because there were insufficient data to estimate individual equations. If performance of individual equations for these seven land uses were evaluated, most would lose any significance, so the approach adopted was to increase the sample size (i.e. using all the final model 88 observations) and assume that the errors were similar in all combinations.

Conclusions

- Our study has evaluated N loss, and although we discuss the limitation of the data, data gaps, and model produced, it is worth noting for future studies that we also capture data for P, sediment, and *E. coli* losses.
- The database contains 56 studies and 155 rows of data. Where explanatory variables associated with N losses were missing, these were obtained from other sources where possible. Data for many explanatory variables were missing, including a low number of observations for some land uses (see below), so further data should be obtained in future.
- This analysis showed N loss (response variable 'Nitrogen' in the final model) could be predicted from the most parsimonious model (and minimum number of assumptions) that used explanatory variables 'total water' (rainfall + irrigation) and annual 'fertiliser N' applied. Both had 88 observations in the final model.
- The final model had good performance with an r^2 of 80%, with 88 observations. Note that the error (uncertainty) is smoothed across all data in the final model to enable better estimates of the error to be made. Therefore, there are no uncertainty

estimates for each of the equations developed to predict N loss for the seven land uses. These results indicate a lack of data in the final model for the combinations of land uses and explanatory variables, hence any predictions from the model are considered to have low confidence.

- A strong caveat is that further data are needed, such as for the land uses and explanatory variables including TN for beef, deer, horticulture, exotic and native forest, sheep, and vegetable land uses. There were no observations for TN for cropping. There were low numbers of observations for nitrate-N for beef, deer, exotic and native forest, gorse, sheep, and vegetable land uses. There were no data for leaching on hill country, and insufficient data on hill country for TN.
- To better predict N losses, future work could address better data for explanatory variables including where proxies were used for missing data, and generate observations of N losses in response to factors such as fertiliser type, organic manures and effluents, supplementary feed and stocking rate. In our analysis it was likely that stocking rate was correlated by other explanatory variables such as fertiliser rate. This report discusses the use of the data, its limitations, and data gaps, which should be considered by users.

Recommendations

- A strong caveat from this study is that users should consider the lack of data in the final model and that the results are considered to have low confidence.
- Another strong caveat from this study is that further data are needed as described in the conclusions. The statistical analysis and equations for the seven land uses could be considered in the future development of future tools (such as an N Risk Index Tool), if the limitations of the study and data gaps, described above, are considered by users.
- Additional work is warranted to obtain more data to fill the data gaps, and for better proxies for missing data, and to generate more observations of N losses for land uses in response to explanatory variables such as soil water retention, fertiliser type, organic manures and effluents, supplementary feed, and stocking rate. Additional data would also enable a cross-validation to be undertaken to understand the ability of the models and suitability of explanatory data to predict N and whether there was over-fitting.
- After considering the database's limitations, it may also be useful to help assess the accuracy and precision of outputs from other tools, such as Overseer[FM] (see example in Appendix 6).

1 Introduction

The Ministry for the Environment (MfE) is investigating diffuse nutrient, sediment, and *E. coli* losses from farmland and the risk to freshwater. Diffuse nutrient discharges from farmland and their risk to freshwater can be assessed using a variety of methods, including modelling.

For example, the OVERSEER farm nutrient modelling tool has been used for nutrient research, farm nutrient planning and regulation. Some concerns have been raised about its use for regulation purposes. These include that the model does not adequately capture overland flow (runoff), which may be an important pathway for losses of non-nitrate forms of nitrogen; and that the OVERSEER model has focused more on leaching of nitrate-nitrogen (PCE 2018; MfE and MPI 2021; MPI 2021).

The Government's response to the findings reported in MPI (2021) included the option of the creation of a new risk index tool, potentially using elements of OVERSEER (MfE and MPI 2021). MfE and MPI (2021) reported in many catchments with a high nitrogen load, nitrate-nitrogen is often the main nitrogen species, so nitrogen lost by overland flow may be less important in areas with low slope and good drainage.

The Ministry's Expert Advisory Group also provided advice on future options. These include a recommendation to develop a Nitrogen (N) Risk Index Tool, as an additional or complementary tool for use (MfE and MPI 2021). Examples of N Risk Indexes are available in the literature (e.g. Heathwaite et al. 2000; McDowell et al. 2002; Drewry et al. 2011). Further details such as examples of weightings and scorings, and potential nutrient risk assessment approaches and nitrogen-risk indexing systems are available in MfE and MPI (2021).

Phase One of the Risk Index Tool will be limited to scoring the risk associated with N losses. There is therefore a need to obtain published literature data that reflect the nitrogen loss likelihood from a combination of farm management and environmental characteristics, by collating a range of New Zealand N loss values and site characteristics such as climate conditions, and farm management factors, from the scientific literature. There is also a need to determine how these site characteristics could influence total nitrogen and nitrate losses.

It is anticipated that the future phases could include P, sediment, and pathogens, so there is also a need to obtain published literature data that reflect their losses by collating a range of New Zealand P, sediment, and *E. coli* loss values, with site characteristics.

2 Objectives

To collate New Zealand published literature values of P, sediment, and *E. coli* annual losses for selected land uses, with site characteristics that may influence their loss, into a database.

To collate predominantly New Zealand published literature values of total nitrogen and nitrate annual losses for selected land uses, with site characteristics that may influence their loss, into a database.

To conduct, in a parsimonious way, a multivariate analysis of characteristics in the database to determine those site characteristics that have the strongest statistical relationship with total nitrogen losses for a range of land uses.

To develop equations for total nitrogen loss for seven land uses for runoff and leaching.

3 Methods

A search of New Zealand-based studies from the published science literature (peer-reviewed) and grey literature (e.g. reports) was conducted. Loss values of N, P, sediment, and *E. coli* were taken from the scientific literature, depending on acceptance criteria used, and expert opinion. Criteria for inclusion are outlined below.

This study had a limited timeframe, so the approach and list of site characteristics were limited to those deemed initially most important for this study and timeframe.

3.1 Criteria for inclusion and method for the characteristics collated

The following criteria were used for inclusion of losses and for site characteristics into the excel database. Details of the method for the site characteristics collated are also presented below:

- Losses were expressed (or could be calculated) as annual yields, e.g., kg N/ha/yr. Where multiple years were identified, we used the mean annual yield for all losses and site characteristics, where appropriate to do so, e.g. consistent land use, treatment, etc.
- Losses spanned at least one year.
- Literature studies were published from 1995 to 2021 for land uses, except forestry, bush, and some sheep/beef pasture studies. For forestry, bush, and some sheep/beef pasture, several earlier studies (back to 1977) were accepted because it was unlikely that management practices, if used, would have changed greatly in those land uses. For the land uses dairy, deer, horticulture, cropping, and vegetables, studies published since and including 1995 were accepted because it is likely they would have used more modern management practices and study techniques, than earlier studies.
- Literature studies were New Zealand-based, or Australian-based from areas with similar climate conditions. Studies were searched for horticulture, cropping and vegetables from Tasmania, and/or Victoria, Australia. These were chosen to increase the number of studies, if considered to have similar climate conditions to those found in New Zealand. There were four observations used from Victoria for cropping land use, and two observations from Tasmania for cropping land use. These six observations were for NO_3-N only. Soil type was not considered for these observations as the New Zealand Soil Classification (Hewitt 2010) was used.

- Losses could be attributed to a dominant land use. Dominant was defined as the largest land use in area and time in the published study.
- Losses were recorded for any one or more of NO_3-N, TN, DRP, TP, sediment, *E. coli.*
- Flow path process was recorded (surface or subsurface (including artificial drainage networks)).
- The spatial scale varied from a 20-cm diameter lysimeter to 30-cm depth up to a second order stream on the same property. Up to 1,000 hectares was used as a general acceptance criterion.
- We recorded data according to their source, namely termed 'measured' observations if the published study had measurements. 'Modelled' estimates were also used. Models that were considered acceptable for this report included APSIM (Agricultural Production Systems sIMulator; Holzworth et al. 2014) and SPASMO (Soil Plant Atmosphere System Model).
- Modelled estimates from literature studies using OVERSEER^{FM} and versions of OVERSEER were excluded because of concerns raised in recent publications (PCE 2018; MPI 2021) as described in the introduction.

The following characteristics were included in the database where available. These were based on expert opinion and characteristics that influence loss of nutrients and sediment.

- Land use (dairy, sheep and beef, deer, forestry, tree crops, dominant arable or horticultural crops within a rotation, vegetables). Several studies on gorse were included.
- Location (latitude and longitude). However, if these were unavailable for the study, or insufficient location details were provided, these were estimated, or an estimate was based on the nearby locality using NZ Topo Map (https://www.topomap.co.nz/).
- Mean annual rainfall (mm).
- Mean annual irrigation (mm).
- Mean annual temperature (°C).
- Mean slope (degrees).
- Soil order (New Zealand soil classification; Hewitt 2010).
- Dominant soil type's infiltration rate (mm/hr), anion storage capacity (ASC; %), bulk density (g/cm^3) from the study. However, if these were unavailable, these were broadly estimated from S-map (https://smap.landcareresearch.co.nz/). Further details are provided below.
- Stocking rate of grazing animals (stock units/ha).
- Annual fertiliser rates (kg N or P/ha/yr). Fertiliser type or use of inhibitors were not considered given the limited scope of this study.
- Mean annual soil nitrogen surplus (kg/ha/yr).
- Use of conventional tillage, minimum tillage, or no tillage.
- Fallow period (yes/no).
- Use of grazed winter forage crops (yes/no).
- Mean soil Olsen P concentration (mg/L) or (mg/kg).
- How the nutrient loss measurement was made, e.g. lysimeter, plot, small catchment.

- Area of lysimeter, plot or small catchment (m^2).
- Stream order (1 or 2).
- Publication details of the study.
- Notes.

If unavailable from the study, soil order and several soil characteristics were estimated from S-map auto-generated factsheets, using location data described above. The S-map tool is available from https://smap.landcareresearch.co.nz/. Note that S-map soil mapping is published at 1:50,000 (regional) scale so only provides a very broad estimate based on the modelling in S-map.

Note that such properties (bulk density, ASC, 'hydrological group') are modelled estimates by MWLR, of inherent soil properties, without the influence of farm management, as indicated on the S-map website. Bulk density values were from the 'topsoil' parameter in S-map factsheets. Soil infiltration rate was from soil 'hydrological group', being four soil classes (A to D) that classify soils according to their vulnerability to runoff, and modelled in S-map. Class A soils have rapid permeability and rainfall, and irrigation is readily transmitted. Class D soils can only infiltrate a small amount of water. Further details on the 'hydrological group' method are available in USDA-NRCS (1986). For this report and analysis, the mid-point of the 'hydrological group' classes were used, namely 1, 3, 6, and >9 mm/h.

Although the term anion storage capacity has been used in the database, it should be noted that the estimates in MWLR S-map factsheets are termed phosphate retention. Further details on phosphate retention used by MWLR are available in Blakemore et al. (1987).

Slope class midpoints were defined as 4°, 12°, 21° >21°, based on slope classes in McDowell and Wilcock (2008). Dairy effluent application was included but not municipal effluent. For use in the statistical analysis, Olsen P (mg/kg) was converted to units of mg/L using soil bulk density.

Where mean annual rainfall and temperature were not available, these were taken from the nearest operating NIWA Cliflo meteorological station approximately 5 to 30 km away from the site.

The database is shown in Appendix 1, with references in Appendix 2.

3.2 Overview of statistical analysis main steps

Analysis of P, sediment and *E. coli* may be conducted in the future so are not included in this report.

This section presents an overview of the steps in the statistical analysis. There were several main steps in this analysis:

- Initially, all available explanatory variables and two response variables were considered individually. The stepwise process removed explanatory variables that were not significant, not required, or had insufficient data.
- Those explanatory variables that were significant individually were then combined, with their interactions, in a combined model. Two response variables were combined into one.
- This combined model was reduced stepwise by removing explanatory variables to simplify the model further. Explanatory variables were combined to increase their utility.
- The process continued until a parsimonious model (i.e. with a minimal number of assumptions, or explanatory variables needed) was found (termed the final model).
- The final model was then disaggregated into individual equations for nitrogen loss for seven land uses for runoff and leaching.

3.3 Initial statistical analysis

An analysis of site characteristics in the database was undertaken to determine those explanatory variables that have the greatest statistical relationship with N losses. Initial response and explanatory variables used in the analysis from the database site characteristics are presented in Table 1.

Table 1. Response and initial explanatory variables used in the nitrogen multivariate analysis

Site characteristic	Variable used in R
Response variables	
TN loss (kg N/ha/yr)	TN
Nitrate-N loss (kg N/ha/yr)	N_N
Explanatory variables	
Dominant land use	landuse
Longitude approx	long
Latitude approx	lat
Measured, modelled	method
Process (note that runoff includes both surface runoff and drainage)	process
Slope	slope_class
Slope (degree)	slope_num
Soil order	soilorder
Dominant soil infiltration rate (mm/hr) class or from S-map	infiltration_class
Soil infiltration (mm/hr)	Infiltration_num
Dominant soil anion storage capacity (%) or from S-map	anionstorage
Dominant soil bulk density (g cm^{-3}) or from S-map	bulkdensity
Rain (mm/yr)	rain
Mean annual irrigation (mm)	irrigation

Site characteristic	Variable used in R
Mean annual temperature (°C)	temp
Stocking rate (stock units/ha)	density
Annual fertiliser rates (kg N/ha/yr)	fertiliser_N
Annual fertiliser rates (kg P/ha/yr)	fertiliser_P
Mean annual soil N surplus (kg/ha/yr)	N_surplus
Conventional, minimum, or no tillage	tillage
Fallow period (yes/no)	fallow
Use of grazed winter forage crops (yes/no)	wintercrop
Mean soil Olsen P concentration (mg/L)	olsen_p
Lysimeter, plot, catchment	scale
Plot or catchment size (m^2)	size

Initially, total nitrogen (TN) without nitrate-N was considered as the response variable, however, this resulted in a considerable loss of data as TN was not recorded in all cases. The nitrate-N was then investigated but there was not an improvement in terms of observations. See the results section which summarises initial results.

The initial full model (with all study characteristics included as individual explanatory variables from Table 1), with their interactions, was reduced stepwise, until a parsimonious (i.e. with a minimal number of assumptions, or explanatory variables needed) model was found (termed the final model; see later section). The residuals were inspected, and no serious breaches of the model assumptions (e.g. distribution, homogeneity of variance) were found.

3.4 Combining variables

This section presents the stepwise process of combining and removing variables. First, explanation is given for combining the variables and their description. Later in this section the process is summarised for the development of the fitted model terms with the explanatory variables used and their interactions. Four equations of the fitted model terms are presented after the description below.

The two response variables from above (TN) and nitrate-N) were then investigated. Where both were recorded, there was a 1:1 relationship so a combination was used for the analysis (response variable 'Nitrogen'). Where the response variable TN was available it was used, otherwise nitrate-N was used as the new response variable 'Nitrogen'.

The resulting N loss response variable 'Nitrogen' was analysed in a simple linear model, the response was found to be highly skewed, so a logarithmic transform was therefore carried out.

A number of explanatory variables that were shown to be not significant, or had insufficient observations, or were not logical (e.g. vegetables on hill country), removed or were combined to increase their utility during the next model run during the stepwise process. Combination of explanatory variables was based on expert opinion and interpretation of the initial analysis results. Several examples of combining initial explanatory variables into new ones during model development are provided below. The land use, ‘gorse’ was removed early on.

For example, for soil order, the explanatory variable (‘soilorder’) was shown to be not useful in the initial model run (see later in the report). Therefore, in the next step, soil orders were then combined to form a new explanatory variable called ‘soilordernew’. Soil orders were then grouped in ‘soilordernew’ based on expert opinion from these initial results. New groupings were created, termed ‘Allophanic’ (15 observations), ‘Pumice’ (11) and all other soil orders (Brown, Gley, Pallic, Recent, Semi-arid, ‘Various’ and Granular) were combined into the new variable termed ‘Other’ (113 observations). The term ‘Various’ soil order refers to sites where there were insufficient details to determine the individual soil order with confidence. For example, all ‘other’ soil orders combined included low numbers of observations for Gley and Semi-arid soils.

Similarly, for ‘lysimeter, plot, catchment’ (explanatory variable ‘scale’), catchment (33 observations) and lysimeter/plot (82 observations) were then combined. For the explanatory variable termed ‘process’ (leaching and runoff), runoff was inputted as a combination of all flow paths where data were provided, i.e. not just surface runoff, so some caution should be applied when interpreting these results. This is also discussed later in the report.

Similarly, for the explanatory variable ‘slope’, the low slope of the sites was retained (‘Flat, 101 observations), whereas the remining slopes (‘Easy’, ‘Rolling’, ‘Steep’) were combined into a new explanatory variable called ‘Hill’ (35 observations).

The explanatory variables ‘rain’ and ‘irrigation’ (Table 1) were summed to create a new ‘total water’ variable (termed ‘rain_plus_irrig’ in the R code). For stocking rate (explanatory variable ‘density’), this was likely to be highly correlated with other variables, so is not included in the final model. Other explanatory variables from Table 1 were also excluded during the stepwise process.

With the above stepwise analysis and combined variables, the model then included the explanatory variables ‘landuse’, ‘process’, ‘flat_hill’, ‘soilordernew’, ‘rain_plus_irrig’, ‘fertiliser_N’ and ‘fertiliser_P’.

During further stepwise analysis, it was found that:

- the slope explanatory variable (‘flat_hill’) was confounded with other variables, so was removed
- ‘Fertiliser_P’ was not significant, so was removed
- interactions with ‘Fertiliser_N’ were not significant, so ‘Fertiliser_N’ remained as a main effect (i.e. interaction effects were removed)
- the new soil order explanatory variable ‘soilordernew’ was not significant, so was removed, and
- any further interactions that were not estimable were removed.

Model terms

This process is summarised below for the development of the fitted model terms with the explanatory variables used and their interactions. Further information and output are in Appendix 4.

> Log(Nitrogen) = landuse*process*flat_hill*soilordernew*rain_plus_irrig*fertiliser_N*fertiliser_P)
>
> Where the response variable is 'Nitrogen'. The explanatory variables are explained above and or Table 1. Note that an asterisk (*) indicates the interaction and all lower terms in the model, i.e. landuse*process includes landuse + process + landuse, process interaction. A colon (:) indicates just the interaction term so the main effects should also be included if it is in the model (Appendix 4).

After the initial process described in the previous section, the procedure started with 'landuse', 'process', 'flat_hill' (slopes flat and combined, as described above), 'soilordernew (soil order with combined categories as described above), 'total water' (rain and irrigation, termed in the R code as 'rain_plus_irrig'), 'fertiliser N', and 'fertiliser P', and all the interaction terms.

Considering the output from the above, the next step was to remove the 'flat_hill' explanatory variable as it was not estimable from the data available. The analysis of variance (Appendix 4) is therefore the same as before, but the slope variable was removed from the model statement:

> Log(Nitrogen) = landuse*process*soilordernew*rain_plus_irrig*fertiliser_N*fertiliser_P)
>
> where the response variable is 'Nitrogen'. The explanatory variables are explained in Table 1 and above.

In the next step, none of the model terms including 'fertiliser_P' were significant so 'fertiliser_P' was removed from the model:

> Log(Nitrogen) = landuse*process*soilordernew*rain_plus_irrig*fertiliser_N)

In the next step, interactions with the soil order ('soilordernew') and 'fertiliser_N' were not significant, so the interactions were removed. Then, soil order (as 'soilordernew') was not significant, so it was completely removed from the model:

> Log(Nitrogen) = landuse + process + rain_plus_irrig + landuse:process + landuse:rain_plus_irrig + process:rain_plus_irrig + fertiliser_N)

In the resulting model all terms were significant. The interactions between 'landuse' and 'process', 'landuse' and 'rain_plus_irrig' ('total water') and 'process' were all significant so were retained.

3.5 Final model

The final model is:

> Log(Nitrogen) = landuse + process + rain_plus_irrig + landuse:process + landuse:rain_plus_irrig + process:rain_plus_irrig + fertiliser_N
>
> where the response variable is 'Nitrogen' (in kg N/ha/yr). The explanatory variables 'land use', 'process', and 'fertiliser_N' are explained in Table 1. Note 'total water' used in the next section is the 'rain_plus_irrig' term in the R code (units mm/yr). The explanatory variable 'fertiliser N' is annual fertiliser rate (in kg N/ha/yr; Table 1). The colon (:) indicates the interaction between the explanatory variables.

The final model analysis results are presented in the results section, including the disaggregated equations for individual land uses. Note that where the linear equations are in log-space, a bias correction (Miller 1984) needs to be applied, with further details presented in the results section.

Further details are presented in Appendix 3 (R code) and Appendix 4 (statistical output). All statistical analyses were carried out using the statistical package 'R' version 4.1.0 (R Core Team 2021).

4 Results

4.1 Spreadsheet database

The spreadsheet database (Appendix 1) supplied from this project contains the data used in the analysis and full details of publication references (Appendix 2).

In summary, in the spreadsheet database there were 56 studies and 155 rows of data values included. Note that there were some items and studies with data not available. In the spreadsheet there were 114 rows of measured data and 41 rows of modelled data; 58 rows of data for TN, 124 rows of data for nitrate-N (29 rows had both TN and nitrate-N). For land use, rows were: beef 3, cropping 31, dairy 47, deer 8, exotic forestry 13, gorse 4, horticulture 21, native forest 7, sheep 10, and vegetables 11. For 'process', there were 104 rows for 'leaching' and 51 for 'runoff'. There were 46 rows with stocking rate data, and 97 rows for annual N fertiliser application. There were 33 rows for catchment studies, 56 for lysimeter studies and 29 for plot studies. By region, rows were Canterbury 24, Otago 19, Waikato 18, Manawatu 13, Bay of Plenty 14, Southland 12, Gisborne 13, Wellington 7, and a smaller number for other regions. The six rows in total from Victoria and Tasmania for cropping land use were for nitrate-N only.

In summary, soil orders in the spreadsheet database included Allophanic (15 rows), Brown (25 rows), Gley (4 rows), Granular (6 rows), Pallic (42 rows), Pumice (14 rows), Recent (7 rows) and Semi-arid (3 rows). There were 13 rows with soil order not available, and 25 rows of data termed 'various' which represented studies where there were multiple soil orders or multiple locations in a study, so the dominant soil order could not be determined.

For this report, P, sediment and *E. coli* annual losses for selected land uses, were collated into a spreadsheet database for future evaluation. In summary, in the spreadsheet database there were 69 rows of annual losses for TP, and 30 rows for DRP; 22 rows of sediment losses and 11 rows of *E. coli* annual losses included. There were 65 rows for annual P fertiliser application included.

For the statistical analysis of TN and nitrate-N, the database had 47 rows of TN observations available, and 124 nitrate-N observations (i.e. some with both). The database had 142 rows of either TN, nitrate-N, or both observations available. Of those 142 rows of observations of TN or nitrate-N, the database had:

- Land use: dairy (37), beef (2), cropping (31), deer (7), exotic forestry (13), gorse (4), horticulture (20), native forest (7), pasture (5), sheep (7) and vegetables (9 observations)
- Measured (102), modelled (40)
- Leaching (100), runoff (42)
- N fertiliser application (88).

4.2 Initial statistical evaluation

All available explanatory variables from Table 1 were considered individually and those variables that were significant individually were then combined, with their interactions, in a combined model, which was then simplified by removing model terms until a parsimonious model with only significant terms was attained.

Examples of combined explanatory variables were presented in the methods.

This section very broadly summarises results for the two individual response variables from Table 1, prior to them being combined into one response variable, as described previously.

For the initial individual evaluation of total nitrogen (response variable 'TN') the results and/or decisions were:

- land use – several land uses were significantly different. Process – significant difference.
- method – not significant. Rainfall+irrigation were combined (later termed 'total water') – not significant.
- 'slope_class' – 'Flat' significantly different, hence decision to combine three classes.
- soil order – several significantly different but low observation numbers so combining was then considered. Further details are in the methods.
- infiltration class C was significantly different from class A. For 'scale' there was a small difference, but this was not considered useful. Anion storage, bulk density,

temperature, 'density', N fertiliser application, P fertiliser application, Olsen P, size, and winter crop were all not significant. For N surplus, tillage, and fallow, there were insufficient data to draw conclusions.

For the initial analysis of nitrate-N (response variable 'N_N') the results and/or decisions were:

- land use – several land uses were significantly different. Process – significant difference. Scale – highly significant.
- 'slope_class' – Flat significantly different, hence decision to combine three classes.
- soil order – several significantly different but some low observation numbers. Combining soils was then considered. Further details and examples are in the methods.
- method, infiltration class, anion storage, bulk density, temperature, 'density', N surplus, size, winter crop, were all not significant.
- N fertiliser application – highly significant. P fertiliser application – significant. Rainfall was significant. Rainfall+irrigation were combined (later termed 'total water') and was marginally significant so considered for further evaluation.
- For Olsen P, tillage and fallow, there were insufficient data to draw conclusions.

Based on the initial evaluation above, during the subsequent stepwise process, response variables total nitrogen and nitrate-N (response variables 'TN' and 'N_N') were then combined. As described in the methods, where the response variable TN was available it was used, otherwise nitrate-N was used as the response variable 'Nitrogen'. The substitution of TN loss by nitrate-N loss in the methods section was supported by the data. The log (TN loss) and log (nitrate-N loss) were highly correlated ($r = 0.85$; slope = 0.966 and was not significantly different to 1). The actual vs predicted values from the final model shows most points were close to the 1:1 line (see Appendix 5).

During the early stages of the process several knowledge and data gaps were identified:

- There were low numbers of observations for TN (<10 observations) for beef, deer, horticulture, exotic and native forest, sheep and vegetable land uses. There were no observations for TN for cropping.
- There were low numbers of observations for nitrate-N (<10 observations) for beef, deer, exotic and native forest, gorse, sheep and vegetable land uses.
- For N surplus, Olsen P, tillage and fallow, there were insufficient data.
- There were no data for leaching on hill country, and insufficient data on hill country for TN, so these interactions were removed.
- Interactions with fertiliser N were not significant so was retained as a main effect. Interactions with soil order were not significant so was retained as a main effect.

4.3 Combining variables

This section presents an example of the analysis of variance results when combining and/or removing explanatory variables, as described in the stepwise process (methods section 3.3). Further analyses of variance outputs are presented in Appendix 4.

The analysis of variance results for the explanatory variables in the model are presented in Table 2. This table presents the results from the procedure as some explanatory variables were combined, starting with 'landuse', 'process', 'flat_hill', 'soilordernew, 'rain_plus_irrig' ('total water'), 'fertiliser N', and 'fertiliser P', and all the interaction terms. A colon (:) indicates the interaction term.

Explanatory variables that were significant were 'landuse' and 'process' (P < 0.001), 'soilordernew', rain_plus_irrig' ('total water'), 'fertiliser N', 'landuse' 'process' interaction (P < 0.01), and 'landuse' 'process' 'fertiliser_N' interaction (P < 0.05).

Table 2. Analysis of variance results as explanatory variables were combined

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Significance
landuse	5	35.056	7.0112	21.8343	0.000388	***
process	1	27.731	27.7311	86.36	3.46E-05	***
soilordernew	1	4.004	4.0045	12.4708	0.00958	**
rain_plus_irrig	1	7.631	7.6311	23.7648	0.001804	**
fertiliser_N	1	3.989	3.9887	12.4216	0.00967	**
fertiliser_P	1	0.589	0.5891	1.8346	0.217677	
landuse:process	2	13.458	6.7288	20.9547	0.001109	**
landuse:soilordernew	1	0.977	0.9771	3.0429	0.124606	
landuse:rain_plus_irrig	4	5.075	1.2688	3.9514	0.054885	.
soilordernew:rain_plus_irrig	1	0.439	0.4388	1.3664	0.280692	
landuse:fertiliser_N	4	0.368	0.0921	0.2867	0.877726	
process:fertiliser_N	1	0.001	0.0008	0.0024	0.962611	
soilordernew:fertiliser_N	1	0.068	0.0676	0.2105	0.66029	
rain_plus_irrig:fertiliser_N	1	0.373	0.3732	1.1622	0.316747	
landuse:fertiliser_P	3	3.973	1.3244	4.1244	0.055924	.
process:fertiliser_P	1	0.638	0.6378	1.9862	0.201585	
rain_plus_irrig:fertiliser_P	1	0.436	0.4363	1.3588	0.281928	
fertiliser_N:fertiliser_P	1	0.017	0.0169	0.0525	0.825248	
landuse:process:fertiliser_N	1	2.857	2.8574	8.8986	0.020426	*
landuse:rain_plus_irrig:fertiliser_N	2	0.3	0.1498	0.4666	0.645328	
landuse:rain_plus_irrig:fertiliser_P	1	1.217	1.2167	3.7892	0.092634	.
landuse:fertiliser_N:fertiliser_P	2	0.53	0.2651	0.8257	0.476474	
rain_plus_irrig:fertiliser_N:fertiliser_P	1	0.559	0.559	1.7409	0.228537	
landuse:rain_plus_irrig:fertiliser_N:fertiliser_P	1	0.832	0.8321	2.5914	0.151481	
Residuals	7	2.248	0.3211			

Notes: explanatory variables are described earlier. Significance: '.' P < 0.1; * P < 0.05; ** P < 0.01; *** P < 0.001. A colon (:) indicates the interaction term.

4.4 Final model

This section presents the final model analysis of variance results, an overview and summary graphs, and a summary of observation numbers and model performance.

4.4.1 Analysis of variance

The analysis of variance results for the final model presented in Table 3. (In this table 'landuseD' is a categorical explanatory variable used to develop the land use equations later and is interpreted here as 'landuse').

Explanatory variables that were significant were 'landuse', 'process', 'landuse' 'process' interaction (P < 0.001), 'rain_plus_irrig' ('total water'), 'fertiliser N', 'landuse' 'rain_plus_irrig' interaction (P < 0.01), and 'rain_plus_irrig' 'process' interaction (P < 0.05).

Table 3. Analysis of variance results of final model

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	Significance
landuseD	6	47.562	7.927	19.3256	3.22E-13	***
rain_plus_irrig	1	4.746	4.7465	11.5717	0.001103	**
process	1	25.584	25.5837	62.372	2.59E-11	***
fertiliser_N	1	4.075	4.0754	9.9358	0.002375	**
landuseD:process	2	26.505	13.2525	32.309	1.05E-10	***
landuseD:rain_plus_irrig	4	7.182	1.7955	4.3774	0.003205	**
rain_plus_irrig:process	1	1.887	1.8875	4.6016	0.035364	*
Residuals	71	29.123	0.4102			

Notes: explanatory variables are described earlier. Significance: * P < 0.05; ** P < 0.01; *** P < 0.001. A colon (:) indicates the interaction term.

4.4.2 Final model overview

Several boxplots are presented to provide an overview of the data in the final model. A summary of the N loss by land use, and by explanatory variable 'method' (i.e. 'measurement' or 'modelled') is shown in Figure 1.

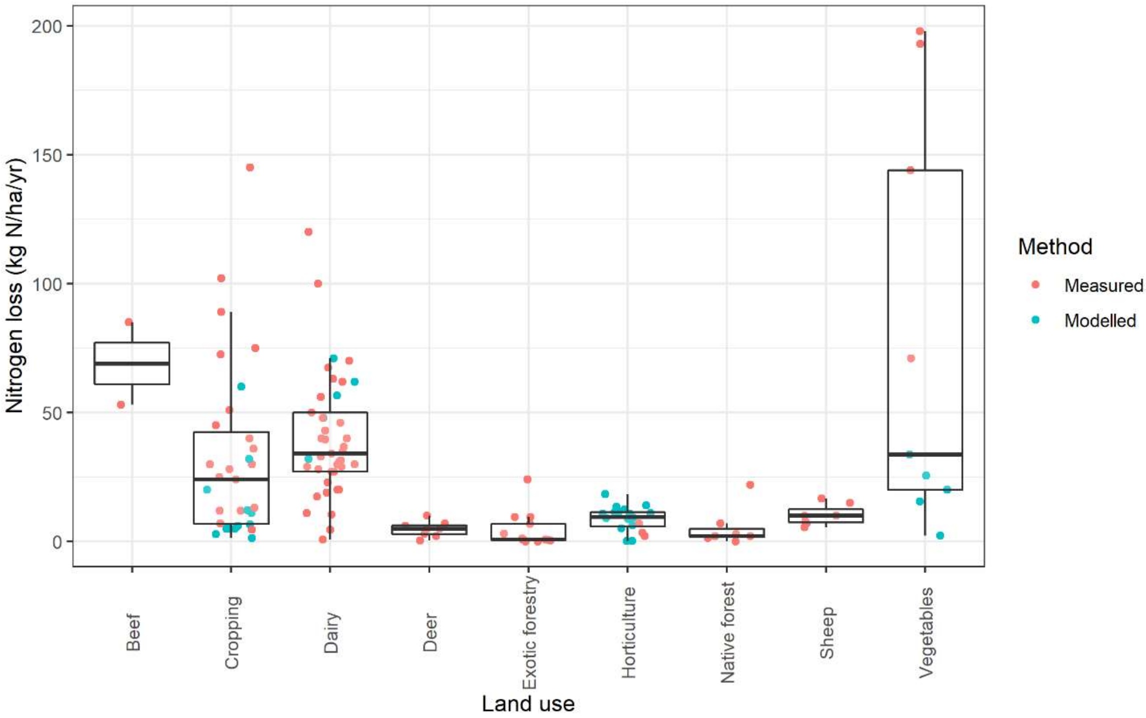

Figure 1. Response variable 'Nitrogen' loss (kg N/ha/yr) by explanatory variables 'land use', and 'method' ('measurement' or 'modelled'). Boxes represent the 25–75th percentiles.

A summary of the response variable 'Nitrogen' loss values observed for explanatory variable 'process' (leaching and runoff), and land use is shown in Figure 2.

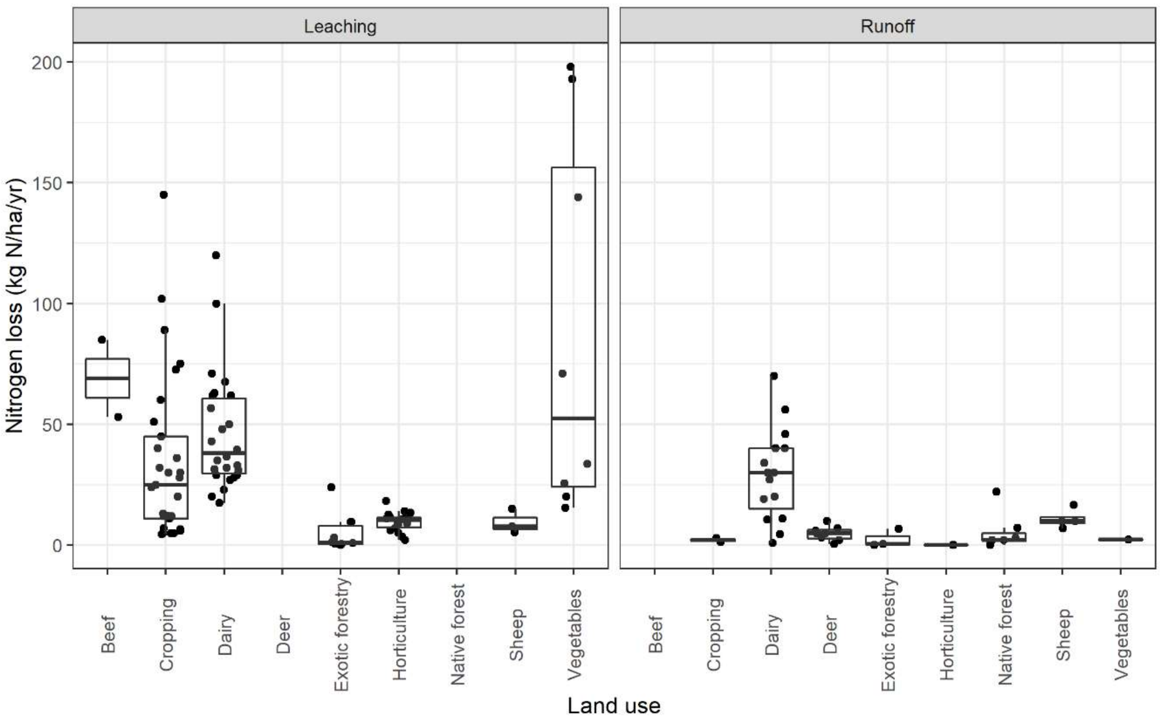

Figure 2. Response variable 'Nitrogen' loss (kg N/ha/yr) for explanatory variables 'process' (leaching and runoff) and 'land use'. Boxes represent the 25–75th percentiles.

4.4.3 Observations and model performance

The final model had 88 observations. This is made up of 22 observations for TN and 66 observations for nitrate-N. Of these, 12 had both TN and nitrate-N values (TN used in model), 10 had TN but no nitrate-N, and 66 had nitrate-N only.

In the final model the observations for explanatory variables 'land use' and 'process' (leaching and runoff) are summarised in Table 4. Dairy had the greatest number of observations for leaching, followed by cropping and horticulture. There were very low numbers of observations (3 or less) for beef, deer, and sheep, with no observations for exotic forestry, gorse, and native forest.

Table 4. Summary of the number of observations in the final model for the explanatory variables 'land use' and 'process' (leaching and runoff)

Land use	Leaching	Runoff	Total
Dairy	22	13	35
Beef	2	0	2
Cropping	20	0	20
Deer	0	3	3
Exotic forestry	0	0	0
Gorse	0	0	0
Horticulture	14	2	16
Native forest	0	0	0
Sheep	3	0	3
Vegetables	8	1	9
Total	**69**	**19**	**88**

In the final model for the explanatory variable 'method', there were 52 observations from 'measurements', and 36 from 'modelled' observations (total is 88). The explanatory variables 'total water' and 'fertiliser N' (see equations section) both have 88 observations in the final model.

The observed values and the response variable 'Nitrogen' loss predicted by the model for explanatory variables 'total water' (rainfall + irrigation), 'process' (leaching and runoff), and 'land use' are shown in Figure 3. Note that Figure 3 shows the response variable 'Nitrogen' loss (and therefore does not show N fertiliser inputs).

As above, Figure 3 shows low numbers of observations, for leaching for beef, deer (nil), native forest (nil), sheep land uses, with dairy land use having most observations. Figure 3 shows low numbers of observations for runoff for most land uses, with dairy land use having the most observations.

These results indicate a lack of data in the final model for the combinations of land uses and explanatory variables, hence the results are considered to have low confidence.

The final model performance had an r^2 of 80%, (with 88 observations). Note that for this type of analysis, the error component is smoothed over all data, and was supported by the residual analysis, as described earlier. Also note the number of observations used in the analysis changes, as explanatory variables are removed, so statistics such as the AICs (Akaike information criterion) or adjusted r^2 are not comparable, hence are not presented.

Similarly, if performance of individual equations for land use (next section) were evaluated, most would lose any significance, so the approach adopted in the methods was to increase the sample size (i.e. using all the final model 88 observations) and assume that the errors were similar in all combinations.

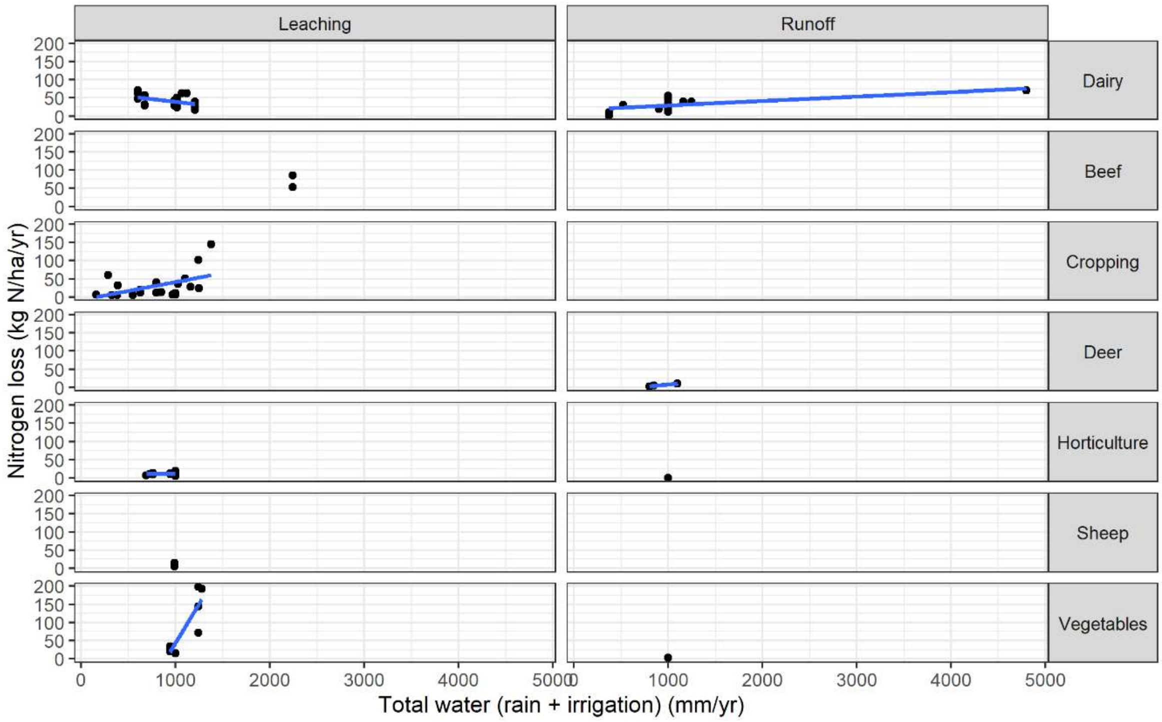

Figure 3. Response variable 'Nitrogen' loss (kg N/ha/yr) predicted by the model (blue line) for explanatory variables 'total water' (i.e. rainfall + irrigation), 'process' (leaching and runoff), and for 'land use'.

Further information is in Appendix 5 which presents the response variable 'Nitrogen' loss (kg N/ha/yr) predicted by the model for the explanatory variable 'Fertiliser N'.

4.5 Equations per land use for leaching and runoff

The equations for the models for seven land use for leaching and runoff are presented in this section.

The final whole model has been disaggregated into equations for simplification. For example, for a land use (e.g. dairy and leaching), the other land uses for leaching will be zero, but the components form part of the whole final model. Note that the whole final model approach was used, as described above and in the methods because there was insufficient data to estimate individual equations.

Note that the error (uncertainty) is smoothed across all data in the final model to enable better estimates of the error to be made. Therefore, there are no uncertainty estimates for each equation below.

The equations are presented in log-space as linear equations to assist the reader. Note that to calculate the estimate of N loss, exponentiate the $\log(\text{Nitrogen})$ (i.e. $e^{\log(\text{Nitrogen})}$), and then apply a bias correction. For bias correction multiply by 1.23.

Further details of the coefficients, standard errors, and significance for the explanatory variables in the final model are presented in Table 5.

For the equations below, the two explanatory variables are 'total water' (rainfall + irrigation, as mm/yr), and 'fertiliser' is annual nitrogen fertiliser rate (kg N/ha/yr) (explanatory variable 'fertiliser_N'). Both have 88 observations in the final model.

For Dairy / Leaching

$$\log(Nitrogen) = 4.13 - 0.001 \times total\ water + 0.003 \times fertiliser$$

For Dairy / Runoff

$$\log(Nitrogen) = 2.10 + 0.001 \times total\ water + 0.003 \times fertiliser$$

For Beef / Leaching

$$\log(Nitrogen) = 6.31 - 0.001 \times total\ water + 0.003 \times fertiliser$$

For Beef / Runoff

$$\log(Nitrogen) = 4.28 + 0.001 \times total\ water + 0.003 \times fertiliser$$

For Cropping / Leaching

$$\log(Nitrogen) = 1.30 + 0.001 \times total\ water + 0.003 \times fertiliser$$

For Cropping / Runoff

$$\log(Nitrogen) = -0.73 + 0.003 \times total\ water + 0.003 \times fertiliser$$

For Deer / Leaching

$$\log(Nitrogen) = -0.80 + 0.003 \times total\ water + 0.003 \times fertiliser$$

For Deer / Runoff

$$\log(Nitrogen) = -2.83 + 0.005 \times total\ water + 0.003 \times fertiliser$$

For Horticulture / Leaching

$$\log(Nitrogen) = 2.56 + 0.0003 \times total\ water + 0.003 \times fertiliser$$

For Horticulture / Runoff

$$\log(Nitrogen) = -3.60 + 0.002 \times total\ water + 0.003 \times fertiliser$$

For Sheep / Leaching

$$\log(Nitrogen) = 2.99 - 0.001 \times total\ water + 0.003 \times fertiliser$$

For Sheep / Runoff

$$\log(Nitrogen) = 0.96 + 0.001 \times total\ water + 0.003 \times fertiliser$$

For Vegetables / Leaching

$$\log(Nitrogen) = -2.42 + 0.005 \times total\ water + 0.003 \times fertiliser$$

For Vegetables / Runoff

$$\log(Nitrogen) = -6.41 + 0.007 \times total\ water + 0.003 \times fertiliser$$

Table 5. Summary of the coefficients for slope estimate, standard error, t values and significance for explanatory variables in the model. Note that the base model is dairy, flat and leaching. Note that the coefficients are presented in log-space (rather than as exponents on the original scale)

	Slope estimate	Standard error	t value	Pr(>|t|)	Significance
(Intercept)	4.131	0.663	6.23	0	***
landuseDBeef	2.177	1.029	2.115	0.038	*
landuseDCropping	-2.836	0.752	-3.77	0	***
landuseDDeer	-4.934	2.623	-1.881	0.064	.
landuseDHorticulture	-1.568	1.491	-1.052	0.297	
landuseDSheep	-1.143	0.408	-2.805	0.006	**
landuseDVegetables	-6.555	1.82	-3.603	0.001	**
rain_plus_irrig	-0.001	0.001	-1.488	0.141	
processRunoff	-2.027	0.714	-2.839	0.006	**
fertiliser_N	0.003	0.001	3.238	0.002	**
landuseDHorticulture:processRunoff	-4.132	0.562	-7.353	0	***
landuseDVegetables:processRunoff	-1.96	0.734	-2.669	0.009	**
landuseDCropping:rain_plus_irrig	0.002	0.001	3.034	0.003	**
landuseDDeer:rain_plus_irrig	0.004	0.003	1.451	0.151	
landuseDHorticulture:rain_plus_irrig	0.001	0.002	0.422	0.674	
landuseDVegetables:rain_plus_irrig	0.006	0.002	3.83	0	***
rain_plus_irrig:processRunoff	0.002	0.001	2.145	0.035	*

Notes: 'D' is dairy; note that 'landuseD' is a categorical variable with a number of land uses; 'rain_plus_irrig' is rainfall + irrigation ('total water'). Rainfall plus irrigation is in mm/yr. Fertiliser N is annual fertiliser rate in (kg N/ha/yr). Significance: '.' P < 0.1; * P < 0.05; ** P < 0.01; *** P < 0.001.

5 Discussion and limitations

This section discusses some of the limitations and some cautions for use of the results.

5.1 Limitations of the data

We used several proxies where data were not available in the published studies. The reader should be aware of their limitations.

If the soil data (soil order, infiltration, anion storage capacity and bulk density) were unavailable from the literature studies, these site soil characteristics were estimated from S-map auto-generated factsheets. However, given that S-map soil mapping is published at 1:50,000 (regional) scale, such estimates only provide a coarse spatial estimate based on the modelling in S-map. It is possible that soil order did not feature in the final model because of this broad or poor site characterisation. For example, studies may have listed multiple soil orders or multiple locations, so the dominant soil order could not be determined.

Other data, such as bulk density and anion storage capacity, have been found to be useful when evaluating nutrient losses (McDowell 2017). Where not reported, these data were obtained from S-map, which in turn derives them from pedo-transfer functions (McNeill et al. 2018). MWLR advises that they are updating the pedo-transfer functions for bulk density and anion storage capacity. Recent updates of water retention modelling, soil depth and available water capacity (shown as profile available water) in S-map may prove useful as these could be used to estimate N leaching.

Our data represent a range of scales. The concept of critical source areas (McDowell et al. 2009) outlines that most losses come from a minority of an area. In farm and catchment studies, the full impact of critical source areas may be missed if important. We have no way of determining the magnitude of importance for critical source areas; hence, our assumption is that the data collected at these scales are representative not only of, for example, the land use, but also of the frequency and magnitude of critical source areas on N losses. It is worthy of noting that data for plot (viz. lysimeter) scales do not suffer this issue as they are either in (e.g. wintering block) or outside of a critical source area.

5.2 Gaps in the data

Several data gaps were identified in this study, which included a limited number of studies and observations for some land uses (partly caused by filtering out old data that may not have the same farm systems or quality control checks as recent data). This resulted in some missing combinations of hydrological processes and explanatory variables in the final model development.

Gaps were identified during the model development, including for runoff for beef, cropping, deer, sheep land use, and 'total water' for beef and sheep land use. Low numbers of observations were also identified for leaching for beef, deer, and sheep land uses. It was also identified that studies with modelled loss results tended to have lower N loss values for vegetables, than for studies with measured values.

The model cannot be used in process by factor combinations within those gaps. Gaps have the potential to be true or false depending on the representativeness of the data and whether a factor does control N loss. However, over several decades of research we outline factors as likely gaps, implying their influence to be true, but not well represented in the dataset and hence cannot be included in the final model:

- Fertiliser type, or fast- and slow-release products. However, it should be noted that fertiliser rate was a sensitive explanatory variable in the model
- Organic manures, effluents, and composts
- Nitrification inhibitors. However, this could be estimated by discounting the rate of fertiliser applied
- Supplementary feed
- Stocking (e.g. Roche et al. 2016). However, this is likely to be correlated and picked up by other variables such as fertiliser rate.

5.3 Use of the data

Given the nature of the study, the final model performed with an r^2 of 80%, with 88 observations. However, a strong caveat is that further data are needed, such as for a number of land uses and explanatory variables as identified earlier in this report. Further data would be valuable.

An independent dataset, if available, could be used to check the model for validation purposes, but this was not undertaken due to the limitations of the data and the limited observation numbers available, and was beyond the scope of this study. If a suitable data set was available for validation, it would be very useful, but it could also be included in the data to improve the predictions. An independent dataset would, however, need to include at least some of the land uses and explanatory variables, plus a range of 'total water' and 'fertiliser N' values. It appears unlikely that this exists.

To aid the reader, the equations are presented as linear equations in log-space. For future use, the instructions presented in the results section when converting equations from log-space, should be adhered to. It is also worth noting that we recorded leaching as losses pertaining to sub-surface flow, but runoff was recorded as the combination of surface runoff and interflow, which may also include sub-surface flow. Where surface runoff (only) data are available, these are noted in the database with an asterisk.

The coefficients could be interpreted to help inform future research or tools, such as a relative weighting for factors contributing to N loss, in the future development of the Nitrogen Risk Index Tool, or other models such as Overseer^FM (see example Appendix 6). However, considerable caution is advised given the nature of this study, the results, the limitations of the data, and the lack of data such as for the land uses and explanatory variables identified earlier.

6 Conclusions

Our study has evaluated N loss, and although we discuss the limitation of the data, data gaps and model produced, it is worth noting for future studies that we also capture data for P, sediment, and *E. coli* losses.

The database contains 56 studies and 155 rows of data. Where explanatory variables associated with N losses were missing, these were obtained from other sources where possible. Data for many explanatory variables were missing, including a low number of observations for some land uses (see below), so further data should be obtained in future.

This analysis showed N loss (response variable 'Nitrogen' in the final model) could be predicted from the most parsimonious model (and minimum number of assumptions) that used explanatory variables 'total water' (rainfall + irrigation) and annual 'fertiliser N' applied. Both had 88 observations in the final model.

The final model had good performance with an r^2 of 80%, with 88 observations. Note that the error (uncertainty) is smoothed across all data in the final model to enable better estimates of the error to be made. Therefore, there are no uncertainty estimates for each of the equations developed to predict N loss for the seven land uses. These results indicate a lack of data in the final model for the combinations of land uses and explanatory variables, hence any predictions from the model are considered to have low confidence.

A strong caveat is that further data are needed, such as for land uses and explanatory variables, including TN for beef, deer, horticulture, exotic and native forest, sheep, and vegetable land uses. There were no observations for TN for cropping. There were low numbers of observations for nitrate-N for beef, deer, exotic and native forest, gorse, sheep, and vegetable land uses. There were no data for leaching on hill country, and insufficient data on hill country for TN.

To better predict N losses, future work could address better data for explanatory variables, including where proxies were used for missing data, and generate observations of N losses in response to factors such as fertiliser type, organic manures and effluents, supplementary feed, and stocking rate. In our analysis it was likely that stocking rate was correlated by other explanatory variables such as fertiliser rate. This report discusses the use of the data, its limitations, and data gaps, which should be considered by users.

7 Recommendations

A strong caveat from this study is that users should consider the lack of data in the final model and that the results are considered to have low confidence.

Another strong caveat from this study is that further data are needed, as described in the conclusions. The statistical analysis and equations for the seven land uses could be considered in the future development of future tools (such as an N Risk Index Tool), if the limitations of the study and data gaps, described above, are considered by users.

Additional work is warranted to obtain more data to fill the data gaps, and for better proxies for missing data, and to generate more observations of N losses for land uses in response to explanatory variables such as soil water retention, fertiliser type, organic manures and effluents, supplementary feed, and stocking rate. Additional data would also enable a cross-validation to be undertaken to understand the ability of the models and suitability of explanatory data to predict N and whether there was over-fitting.

After considering the database's limitations, it may also be useful to help assess the accuracy and precision of outputs from other tools, such as Overseer^{FM} (see example in Appendix 6).

8 Acknowledgements

We thank Anne Austin for editing and Kate Boardman for assistance with formatting.

9 References

Blakemore LC, Searle PL, Daly BK 1987. Methods for chemical analysis of soils New Zealand Soil Bureau Scientific Report 80. 103 p.

Correndo AA, Hefley TJ, Holzworth DP, Ciampitti IA 2021. Revisiting linear regression to test agreement in continuous predicted-observed datasets. Agricultural Systems 192: 103194.

Drewry JJ, Newham LTH, Greene RSB 2011. Index models to evaluate the risk of phosphorus and nitrogen loss at catchment scales. Journal of Environmental Management 92: 639–649.

Heathwaite L, Sharpley A, Gburek W 2000. A conceptual approach for integrating phosphorus and nitrogen management at watershed scales. Journal of Environmental Quality 29(1): 158–166.

Hewitt A 2010. New Zealand soil classification. Third edn. Lincoln, New Zealand: Landcare Research. 136 p.

Holzworth DP, Huth NI, deVoil PG, Zurcher EJ, Herrmann NI, McLean G, Chenu K, van Oosterom EJ, Snow V, Murphy C and others 2014. APSIM – Evolution towards a new generation of agricultural systems simulation. Environmental Modelling & Software 62: 327–350.

McDowell RW 2017. Does variable rate irrigation decrease nutrient leaching losses from grazed dairy farming? Soil Use and Management 33(4): 530–537.

McDowell RW, Sharpley AN, Kleinman PJA 2002. Integrating phosphorus and nitrogen decision management at watershed scales. Journal of the American Water Resources Association 38(2): 479–491.

McDowell RW, Srinivasan MS 2009. Identifying critical source areas for water quality: 2. Validating the approach for phosphorus and sediment losses in grazed headwater catchments. Journal of Hydrology 379: 68–80.

McDowell RW, Wilcock RJ 2008. Water quality and the effects of different pastoral animals. New Zealand Veterinary Journal 56(6): 289–296.

McNeill SJ, Lilburne LR, Carrick S, Webb TH, Cuthill T 2018. Pedotransfer functions for the soil water characteristics of New Zealand soils using S-map information. Geoderma 326: 96–110.

Miller DM 1984. Reducing transformation bias in curve fitting. The American Statistician 38(2): 124–126.

MPI 2021. Overseer whole-model review Assessment of the model approach. Prepared for the Ministry for Primary Industries and the Ministry for the Environment by the Science Advisory Panel MPI Technical Paper no: 2021/12. Wellington: Ministry for Primary Industries. 118 p.

PCE 2018. Overseer and regulatory oversight: Models, uncertainty and cleaning up our waterways. Wellington: Parliamentary Commissioner for the Environment. 138 p.

R Core Team 2021. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Roche J, Ledgard S, Sprosen M, Lindsey S, Penno J, Horan B, Macdonald K 2016. Increased stocking rate and associated strategic dry-off decision rules reduced the amount of nitrate-N leached under grazing. Journal of Dairy Science 99: 5916–5925.

USDA-NRCS 1986. Urban hydrology for small watersheds TR-55. Washington, DC: United States Department of Agriculture Natural Resources Conservation Service. 164 p.

Appendix 1 – spreadsheet database

The spreadsheet database was supplied electronically as an output of this study.

Screen images of the spreadsheet are on the following pages. Where surface runoff (only) data are available, these are noted in the database with an asterisk. References are listed in Appendix 2. Due to page space constraints, all text within some cells may not be visible, so the electronic version supplied should be referred to. Comments have been excluded from the spreadsheet screen images below due to their length.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Dominant land use	TN loss (kg N/ha/yr)	Nitrate-N loss (kg N/ha/yr)	DRP loss (kg P/ha/yr)	TP loss (kg P/ha/yr)	Sediment loss (kg/ha/yr)	E. coli loss (MPN ha^{-1})	Longitude approx	Latitude approx	Location	Region	Management	Measured, modelled	Model type	Duration (yr)
2	Exotic forestry	6.76	2.55	0.26	1.33	2632		175.154	-37.78	Whatawhata	Waikato	Pine agro-forest, m	Measured		2
3	Sheep	16.6	6.89	0.28	2.4	2100		175.25	-37.78	Whatawhata	Waikato	Sheep pasture with	Measured		2
4	Exotic forestry		3					176.2	-38.6	Purukohukoi	Bay of Plenty	Pine	Measured		
5	Exotic forestry		0.5					176.2	-38.6	Purukohukoi	Bay of Plenty	Pine, post clear cu	Measured		
6	Exotic forestry		24					176.2	-38.6	Purukohukoi	Bay of Plenty	Pine, pre-clearing	Measured		
7	Exotic forestry		0.8					176.347	-38.06	Rotorua	Bay of Plenty	20 yr old pine, 1.7 y	Measured		
8	Exotic forestry		0.7					176.347	-38.06	Rotorua	Bay of Plenty	7 yr old pine, 1.7 yr	Measured		
9	Exotic forestry	9.5	4.5					175.617	-40.385	Massey	Manawatu-W	Pine. Highest after	Measured		9
10	Exotic forestry		9.5					175.617	-40.385	Massey	Manawatu-W	Pine. Highest after	Measured		9
11	Native forest	2.07	0.55	0.27	0.58	320		175.25	-37.78	Whatawhata	Waikato	Podocarp-hardwo	Measured		2
12	Native forest	7.1			0.6			170.658	-43.016	Dunlop catc	West Coast	Native forest and s	Measured		1.1
13	Native forest	22	1.2	0.71	2.1			170.707	-42.955	Fergusson c	West Coast	Native forest and s	Measured		1.1
14	Cropping		89					172.475	-43.635		Canterbury	Mixed, early plough	Measured		
15	Cropping		30					172.475	-43.635		Canterbury	Mixed, autumn plo	Measured		
16	Cropping		75					172.475	-43.635		Canterbury	Lentils, beans, pea	Measured		
17	Cropping		25					172.475	-43.635		Canterbury	Lupins	Measured		
18	Cropping		30					172.475	-43.635		Canterbury	Rape	Measured		
19	Cropping		45					172.475	-43.635		Canterbury	Lupins	Measured		
20	Cropping		12					176.344	-40.024		Hawke's Bay	Mixed site, irrigated	Measured		2
21	Cropping		6.6		0			177.966	-38.639	Poverty Bay	Gisborne	Maize	Modelled	Spasmo	
22	Cropping		10.9		0			177.966	-38.639	Poverty Bay	Gisborne	Squash, irrigated	Modelled	Spasmo	
23	Cropping		1.3		2.4			177.966	-38.639	Poverty Bay	Gisborne	Maize	Modelled	Spasmo	
24	Cropping		2.8		2.9			177.966	-38.639	Poverty Bay	Gisborne	Squash, irrigated	Modelled	Spasmo	
25	Vegetables		71					174.863	-37.204	Pukekohe	Auckland	Winter spinach cor	Measured		3
26	Vegetables		144					174.863	-37.204	Pukekohe	Auckland	Winter spinach N a	Measured		3
27	Vegetables		198					174.863	-37.204	Pukekohe	Auckland	Winter spinach N a	Measured		3
28	Vegetables	15.5			0			177.966	-38.639	Poverty Bay	Gisborne	Summer broccoli, w	Modelled	Spasmo	
29	Vegetables	25.5						173.132	-41.334	Waimea Plai	Tasman	Cabbage and lettu	Modelled	Spasmo	40
30	Vegetables	20						173.132	-41.334	Waimea Plai	Tasman	Pumpkin and lettu	Modelled	Spasmo	40
31	Vegetables	33.7						173.132	-41.334	Waimea Plai	Tasman	Lettuce, cabbage,	Modelled	Spasmo	40
32	Vegetables		2.2		5.15			177.966	-38.639	Poverty Bay	Gisborne	Summer broccoli, w	Modelled	Spasmo	
33	Horticulture	2	3.5					173.904	-41.454	Rapaura	Marlborough	Viticulture	Measured	Spasmo	2
34	Horticulture		5.2		0			177.966	-38.639	Poverty Bay	Gisborne	Viticulture	Modelled	Spasmo	
35	Horticulture		18.3		0			177.966	-38.639	Poverty Bay	Gisborne	Citrus orange	Modelled	Spasmo	
36	Horticulture		9.9		0			177.966	-38.639	Poverty Bay	Gisborne	Kiwifruit, irrigated	Modelled	Spasmo	
37	Horticulture		10.8					173.132	-41.334	Waimea Plai	Tasman	Apples	Modelled	Spasmo	40
38	Horticulture		11.5					173.132	-41.334	Waimea Plai	Tasman	Viticulture	Modelled	Spasmo	40
39	Horticulture		11.3					173.132	-41.334	Waimea Plai	Tasman	Viticulture	Modelled	Spasmo	40
40	Horticulture		10.7					173.132	-41.334	Waimea Plai	Tasman	Apples (+berries, h	Modelled	Spasmo	40
41	Horticulture		8.7		0.26			178.019	-38.662		Gisborne	Viticulture	Modelled	Spasmo	44
42	Horticulture		6.2		0.51			176.847	-39.638		Hawke's Bay	Viticulture	Modelled	Spasmo	44
43	Horticulture		14		0.09			175.459	-41.218		Martinboroug	Viticulture	Modelled	Spasmo	44
44	Horticulture		9.2		0.25			173.868	-41.517		Marlborough	Viticulture	Modelled	Spasmo	44
45	Horticulture		9		0.03			173.282	-41.273		Nelson	Viticulture	Modelled	Spasmo	44
46	Horticulture		11		0.09			168.83	-44.942	Central Otag	Otago	Viticulture	Modelled	Spasmo	44
47	Horticulture		13.5		0.14			173.868	-41.517		Marlborough	Viticulture	Modelled	Spasmo	44
48	Horticulture		0.1		0.4			177.966	-38.639	Poverty Bay	Gisborne	Viticulture	Modelled	Spasmo	
49	Horticulture		0.1		0.3			177.966	-38.639	Poverty Bay	Gisborne	Citrus orange	Modelled	Spasmo	
50	Horticulture				0.2			177.966	-38.639	Poverty Bay	Gisborne	Kiwifruit, irrigated	Modelled	Spasmo	
51	Gorse		47					176.347	-38.06	Rotorua	Bay of Plenty	Gorse	Measured		1.7
52	Gorse		51					176.347	-38.06	Rotorua	Bay of Plenty	Gorse	Measured		1.7
53	Gorse	28.9						175.253	-41.296	Ruamahang	Wellington	Gorse	Modelled		

	P	Q	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM
1	Process	Slope	Annual fertiliser rates (kg N/ha/yr)	Annual fertiliser rates (kg P/ha/yr)	Mean annual soil N surplus (kg/ha/yr)	Conventional, minimum, or no tillage	Fallow period (yes/no)	Use of grazed winter forage	Mean soil Olsen P concn (mg/L)	Mean soil Olsen P concn (mg/kg)	Lysimeter, plot, catchment	Lysimeter type	Plot or catchment size (m2)	Stream order (1 or 2)	Reference
2	Runoff	Steep						no			Catchment		2960000		Quinn and Stroud (2002)
3	Runoff	Steep		20.7				no			Catchment		2960000		Quinn and Stroud (2002)
4	Leaching	Steep						no			Lysimeter	Barrel	0.0005		Parfitt et al. (2002)
5	Leaching	Steep						no			Lysimeter	Barrel	0.0005		Parfitt et al. (2002)
6	Leaching	Steep						no			Lysimeter	Barrel	0.0005		Parfitt et al. (2002)
7	Leaching							no			Lysimeter	Barrel	0.0005		Magesan and Wang (2008)
8	Leaching							no			Lysimeter	Barrel	0.0005		Magesan and Wang (2008)
9	Leaching	Flat						no			Lysimeter	Barrel	0.0005		Parfitt and Ross (2011)
10	Leaching	Flat						no			Lysimeter	Barrel	0.0005		Parfitt and Ross (2011) but refers to c
11	Runoff	Steep						no			Catchment		2960000	2	Quinn and Stroud (2002)
12	Runoff	Flat						no			Catchment		1800000		Davies-Colley and Nagels (2002)
13	Runoff	Flat						no			Catchment		2100000		Davies-Colley and Nagels (2002)
14	Leaching	Flat					yes	no			Plot		2100000		Francis (1995)
15	Leaching	Flat						no			Plot		0.0005		Francis (1995)
16	Leaching	Flat						no			Plot		0.0005		Francis (1995)
17	Leaching	Flat						no			Plot		0.0005		Francis (1995)
18	Leaching	Flat						no			Plot		0.0005		Francis (1995)
19	Leaching	Flat						no			Plot		0.0005		Francis (1995)
20	Leaching						yes	no			Lysimeter	Porous	0.031		Norris et al. (2017)
21	Leaching	Flat	162	25				no							Gentile et al. (2014)
22	Leaching	Flat	129	20				no							Gentile et al. (2014)
23	Runoff	Flat						no			-				Gentile et al. (2014)
24	Runoff	Flat						no			-				Gentile et al. (2014)
25	Leaching	Flat	0	0				no			Lysimeter	Porous	0.0005		Williams et al. (2003)
26	Leaching	Flat	200	270				no		72	Lysimeter	Porous	0.0005		Williams et al. (2003)
27	Leaching	Flat	400	270				no			Lysimeter	Porous	0.0005		Williams et al. (2003)
28	Leaching	Flat	100	25				no							Gentile et al. (2014)
29	Leaching	Flat	200	77				no							Fenemor et al. (2015)
30	Leaching	Flat	141	51				no							Fenemor et al. (2015)
31	Leaching	Flat	152	63				yes							Fenemor et al. (2016)
32	Runoff	Flat	100	25				no			-				Gentile et al. (2014)
33	Leaching	Flat						no			Lysimeter	Drainag	0.031		Green et al. (2014)
34	Leaching	Flat	0	0				no							Gentile et al. (2014)
35	Leaching	Flat	70	15				no							Gentile et al. (2014)
36	Leaching	Flat	92					no							Gentile et al. (2014)
37	Leaching	Flat	40					no							Fenemor et al. (2015)
38	Leaching	Flat	5					no							Fenemor et al. (2015)
39	Leaching	Flat	5					no							Fenemor et al. (2016)
40	Leaching	Flat	40					no							Fenemor et al. (2016)
41	Leaching	Flat	5	5.5				no							Clothier and Green (2017)
42	Leaching	Flat	5	5.5				no							Clothier and Green (2017)
43	Leaching	Flat	5	5.5				no							Clothier and Green (2017)
44	Leaching	Flat	5	5.5				no							Clothier and Green (2017)
45	Leaching	Flat	5	5.5				no							Clothier and Green (2017)
46	Leaching	Flat	5	5.5				no							Clothier and Green (2017)
47	Leaching	Flat	5	5.5				no							Clothier and Green (2017)
48	Runoff	Flat	0	0				no			-				Gentile et al. (2014)
49	Runoff	Flat	70	15				no			-				Gentile et al. (2014)
50	Runoff	Flat	92					no			-				Gentile et al. (2014)
51	Leaching							no		31	Lysimeter	Porous	0.0005		Magesan and Wang (2008)
52	Leaching							no		60	Lysimeter	Porous	0.0005		Magesan and Wang (2008)
53	Leaching							no							Mason et al. (2016)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM
1	Dominant land use	TN loss (kg N/ha/yr)	Nitrate-N loss (kg N/ha/yr)	DRP loss (kg P/ha/yr)	TP loss (kg P/ha/yr)	Sediment loss (kg/ha/yr)	E. coli loss (MPN ha^{-1})	Longitude approx	Latitude approx	Location	Region	Management	Measured, modelled	Model type	Duration (yr)	Process	Slope	Annual fertiliser rates (kg N/ha/yr)	Annual fertiliser rates (kg P/ha/yr)	Mean annual soil N surplus (kg/ha/yr)	Conventional, minimum, or no tillage	Fallow period (yes/no)	Use of grazed winter forage	Mean soil Olsen P concn (mg/L)	Mean soil Olsen P concn (mg/kg)	Lysimeter, plot, catchment	Lysimeter type	Plot or catchment size (m2)	Stream order (1 or 2)	Reference
52	Gorse		51					176.347	-38.06	Rotorua	Bay of Plenty	Gorse	Measured		1.7	Leaching							no		60	Lysimeter	Porous	0.0005		Magesan and Wang (2008)
53	Gorse	28.9						175.253	-41.296	Ruamahang	Wellington	Gorse	Modelled			Leaching							no							Mason et al. (2016)
54	Cropping		72.6					175.246	-37.952	Te Awamutu	Waikato	Maize, fallow	Measured		2	Leaching						yes	no			Lysimeter	Barrel	0.785		Tsimba et al. (2021)
55	Cropping		4.6					175.246	-37.952	Te Awamutu	Waikato	Maize, ryegrass	Measured		3	Leaching						yes	no			Lysimeter	Barrel	0.785		Tsimba et al. (2021)
56	Vegetables					7000		175.311	-39.414	Ohakune	Manawatu-W	Carrots	Measured		4	Runoff*	Easy				Conventional		no			Plot		0.015		Basher et al. (2004)
57	Vegetables					16000		174.881	-37.129	Pukekohe	Auckland	Onions, potatoes,	Measured		1	Runoff*	Easy				Convention	yes	no			Plot		0.015		Basher & Ross (2002)
58	Horticulture		3.4					176.846	-39.638	Hastings	Hawke's Bay	Apples	Measured		38	Leaching	Flat						no			Lysimeter	Drainag	0.031		Green et al. (2012)
59	Horticulture		7					176.846	-39.638	Hastings	Hawke's Bay	Apples	Measured	Spasmo	38	Leaching	Flat						no			Lysimeter	Drainag	0.031		Green et al. (2012)
60	Horticulture		12.5								Various	Kiwifruit	Modelled	Spasmo	37	Leaching	Flat						no							Deurer et al. (2011)
61	Cropping		20					171.815	-43.973	Wakanui	Canterbury	Barley	Modelled	APSIM	4	Leaching	Flat	199					no							Khaembah & Horrocks (2018)
62	Cropping		12.2					171.414	-43.821	Mayfield	Canterbury	Barley	Modelled	APSIM	4	Leaching	Flat	144					no							Khaembah & Horrocks (2018)
63	Cropping		32					146.43	-41.2		Tasmania	Potatoes	Modelled	APSIM	25	Leaching		397		89	Minimum		no							Lisson & Cotching (2011)
64	Cropping		<10					146.43	-41.2		Tasmania	Bean, pea, maize,	Modelled	APSIM	25	Leaching		207			Minimum		no							Lisson & Cotching (2011)
65	Cropping		36		0.1			171.646	-43.635	Methven	Canterbury	Ryegrass seed, pa	Measured		6	Leaching	Flat	414	24.4		Convention	no	no	12		Lysimeter	Drainag	0.031		Trolove et al. (2021)
66	Cropping		13		0.05			171.646	-43.635	Methven	Canterbury	Cereal, carrots, rye	Measured		6	Leaching	Flat	165	22.2		Convention	no	no	16		Lysimeter	Drainag	0.031		Trolove et al. (2021)
67	Cropping		40		0.1			172.25	-43.808	Southbridge	Canterbury	Mixed cropping	Measured		6	Leaching	Flat	85.78	17.8		Minimum	no	no	26		Lysimeter	Drainag	0.031		Trolove et al. (2021)
68	Cropping		145		0.33			175.382	-40.179	Bulls	Manawatu	Mixed cropping	Measured		6	Leaching	Flat	276	47.8		Convention	yes	no	28		Lysimeter	Drainag	0.031		Trolove et al. (2021)
69	Cropping		24		0.07			175.4	-39.417	Ohakune	Manawatu	Mixed cropping	Measured		6	Leaching	Flat	150.8	48.8		Minimum	yes	no	26		Lysimeter	Drainag	0.031		Trolove et al. (2021)
70	Cropping		12		0.16			176.628	-39.896	Otane	Hawke's Bay	Mixed cropping	Measured		6	Leaching	Flat	158.2	36.4		Convention	yes	no	25.5		Lysimeter	Drainag	0.031		Trolove et al. (2021)
71	Cropping		7					176.847	-39.639	Hastings	Hawke's Bay	Mixed cropping	Measured		1.5	Leaching	Flat	106.8	6		Convention	no	no	51		Lysimeter	Drainag	0.031		Trolove et al. (2021)
72	Cropping		28		0.31			176.349	-40.027	Takapau	Hawke's Bay	Mixed cropping	Measured		6	Leaching	Flat	141.8	33.8		Convention	yes	no	22		Lysimeter	Drainag	0.031		Trolove et al. (2021)
73	Cropping		51		0.06			175.773	-37.811	Matamata	Waikato	Mixed cropping	Measured		6	Leaching	Flat	174.5	114		Convention	yes	no	43		Lysimeter	Drainag	0.031		Trolove et al. (2021)
74	Vegetables		193		0.1			174.904	-37.203	Pukekohe	Auckland	Lettuce, potato	Measured		6	Leaching	Flat	255	62		Convention	yes	no	188		Lysimeter	Drainag	0.031		Trolove et al. (2021)
75	Cropping		102		0.16			174.942	-37.267	Tuakau	Waikato	Mixed cropping	Measured		6	Leaching	Flat	107	30.333		Convention	yes	no	114		Lysimeter	Drainag	0.031		Trolove et al. (2021)
76	Dairy		27					172.33	-43.65	Lincoln	Canterbury	Dairy	Measured		2	Leaching	Flat	148.5		237			no			Catchment		225000		Al-Marashdeh et al. (2021)
77	Dairy		17.5					172.33	-43.65	Lincoln	Canterbury	Dairy	Measured		2	Leaching	Flat	150.5		236			no			Catchment		225000		Al-Marashdeh et al. (2021)
78	Dairy		39.5					172.33	-43.65	Lincoln	Canterbury	Dairy	Measured		2	Leaching	Flat	279		372			no			Catchment		225000		Al-Marashdeh et al. (2021)
79	Dairy		29					172.43	-43.63		Canterbury	Dairy	Measured		4	Leaching	Flat	156		176			no			Lysimeter	Barrel			Chapman et al. (2020)
80	Dairy		43					172.43	-43.63		Canterbury	Dairy	Measured		4	Leaching	Flat	304		323			no			Lysimeter	Barrel			Chapman et al. (2020)
81	Dairy		48					172.34	-43.64	Lincoln	Canterbury	Dairy	Modelled		1	Leaching	Flat	0					no			Lysimeter	Barrel			Beukes et al. (2020)
82	Dairy		62					172.34	-43.64	Lincoln	Canterbury	Dairy	Modelled		1	Leaching	Flat	150					no			Lysimeter	Barrel			Beukes et al. (2020)
83	Dairy		71					172.34	-43.64	Lincoln	Canterbury	Dairy	Modelled		1	Leaching	Flat	300					no			Lysimeter	Barrel			Beukes et al. (2020)
84	Dairy		62					175.36	-37.76	Hamilton	Waikato	Dairy	Measured		4	Leaching		137					no			Lysimeter	Porous ceramic cups			Beukes et al. (2019)
85	Dairy		29					169.71	-46.28		Otago	Dairy	Measured		4	Leaching		109					no			Lysimeter	Porous ceramic cups			Beukes et al. (2019)
86	Dairy		56.7					175.37	-37.77	Hamilton	Waikato	Dairy	Modelled	APSIM	3	Leaching	Flat	180					no							Beukes et al. (2017)
87	Dairy		32					175.37	-37.77	Hamilton	Waikato	Dairy	Modelled	APSIM	3	Leaching	Flat	50					no							Beukes et al. (2017)
88	Dairy	19	14					175.6	-40.4	Palmerston N	Manawatu-W	Dairy	Measured		3	Runoff	Flat	72					no	33.17						Christensen et al. (2019)
89	Dairy	11.02	6.72					175.6	-40.4	Palmerston N	Manawatu-W	Dairy	Measured		3	Runoff	Flat	72					no	33.6						Christensen et al. (2019)
90	Dairy	10.5	5.76		0.42			175.2	-37.5	Ranfurly	Otago	Dairy	Measured		1	Runoff	Flat	150	35				no			Catchment		1430000		McDowell (2017)
91	Dairy	0.81	0.22		0.03			175.2	-37.5	Ranfurly	Otago	Dairy	Measured		1	Runoff	Flat	140	26				no			Catchment		1430000		McDowell (2017)
92	Dairy			0.071	0.25			172.33	-43.65	Lincoln	Canterbury	Dairy	Measured		14	Leaching	Flat	170	40				no	17		Lysimeter	Barrel	0.78525		McDowell et al. (2019)
93	Dairy			0.319	1.46			172.33	-43.65	Lincoln	Canterbury	Dairy	Measured		14	Leaching	Flat	170	40				no	26		Lysimeter	Barrel	0.78525		McDowell et al. (2019)
94	Dairy			0.048	0.117			172.33	-43.65	Lincoln	Canterbury	Dairy	Measured		14	Leaching	Flat	170	40				no	19		Lysimeter	Barrel	0.78525		McDowell et al. (2019)
95	Dairy			0.05	0.122			172.33	-43.65	Lincoln	Canterbury	Dairy	Measured		14	Leaching	Flat	170	40				no	20		Lysimeter	Barrel	0.78525		McDowell et al. (2019)
96	Dairy		50					175.31	-37.78	Hamilton	Waikato	Dairy	Measured		2	Leaching		200					no			Lysimeter	Porous	81000		Roche et al. (2016)
97	Dairy		35					175.31	-37.78	Hamilton	Waikato	Dairy	Measured		2	Leaching		200					no			Lysimeter	Porous	73000		Roche et al. (2016)
98	Dairy		48					175.31	-37.78	Hamilton	Waikato	Dairy	Measured		2	Leaching		200					no			Lysimeter	Porous	61000		Roche et al. (2016)
99	Dairy		33					175.31	-37.78	Hamilton	Waikato	Dairy	Measured		2	Leaching		200					no			Lysimeter	Porous	49000		Roche et al. (2016)
100	Dairy		23					175.31	-37.78	Hamilton	Waikato	Dairy	Measured		2	Leaching		200					no			Lysimeter	Porous	45000		Roche et al. (2016)
101	Dairy		63					175.31	-37.78		Waikato	Dairy	Measured		3	Leaching		180					no					250000		Vogeler et al. (2013)
102	Dairy	100						168.32	-46.31	Wallacetown	Southland	Fodder beet graze	Measured		2	Leaching	Flat						yes			Lysimeter	Porous ceramic cups			Smith and Monaghan (2020)
103	Dairy	67.5						168.32	-46.31	Wallacetown	Southland	Fodder beet graze	Measured		2	Leaching	Flat						yes			Lysimeter	Porous ceramic cups			Smith and Monaghan (2020)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM
1	Dominant land use	TN loss (kg N/ha/yr)	Nitrate-N loss (kg N/ha/yr)	DRP loss (kg P/ha/yr)	TP loss (kg P/ha/yr)	Sediment loss (kg/ha/yr)	E. coli loss (MPN ha^{-1})	Longitude approx	Latitude approx	Location	Region	Management	Measured, modelled	Model type	Duration (yr)	Process	Slope	Annual fertiliser rates (kg N/ha/yr)	Annual fertiliser rates (kg P/ha/yr)	Mean annual soil N surplus (kg/ha/yr)	Conventional, minimum, or no tillage	Fallow period (yes/no)	Use of grazed winter forage	Mean soil Olsen P concn (mg/L)	Mean soil Olsen P concn (mg/kg)	Lysimeter, plot, catchment	Lysimeter type	Plot or catchment size (m2)	Stream order (1 or 2)	Reference
102	Dairy	100						168.32	-46.31	Wallacetowr	Southland	Fodder beet graze	Measured		2	Leaching	Flat						yes			Lysimeter	Porous ceramic cups			Smith and Monaghan (2020)
103	Dairy	67.5						168.32	-46.31	Wallacetowr	Southland	Fodder beet graze	Measured		2	Leaching	Flat						yes			Lysimeter	Porous ceramic cups			Smith and Monaghan (2020)
104	Dairy	120						168.32	-46.31	Wallacetowr	Southland	Kale grazed in Win	Measured		2	Leaching	Flat						yes			Lysimeter	Porous ceramic cups			Smith and Monaghan (2020)
105	Dairy	20						168.32	-46.31	Wallacetowr	Southland	Dairy	Measured		2	Leaching	Flat						no			Lysimeter	Porous ceramic cups			Smith and Monaghan (2020)
106	Dairy	40			1.3			175.6	-37.7	Toenepi	Waikato	Dairy	Measured		15	Runoff	Flat	120	37				no			Catchment		1.6E+07		Wilcock et al. (2013)
107	Dairy	40			1.3			174.1	-39.5	Waiokura	Taranaki	Dairy	Measured		15	Runoff	Flat	151	41				no			Catchment		2.1E+07		Wilcock et al. (2013)
108	Dairy	30			1.2			171	-44.9	Waikakahi	Canterburry	Dairy	Measured		15	Runoff	Flat	177	34				no			Catchment		2.3E+07		Wilcock et al. (2013)
109	Dairy	70			1.7			171.4	-42.7	Inchbonnie	West Coast	Dairy	Measured		15	Runoff	Flat	97	30				no			Catchment		5980000		Wilcock et al. (2013)
110	Dairy	20			2			168.2	-46	Bog Burn	Southland	Dairy	Measured		15	Runoff	Flat	103	37				no			Catchment		2.5E+07		Wilcock et al. (2013)
111	Dairy				1			168.4	-46.2	Invercargill	Southland	Dairy	Measured		3	Runoff*	Flat		30				no			Plot		216		McDowell and Smith (2012)
112	Dairy				0.82	356	2.09E+10	169.69	-46.09	Balclutha	Otago	Winter cropping	Measured		2	Runoff	Steep	78	58				yes			Catchment		43000	1	McDowell (2006)
113	Dairy	27.1			3.83	3740	4.3E+10	169.74	-46.289	Telford	Otago	Winter cropping	Measured		2	Runoff	Steep						yes			Catchment		20000		Monaghan et al. (2017)
114	Dairy	4.5			0.62	219	1.5E+10	169.74	-46.289	Telford	Otago	Winter cropping	Measured		2	Runoff	Steep						yes			Catchment		20000		Monaghan et al. (2017)
115	Sheep			0.27	0.95	400		170.735	-44.952	Rolling Dowr	Otago	Sheep	Measured		1.5	Runoff*	Rolling	75	22				yes		27	Plot		1000		McDowell and Houlbrooke (2009)
116	Sheep			0.33	0.65			170.735	-44.952	Rolling Dowr	Otago	Sheep	Measured		1.5	Runoff*	Rolling	75	22				no		28.7	Plot		1000		McDowell and Houlbrooke (2009)
117	Beef			0.45	1.42	467		170.735	-44.952	Rolling Dowr	Otago	Cattle grazed	Measured		1.5	Runoff*	Rolling	75	22				yes		26	Plot		1000		McDowell and Houlbrooke (2009)
118	Sheep				0.1			168.79	-44.98		Otago	Sheep	Measured			Runoff	Easy						no			Catchment		4.8E+07	2	Caruso (2000)
119	Sheep	7	0.6			128		-176.29	-43.76	Rotorua	Bay of Plenty		Measured			Runoff	Rolling									Catchment		7.3E+07		Williamson et al. (1996)
120	Deer	6	2.47	0.028	0.864	4480	3.41E+09	170.406	-45.866	Mosgiel	Otago	Deer	Measured		1	Runoff	Easy		25				no		22.5	Catchment		41000	2	McDowell (2007)
121	Deer	3	0.6	0.036	2.994	3950	9.79E+08	169.709	-46.302	Balclutha	Otago	Deer	Measured		1	Runoff	Rolling		35				yes		25	Catchment		61000	2	McDowell (2007)
122	Deer	0.415	0.12	0.008	0.22	1560	3.13E+08	169.709	-46.302	Balclutha	Otago	Deer	Measured		1	Runoff	Rolling		35				no		25	Catchment		361000	2	McDowell (2007)
123	Deer	5	2.33	0.2	1.5	3356	3.13E+11	170.406	-45.866	Mosgiel	Otago	Deer	Measured		2	Runoff	Easy		25						22.5	Catchment		41000	2	McDowell (2008)
124	Deer	10	7	0.32	1.4	88	2.90E+10	169.66	-46.38	Glenomaru	Otago	Deer	Measured		2	Runoff	Rolling	46	23						38	Catchment		190	2	McDowell et al (2009)
125	Deer	5	3	0.28	0.22	11	2.10E+11	168.47	-45.71	Waimea	Southland	Deer	Measured		3	Runoff	Rolling	65	20						33	Catchment		290	2	McDowell et al (2009)
126	Dairy		30	0.32				168.79	-46.32	Edendale	Southland	Dairy	Measured		4	Runoff	Flat	0					no			Plot		900		Monaghan et al (2000)
127	Dairy		34	0.1				168.79	-46.32	Edendale	Southland	Dairy	Measured		4	Runoff	Flat	100					no			Plot		900		Monaghan et al (2000)
128	Dairy		46	0.17				168.79	-46.32	Edendale	Southland	Dairy	Measured		4	Runoff	Flat	200					no			Plot		900		Monaghan et al (2000)
129	Dairy		56	0.32				168.79	-46.32	Edendale	Southland	Dairy	Measured		4	Runoff	Flat	400					no			Plot		900		Monaghan et al (2000)
130	Sheep	10	5	0.15	0.94	1695		175.84	-40.31	Ballantrae	Manawatu	Sheep	Measured		4	Runoff	Easy		65				no			Catchment		8500	1	Lambert et al (1985)
131	Sheep	10	3	0.1	0.95	1754		175.84	-40.31	Ballantrae	Manawatu	Sheep	Measured		4	Runoff	Easy		11				no			Catchment		7000	1	Lambert et al (1985)
132	Beef	85	58	0.53	1.3			168.92	-44.6	Matukituki	Otago	Winter cropping	Measured		1	Leaching	Flat	130	71				yes	24		Plot		1000		Smith et al (2016)
133	Beef	53	12	0.5	1			168.92	-44.6	Matukituki	Otago	Cattle	Measured		1	Leaching	Flat	9	12				no	14		Plot		1000		Smith et al (2016)
134	Cropping		60					142.09	-34.24	Mildura	Victoria	Wheat	Modelled	APSIM	27	Leaching	Flat	100					no			Plot		1		Nazrollahi et al. (2021)
135	Cropping		5					142.92	-35.99	Birchip	Victoria	Wheat	Modelled	APSIM	27	Leaching	Flat	100					no			Plot		1		Nazrollahi et al. (2021)
136	Cropping		5					142.3	-36.67	Longerenno	Victoria	Wheat	Modelled	APSIM	27	Leaching	Flat	100					no			Plot		1		Nazrollahi et al. (2021)
137	Cropping		5					142.84	-37.71	Lake Bolac	Victoria	Wheat	Modelled	APSIM	27	Leaching	Flat	100					no			Plot		1		Nazrollahi et al. (2021)
138	Pasture	7	6		0.56	58	1.07E+10	169.668	-46.385	Glenomaru	Otago	Deer	Measured		1	Runoff	Flat		31							Catchment		320000	2	McDowell et al. (2006)
139	Deer	2			3	1012	5.1E+10	168.492	-45.762	Lumsden	Southland	Deer	Measured		<1	Runoff*	Rolling	45	60				yes	32		Plot		16		McDowell and Stevens (2008)
140	Pasture	28	24.5		0.34			175.617	-40.395	Palmerston N	Manawatu	Dairy	Measured		<1	Leaching	Flat	72	49				no			Plot		1600		Houlbrooke et al. (2003)
141	Pasture	31	26.4		0.86			175.617	-40.395	Palmerston N	Manawatu	Dairy	Measured		<1	Leaching	Flat	72	38				no			Plot		1600		Houlbrooke et al. (2003)
142	Pasture	31.4	24.5	0.14	0.65			175.617	-40.395	Palmerston N	Manawatu	Dairy	Measured		<1	Leaching	Flat	135	38				no			Plot		1600		Houlbrooke et al. (2008)
143	Pasture	36.7	22.1	1.08	2.55			175.617	-40.395	Palmerston N	Manawatu	Dairy	Measured		<1	Leaching	Flat	154	31				no			Plot		1600		Houlbrooke et al. (2008)
144	Native forest		0.01		0.2			174.96	-41.183	Taita	Wellington	Native forest	Measured		2	Runoff	Easy									Catchment		109300		McColl et al. (1977)
145	Native forest		1.4		0.1			174.96	-41.183	Taita	Wellington	Native forest	Measured		2	Runoff	Easy									Catchment		46900		McColl et al. (1977)
146	Exotic forestry		0.04		0.07			174.96	-41.183	Taita	Wellington	Exotic forest	Measured		2	Runoff	Easy									Catchment		36000		McColl et al. (1977)
147	Native forest	2	0.55	0.27	0.6	320		175.25	-37.78	Whatawhata	Waikato	Native forest	Measured		1	Runoff	Steep									Catchment		3000000		Quinn and Stroud (2002)
148	Native forest	2.84	3.67	0.017	0.12			176.23	-38.45	Purukohukol	Bay of Plenty	Native forest	Measured		3	Runoff	Easy									Catchment		280000		Cooper and Thomsen (1988)
149	Exotic forestry	0.55	1.31	0.036	0.095			176.23	-38.45	Purukohukol	Bay of Plenty	Exotic forest	Measured		3	Runoff	Easy									Catchment		340000		Cooper and Thomsen (1988)
150	Exotic forestry		0.04					176.258	-38.871	Kaingaroa	Bay of Plenty	Exotic forest	Measured		2	Leaching	Flat									Lysimeter	Porous ceramic cups			Dyck et al. (1983)
151	Exotic forestry		1.2					176.258	-38.871	Kaingaroa	Bay of Plenty	Exotic forest	Measured		2	Leaching	Flat									Lysimeter	Porous ceramic cups			Dyck et al. (1983)
152	Exotic forestry		0.36					176.52	-38.036	Rotoehu	Bay of Plenty	Exotic forest	Measured		1	Leaching	Easy									Lysimeter	Porous ceramic cups			Dyck et al. (1983)
153	Gorse		22.9					176.742	-38.411	Rangitaiki	Bay of Plenty	Gorse	Measured		1	Leaching	Flat									Lysimeter	Porous ceramic cups			Dyck et al. (1983)
154	Sheep		5.4					176.1	-40.55	Wairarapa	Wellington	Sheep	Measured		3	Leaching	Easy	0					no			Lysimeter	Barrel	0.03		Crofoot et al. (2010)
155	Sheep		7.7					176.1	-40.55	Wairarapa	Wellington	Sheep	Measured		3	Leaching	Easy	60					no			Lysimeter	Barrel	0.03		Crofoot et al. (2010)
156	Sheep		15					176.1	-40.55	Wairarapa	Wellington	Sheep	Measured		3	Leaching	Easy	120					no			Lysimeter	Barrel	0.03		Crofoot et al. (2010)

Appendix 2 – study references used in database

Al-Marashdeh O, Cameron K, Hodge S, Gregorini P, Edwards G 2021. Integrating plantain (Plantago lanceolata L.) and italian ryegrass (Lolium multiflorum Lam.) into New Zealand grazing dairy system: the effect on farm productivity, profitability, and nitrogen losses. Animals 11(2): 376.

Basher LR, Ross CW 2002. Soil erosion rates under intensive vegetable production on clay loam, strongly structured soils at Pukekohe, New Zealand. Australian Journal of Soil Research 40(6): 947–961.

Basher LR, Ross CW, Dando J 2004. Effects of carrot growing on volcanic ash soils in the Ohakune area, New Zealand. Soil Research 42(3): 259–272.

Beukes P, Romera A, Hutchinson K, van der Weerden T, de Klein C, Dalley D, Chapman D, Glassey C, Dynes R 2019. Benefits and trade-offs of dairy system changes aimed at reducing nitrate leaching. Animals 9(12): 1158.

Beukes PC, Gregorini P, Cameron K, Attwood GT 2020. Farm-scale carbon and nitrogen fluxes in pastoral dairy production systems using different nitrogen fertilizer regimes. Nutrient Cycling in Agroecosystems 117(1): 1–12.

Beukes PC, Romera AJ, Gregorini P, Macdonald KA, Glassey CB, Shepherd MA 2017. The performance of an efficient dairy system using a combination of nitrogen leaching mitigation strategies in a variable climate. Science of The Total Environment 599-600: 1791–1801.

Caruso BS 2000. Spatial and temporal variability of stream phosphorus in a New Zealand high-country agricultural catchment. New Zealand Journal of Agricultural Research 43(2): 235–249.

Chapman DF, Dalley DE, Edwards GR, Cameron KC, Malcolm BJ, Clement A, Romera AJ, Pinxterhuis IB, Beukes PC, Di HJ and others 2021. Production, profit and nitrogen flows in irrigated dairy systems representing different industry development pathways: the Pastoral 21 experience in Canterbury. New Zealand Journal of Agricultural Research 64(1): 3–35.

Christensen, C.L., Hedley, M.J., Hanly, J.A., Horne, D.J., 2019. Duration-controlled grazing of dairy cows. 2: nitrogen losses in sub-surface drainage water and surface runoff. New Zealand Journal of Agricultural Research 62(1): 48–68.

Clothier BE, Green SR 2017. The leaching and runoff of nutrients from vineyards. In: Currie LD, Hedley MJ ed. Science and policy: nutrient management challenges for the next generation. Palmerston North, Fertilizer and Lime Research Centre, Massey University. Pp. 7.

Cooper AB, Thomsen CE 1988. Nitrogen and phosphorus in streamwaters from adjacent pasture, pine, and native forest catchments. New Zealand Journal of Marine and Freshwater Research 22(2): 279–291.

Crofoot AN, Crofoot EW, Hoogendoorn CJ, Litherland AJ, Garland CB 2010. N-leaching in hill country; farmer led research. Proceedings of the New Zealand Grassland Association (72): 55–60.

Davies-Colley RJ, Nagels JW 2002. Effects of dairying on water quality of lowland streams in Westland and Waikato. Proceedings of the New Zealand Grassland Association 64: 107–114.

Deurer M, Green SR, Clothier BE, Mowat A 2011. Can product water footprints indicate the hydrological impact of primary production? – A case study of New Zealand kiwifruit. Journal of Hydrology 408(3): 246–256.

Dyck W, Gosz J, Hodgkiss P 1983. Nitrate losses from disturbed ecosystems in New Zealand. A comparative analysis. New Zealand Journal of Forestry Science 13(1): 14–24.

Fenemor A, Green S, Dryden G, Samarasinghe O, Newsome P, Price R, Betts H, Lilburne L 2015. Crop production, profit and nutrient losses in relation to irrigation water allocation and reliability – Waimea Plains, Tasman District. Prepared for Ministry for Primary Industries MPI Technical Paper No: 2015/36.

Fenemor A, Price R, Green S 2016. Modelling the source and fate of nitrate-nitrogen losses from Waimea Plains land uses. Envirolink Advice Grant: 1592-TSDC116. Prepared for Tasman District Council. 22 p.

Francis G 1995. Management practices for minimising nitrate leaching after ploughing temporary leguminous pastures in Canterbury, New Zealand. Journal of Contaminant Hydrology 20(3-4): 313–327.

Gentile R, Green S, Mason K, van den Dijssel C, Johnstone P, Clothier B 2014. Land management practices and nutrient losses from farms on the Poverty Bay Flats. A Plant & Food Research report prepared for Gisborne District Council Contract No. 30930. PFR SPTS No. 10506.

Green S, Agnew R, Greven M 2014. Monitoring nitrate loss under vineyard soils on the Wairau Plains, Marlborough. Report for Marlborough District Council PFR SPTS No. 10328. 34 p.

Green S, Deurer M, Clothier B 2012. Consumptive water footprint analysis for apple fruit production. Acta Horticulturae 951: 197–204.

Houlbrooke DJ, Horne DJ, Hedley CB, Hanly JA, Snow V 2003. The impact of intensive dairy farming on the leaching losses of nitrogen and phosphorus from a mole and pipe drained soil. Proceedings of the New Zealand Grassland Association 65: 179–184.

Houlbrooke DJ, Horne DJ, Hedley MJ, Snow VO, Hanly JA 2008. Land application of farm dairy effluent to a mole and pipe drained soil: implications for nutrient enrichment of winter-spring drainage. Soil Research 46(1): 45–52.

Khaembah EN, Horrocks A 2018. A modelling approach to assessment and improvement of nitrogen management on New Zealand arable farms: a case study. Agronomy New Zealand 48: 1–11.

Lambert MG, Devantier BP, Nes P, Penny PE 1985. Losses of nitrogen, phosphorus, and sediment in runoff from hill country under different fertiliser and grazing management regimes. New Zealand Journal of Agricultural Research 28(3): 371–379.

Lisson SN, Cotching WE 2011. Modelling the fate of water and nitrogen in the mixed vegetable farming systems of northern Tasmania, Australia. Agricultural Systems 104(8): 600–608.

Magesan GN, Wang H 2008. Nitrogen leaching from gorse. Final Report. Contract report for Environment Bay of Plenty. Scion. 16 p.

Mason N, Mudge P, Palmer D, McLeod M, Ausseil A, Dymond J 2016. Catchment-scale contribution of invasive nitrogen fixing shrubs to nitrate leaching: a scoping study. Journal of the Royal Society of New Zealand 46(2): 85–102.

McColl RHS, White E, Gibson AR 1977. Phosphorus and nitrate run-off in hill pasture and forest catchments, Taita, New Zealand. New Zealand Journal of Marine and Freshwater Research 11(4): 729–744.

McDowell RW 2006. Phosphorus and Sediment Loss in a Catchment with Winter Forage Grazing of Cropland by Dairy Cattle. Journal of Environmental Quality 35(2): 575–583.

McDowell RW 2007. Water Quality in Headwater Catchments with Deer Wallows. Journal of Environmental Quality 36(5): 1377–1382.

McDowell RW 2008. Water quality of a stream recently fenced-off from deer. New Zealand Journal of Agricultural Research 51(3): 291–298.

McDowell RW 2009. Maintaining good water and soil quality in catchments containing deer farms. International Journal of River Basin Management 7(3): 187–195.

McDowell RW 2017. Does variable rate irrigation decrease nutrient leaching losses from grazed dairy farming? Soil Use and Management 33(4): 530–537.

McDowell RW, Gray CW, Cameron KC, Di HJ, Pellow R 2019. The efficacy of good practice to prevent long-term leaching losses of phosphorus from an irrigated dairy farm. Agriculture, Ecosystems & Environment 273: 86–94.

McDowell RW, Houlbrooke DJ 2009. Management options to decrease phosphorus and sediment losses from irrigated cropland grazed by cattle and sheep. Soil Use and Management 25(3): 224–233.

McDowell RW, McGrouther N, Morgan G, Srinivasan MS, Stevens DR, Johnson M, Copland R 2006. Monitoring the impact of farm practices on water quality in the Otago and Southland deer focus farms. Proceedings of the New Zealand Grassland Association 68: 183–188.

McDowell RW, Smith LC 2012. Potential water quality impact and agronomic effectiveness of different phosphorus fertilisers under grazed dairying in Southland. Proceedings of the New Zealand Grassland Association 74: 225–230.

McDowell RW, Stevens DR 2008. Potential waterway contamination associated with wintering deer on pastures and forage crops. New Zealand Journal of Agricultural Research 51(3): 287–290.

Monaghan RM, Laurenson S, Dalley DE, Orchiston TS 2017. Grazing strategies for reducing contaminant losses to water from forage crop fields grazed by cattle during winter. New Zealand Journal of Agricultural Research 60(3): 333–348.

Monaghan RM, Paton RJ, Smith LC, Binet C 2000. Nutrient losses in drainage and surface runoff from a cattle-grazed pasture in Southland. Proceedings of the New Zealand Grassland Association 62: 99–104.

Nasrollahi N, Hunt J, Tang C, Cann D 2021. Modelled quantification of different sources of nitrogen inefficiency in semi-arid cropping systems. Agronomy 11(6): 1222.

Norris M, Johnstone P, Green S, Van der Klei G, van den Dijssel C, Wright P, Clark G, Thomas S, Williams R, Mathers D, and Halliday A 2017. Rootzone Reality – A network of fluxmeters measuring nutrient losses under cropping rotations. Summary of Year 1 and Year 2 results. In: Currie LD, Hedley MJ eds Science and policy: nutrient management challenges for the next generation. Palmerston North, NZ: Fertilizer and Lime Research Centre, Massey University. 10 p.

Parfitt R, Ross D 2011. Long-term effects of afforestation with Pinus radiata on soil carbon, nitrogen, and pH: a case study. Soil Research 49(6): 494–503.

Parfitt R, Salt G, Hill L 2002. Clear-cutting reduces nitrate leaching in a pine plantation of high natural N status. Forest Ecology and Management 170(1-3): 43–53.

Quinn JM, Stroud MJ 2002. Water quality and sediment and nutrient export from New Zealand hill-land catchments of contrasting land use. New Zealand Journal of Marine and Freshwater Research 36(2): 409–429.

Roche JR, Ledgard SF, Sprosen MS, Lindsey SB, Penno JW, Horan B, Macdonald KA 2016. Increased stocking rate and associated strategic dry-off decision rules reduced the amount of nitrate-N leached under grazing. Journal of Dairy Science 99(7): 5916–5925.

Smith C, Monaghan RM, McDowell R, Carrick S 2016. Nitrogen and phosphorus leaching losses from pasture, winter forage crop and native bush sites in the West Matukituki Valley. AgResearch E500/2016/013. 26 p.

Smith LC, Monaghan RM 2020. Nitrogen leaching losses from fodder beet and kale crops grazed by dairy cows in southern Southland. Journal of New Zealand Grasslands 82: 61–71.

Trolove S, Wallace D, Johnstone P, Sorensen I, Arnold N, van der Weyden J, van den Dijssel C, Dellow S, Wright P, Clark G and others 2021. Protecting our groundwater: Fluxmeter network summary report. A Plant and Food Research report prepared for: Foundation for Arable Research. Milestone No. 78521. Contract No. 35902. Job code: P/411139/01. PFR SPTS No. 20648. 62 p.

Tsimba R, Gunn TD, Densley RJ, Millar JA, Williams ID, Edmeades GO 2021. Quantification and mitigation of nitrogen leaching in a maize silage cropping system. Journal of New Zealand Grasslands 83: 163–170.

Vogeler I, Beukes P, Burggraaf V 2013. Evaluation of mitigation strategies for nitrate leaching on pasture-based dairy systems. Agricultural Systems 115: 21–28.

Wilcock RJ, Monaghan RM, Quinn JM, Srinivasan MS, Houlbrooke DJ, Duncan MJ, Wright-Stow AE, Scarsbrook MR 2013. Trends in water quality of five dairy farming streams in response to adoption of best practice and benefits of long-term monitoring at the catchment scale. Marine and Freshwater Research 64(5): 401–412.

Williams P, Tregurtha R, Francis G 2003. Fate of urea applied to winter spinach in New Zealand. Nutrient Cycling in Agroecosystems 67(3): 245–254.

Williamson RB, Smith CM, Cooper AB 1996. Watershed riparian management and its benefits to a eutrophic lake. Journal of Water Resources Planning and Management 122(1): 24–32.

Appendix 3 – statistical analysis code for 'R' software

Details of the statistical 'R' software code are copied below.

```
library(ggplot2)
library(here)
library(readxl)

theme_set(theme_bw())

filename <- "N_loss_data_recent_3_Jun.xlsx"
coltypes <-
c("numeric","numeric","text","numeric","numeric",rep("text",4),"numeric",rep("text",3),rep("nu
meric",9),
        "text","text","text","numeric","text","numeric","text","text")

Nloss <- read_excel(here(filename),skip=1,col_types=coltypes)

Nloss$N_N[63] <- 6

Nloss$landuseD <- relevel(as.factor(Nloss$landuse),"Dairy")

# Replace NA in rain and irrigation variables with 0
Nloss$irrigationNA0 <- Nloss$irrigation
Nloss$irrigationNA0[is.na(Nloss$irrigationNA0)] <- 0
Nloss$rainNA0 <- Nloss$rain
Nloss$rainNA0[is.na(Nloss$rainNA0)] <- 0

# Then sum rain and irrigation values to give water added.

Nloss$rain_plus_irrig <- Nloss$rainNA0 + Nloss$irrigationNA0
Nloss$lnTN <- log(Nloss$TN)
Nloss$lnN_N <- log(Nloss$N_N)
NlossA <- Nloss

# NOTES

### Remove gorse and keep an eye on Horticulture
### Make new variable, Total N where it is available N_N otherwise
### Combine Easy, Rolling and Steep slope classes
### Combine Lysimeter and Plot
### Soil orders Allophanic, Pummice and Other

NlossA <- NlossA[which(Nloss$landuse!="Gorse"),]

NlossA$flat_hill <- NlossA$slope_class
NlossA$flat_hill[which(NlossA$slope_class=="Easy")] <- "Hill"
NlossA$flat_hill[which(NlossA$slope_class=="Rolling")] <- "Hill"
NlossA$flat_hill[which(NlossA$slope_class=="Steep")] <- "Hill"

NlossA$Nitrogen <- NlossA$TN
```

```
NlossA$Nitrogen[is.na(NlossA$TN)] <- NlossA$N_N[is.na(NlossA$TN)]
NlossA <- as.data.frame(NlossA)
NlossA$LnNitrogen <- log(NlossA$Nitrogen)

NlossA$scale[which(NlossA$scale=="Lysimeter")] <- "Lysimeter/Plot"
NlossA$scale[which(NlossA$scale=="Plot")] <- "Lysimeter/Plot"

NlossA$soilordernew <- NlossA$soilorder
NlossA$soilordernew[which(NlossA$soilorder=="Brown")] <- "Other"
NlossA$soilordernew[which(NlossA$soilorder=="Gley")] <- "Other"
NlossA$soilordernew[which(NlossA$soilorder=="Pallic")] <- "Other"
NlossA$soilordernew[which(NlossA$soilorder=="Recent")] <- "Other"
NlossA$soilordernew[which(NlossA$soilorder=="Semi-arid")] <- "Other"
NlossA$soilordernew[which(NlossA$soilorder=="Various")] <- "Other"
NlossA$soilordernew[which(NlossA$soilorder=="Granular")] <- "Other"

NlossAll <- NlossA

FinalData <- NULL
FinalData$lat <- NlossAll$lat
FinalData$long <- NlossAll$long
FinalData$region <- NlossAll$region
FinalData$LnNitrogen <- NlossAll$LnNitrogen
FinalData$landuseD <- NlossAll$landuseD
FinalData$process <- NlossAll$process
FinalData$rain_plus_irrig <- NlossAll$rain_plus_irrig

FinalData <- as.data.frame(FinalData)
FD <- FinalData
FinalData$Comp_cases <- complete.cases(FinalData)
FinalData$fertiliser_N <- NlossAll$fertiliser_N
FinalData$TN <- NlossAll$TN
FinalData$isTN[!is.na(FinalData$TN)] <- "TN"
table(FinalData$isTN)
FinalData$N_N <- NlossAll$N_N
FinalData$isN_N[!is.na(FinalData$N_N)] <- "N_N"
table(FinalData$isN_N)
table(FinalData$isTN,FinalData$isN_N, useNA = "ifany")

FinalData <- FinalData[FinalData$Comp_cases,]
table(FinalData$isTN,FinalData$isN_N, useNA = "ifany")
FinalData$ID <- seq(1:length(FinalData$long))

Final <- lm(LnNitrogen ~ landuseD+rain_plus_irrig+process+fertiliser_N+
       landuseD:process+landuseD:rain_plus_irrig+process:rain_plus_irrig,
      data=FinalData )
anova(Final)
sFinal <- summary(Final)
sFinal
# str(sFinal)
round(sFinal$coefficients,3)
```

```
pN1 <- ggplot(NlossAll, aes(x=landuse, y=Nitrogen)) + geom_jitter(height=0,width=0.25)
pN1 <- pN1 + geom_boxplot(alpha=0.2,outlier.shape = NA)
pN1 <- pN1 + facet_wrap(~process)
pN1 <- pN1 + theme(axis.text.x=element_text(angle=90, vjust=0.5))
pN1 <- pN1 + ylab("Nitrogen (kg N/ha/yr)") + xlab("Land use")
pN1

# ggsave(here("Report_Boxplot.tiff"), pN1, width=8, height =5)


dev.off()

pN2a <- ggplot(NlossAll, aes(x=rain_plus_irrig, y=Nitrogen)) + geom_jitter(height=0,width=0.25)
pN2a <- pN2a + facet_grid(process~landuseD)
pN2a <- pN2a + geom_smooth(method="lm", se=FALSE)
pN2a <- pN2a + ylab("Nitrogen (kg N/ha/yr)") + xlab("Water Added (rain and irrigation) (mm/yr)")
pN2a <- pN2a + scale_x_continuous(breaks=seq(0, 8000, 4000),limits=c(0, 8000))
pN2a <- pN2a + theme(axis.text.x  = element_text(angle=90, vjust=0.5))
pN2a

# ggsave(here("PN2a.tiff"),pN2a,  height=4, width=9, dpi=600)

pN2b <- ggplot(NlossAll, aes(x=rain_plus_irrig, y=Nitrogen)) + geom_jitter(height=0,width=0.25)
pN2b <- pN2b + facet_grid(landuseD~process)
pN2b <- pN2b + geom_smooth(method="lm", se=FALSE)
pN2b <- pN2b + theme(strip.text.y = element_text(angle=0))
pN2b <- pN2b + ylab("Nitrogen (kg N/ha/yr)") + xlab("Water Added (rain and irrigation) (mm/yr)")
pN2b <- pN2b + scale_y_continuous(breaks=seq(0, 200, 100))
pN2b

# ggsave(here("PN2b.tiff"),pN2b,  height=4, width=6, dpi=600)

FinalData2 <- FinalData[!is.na(FinalData$fertiliser_N),]
FinalData2$Predicted <- exp(predict(Final))

pN5 <- ggplot(FinalData2, aes(x=exp(LnNitrogen), y=Predicted)) + geom_point()
pN5 <- pN5 + geom_abline()
  # geom_smooth(method="lm", se=FALSE)
pN5 <- pN5 + ylab("Predicted Nitrogen (kg N/ha/yr)") + xlab("Measured Nitrogen (kg N/ha/yr)")
pN5
# ggsave(here("PN5.tiff"),pN5,  height=4, width=6, dpi=600)

table(FinalData2$isTN,FinalData2$isN_N, useNA = "ifany")
```

Appendix 4 – statistical output

Details of the statistical output while combining explanatory variables, prior to the final model, are copied below.

```
Allm1 <- lm(LnNitrogen ~
landuse*process*flat_hill*soilordernew*rain_plus_irrig*fertiliser_N*fertiliser_P,
data=NlossAll )
anova(Allm1)
Analysis of Variance Table

Response: LnNitrogen
                                                 Df Sum Sq Mean Sq F value    Pr(>F)
landuse                                           5 35.056  7.0112 21.8343 0.0003878 ***
process                                           1 27.731 27.7311 86.3600 3.462e-05 ***
soilordernew                                      1  4.004  4.0045 12.4708 0.0095798 **
rain_plus_irrig                                   1  7.631  7.6311 23.7648 0.0018041 **
fertiliser_N                                      1  3.989  3.9887 12.4216 0.0096699 **
fertiliser_P                                      1  0.589  0.5891  1.8346 0.2176771
landuse:process                                   2 13.458  6.7288 20.9547 0.0011091 **
landuse:soilordernew                              1  0.977  0.9771  3.0429 0.1246059
landuse:rain_plus_irrig                           4  5.075  1.2688  3.9514 0.0548850 .
soilordernew:rain_plus_irrig                      1  0.439  0.4388  1.3664 0.2806918
landuse:fertiliser_N                              4  0.368  0.0921  0.2867 0.8777259
process:fertiliser_N                              1  0.001  0.0008  0.0024 0.9626108
soilordernew:fertiliser_N                         1  0.068  0.0676  0.2105 0.6602899
rain_plus_irrig:fertiliser_N                      1  0.373  0.3732  1.1622 0.3167469
landuse:fertiliser_P                              3  3.973  1.3244  4.1244 0.0559238 .
process:fertiliser_P                              1  0.638  0.6378  1.9862 0.2015854
rain_plus_irrig:fertiliser_P                      1  0.436  0.4363  1.3588 0.2819278
fertiliser_N:fertiliser_P                         1  0.017  0.0169  0.0525 0.8252483
landuse:process:fertiliser_N                      1  2.857  2.8574  8.8986 0.0204261 *
landuse:rain_plus_irrig:fertiliser_N              2  0.300  0.1498  0.4666 0.6453284
landuse:rain_plus_irrig:fertiliser_P              1  1.217  1.2167  3.7892 0.0926337 .
landuse:fertiliser_N:fertiliser_P                 2  0.530  0.2651  0.8257 0.4764743
rain_plus_irrig:fertiliser_N:fertiliser_P         1  0.559  0.5590  1.7409 0.2285371
landuse:rain_plus_irrig:fertiliser_N:fertiliser_P 1  0.832  0.8321  2.5914 0.1514805
Residuals                                         7  2.248  0.3211
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
>  Allm1a <- lm(LnNitrogen ~
landuse*process*soilordernew*rain_plus_irrig*fertiliser_N*fertiliser_P, data=NlossAll )
>  anova(Allm1a)
Analysis of Variance Table
```

```
Response: LnNitrogen
                                                Df Sum Sq Mean Sq F value    Pr(>F)
landuse                                          5 35.056  7.0112 21.8343 0.0003878 ***
process                                          1 27.731 27.7311 86.3600 3.462e-05 ***
soilordernew                                     1  4.004  4.0045 12.4708 0.0095798 **
rain_plus_irrig                                  1  7.631  7.6311 23.7648 0.0018041 **
fertiliser_N                                     1  3.989  3.9887 12.4216 0.0096699 **
fertiliser_P                                     1  0.589  0.5891  1.8346 0.2176771
landuse:process                                  2 13.458  6.7288 20.9547 0.0011091 **
landuse:soilordernew                             1  0.977  0.9771  3.0429 0.1246059
landuse:rain_plus_irrig                          4  5.075  1.2688  3.9514 0.0548850 .
soilordernew:rain_plus_irrig                     1  0.439  0.4388  1.3664 0.2806918
landuse:fertiliser_N                             4  0.368  0.0921  0.2867 0.8777259
process:fertiliser_N                             1  0.001  0.0008  0.0024 0.9626108
soilordernew:fertiliser_N                        1  0.068  0.0676  0.2105 0.6602899
rain_plus_irrig:fertiliser_N                     1  0.373  0.3732  1.1622 0.3167469
landuse:fertiliser_P                             3  3.973  1.3244  4.1244 0.0559238 .
process:fertiliser_P                             1  0.638  0.6378  1.9862 0.2015854
rain_plus_irrig:fertiliser_P                     1  0.436  0.4363  1.3588 0.2819278
fertiliser_N:fertiliser_P                        1  0.017  0.0169  0.0525 0.8252483
landuse:process:fertiliser_N                     1  2.857  2.8574  8.8986 0.0204261 *
landuse:rain_plus_irrig:fertiliser_N             2  0.300  0.1498  0.4666 0.6453284
landuse:rain_plus_irrig:fertiliser_P             1  1.217  1.2167  3.7892 0.0926337 .
landuse:fertiliser_N:fertiliser_P                2  0.530  0.2651  0.8257 0.4764743
rain_plus_irrig:fertiliser_N:fertiliser_P        1  0.559  0.5590  1.7409 0.2285371
landuse:rain_plus_irrig:fertiliser_N:fertiliser_P 1  0.832  0.8321  2.5914 0.1514805
Residuals                                        7  2.248  0.3211
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>


>  Allm2 <- lm(LnNitrogen ~ landuse*process*soilordernew*rain_plus_irrig*fertiliser_N,
data=NlossAll )
>  anova(Allm2)
Analysis of Variance Table

Response: LnNitrogen
                                 Df Sum Sq Mean Sq F value    Pr(>F)
landuse                           6 43.714  7.2856 17.6294 2.823e-10 ***
process                           1 22.007 22.0072 53.2520 4.200e-09 ***
soilordernew                      1  0.105  0.1050  0.2540 0.6167532
rain_plus_irrig                   1  6.305  6.3050 15.2565 0.0003193 ***
fertiliser_N                      1  3.020  3.0197  7.3070 0.0097248 **
landuse:process                   2 27.028 13.5141 32.7009 1.983e-09 ***
landuse:soilordernew              1  0.052  0.0519  0.1255 0.7248376
process:soilordernew              1  0.547  0.5466  1.3227 0.2563273
landuse:rain_plus_irrig           4  6.477  1.6193  3.9184 0.0083155 **
process:rain_plus_irrig           1  1.809  1.8094  4.3783 0.0422011 *
soilordernew:rain_plus_irrig      1  0.002  0.0016  0.0039 0.9502161
```

```
landuse:fertiliser_N                         5  0.305  0.0609  0.1474 0.9797811
process:fertiliser_N                         1  0.047  0.0473  0.1145 0.7367158
soilordernew:fertiliser_N                    1  0.101  0.1014  0.2454 0.6228309
rain_plus_irrig:fertiliser_N                 1  0.746  0.7455  1.8040 0.1861165
landuse:soilordernew:rain_plus_irrig         1  0.383  0.3830  0.9267 0.3409892
process:soilordernew:rain_plus_irrig         1  0.041  0.0411  0.0994 0.7539872
landuse:process:fertiliser_N                 1  0.047  0.0474  0.1147 0.7364748
landuse:rain_plus_irrig:fertiliser_N         3  0.287  0.0958  0.2318 0.8737227
process:rain_plus_irrig:fertiliser_N         1  1.342  1.3417  3.2466 0.0784273 .
soilordernew:rain_plus_irrig:fertiliser_N  1  0.369  0.3694  0.8938 0.3496199
Residuals                                   44 18.184  0.4133
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
>  Allm3 <- lm(LnNitrogen ~ landuse*process*rain_plus_irrig+soilordernew+fertiliser_N,
data=NlossAll )
>  anova(Allm3)
Analysis of Variance Table

Response: LnNitrogen
                        Df Sum Sq Mean Sq F value    Pr(>F)
landuse                  6 43.714  7.2856 20.4507 3.991e-13 ***
process                  1 22.007 22.0072 61.7741 6.280e-11 ***
rain_plus_irrig          1  6.342  6.3420 17.8021 8.004e-05 ***
soilordernew             1  0.068  0.0679  0.1907   0.66387
fertiliser_N             1  3.020  3.0197  8.4763   0.00497 **
landuse:process          2 27.028 13.5141 37.9341 1.539e-11 ***
landuse:rain_plus_irrig  4  6.499  1.6247  4.5606   0.00268 **
process:rain_plus_irrig  1  1.796  1.7964  5.0426   0.02825 *
Residuals               63 22.444  0.3563
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
>   Allm4 <- lm(LnNitrogen ~
landuse+process+rain_plus_irrig+landuse:process+landuse:rain_plus_irrig+process:rain_plus_i
rrig+fertiliser_N,
+                data=NlossAll )
>   anova(Allm4)
Analysis of Variance Table

Response: LnNitrogen
                        Df Sum Sq Mean Sq F value    Pr(>F)
landuse                  6 47.562  7.9270 19.3256 3.222e-13 ***
process                  1 22.007 22.0072 53.6526 2.969e-10 ***
rain_plus_irrig          1  8.323  8.3230 20.2911 2.553e-05 ***
fertiliser_N             1  4.075  4.0754  9.9358  0.002375 **
landuse:process          2 26.505 13.2525 32.3090 1.053e-10 ***
```

```
landuse:rain_plus_irrig  4  7.182  1.7955  4.3774  0.003205 **
process:rain_plus_irrig  1  1.887  1.8875  4.6016  0.035364 *
Residuals               71 29.123  0.4102
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
```

Details of the final statistical output are copied below.

N_loss_data June 21st

Alasdair Noble

2022-06-21

Version for report

```
 TN
46


N_N
121


       N_N <NA>
  TN    29   17
  <NA>  92   13


       N_N <NA>
  TN    29   17
  <NA>  91    0

Analysis of Variance Table

Response: LnNitrogen
                         Df Sum Sq Mean Sq F value    Pr(>F)
landuseD                  6 47.562  7.9270 19.3256 3.222e-13 ***
rain_plus_irrig           1  4.746  4.7465 11.5717  0.001103 **
process                   1 25.584 25.5837 62.3720 2.589e-11 ***
fertiliser_N              1  4.075  4.0754  9.9358  0.002375 **
landuseD:process          2 26.505 13.2525 32.3090 1.053e-10 ***
landuseD:rain_plus_irrig  4  7.182  1.7955  4.3774  0.003205 **
rain_plus_irrig:process   1  1.887  1.8875  4.6016  0.035364 *
Residuals                71 29.123  0.4102
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Call:
lm(formula = LnNitrogen ~ landuseD + rain_plus_irrig + process +
    fertiliser_N + landuseD:process + landuseD:rain_plus_irrig +
    process:rain_plus_irrig, data = FinalData)
```

```
Residuals:
     Min       1Q   Median       3Q      Max 
-2.91298 -0.28259 -0.00299  0.23720  2.10758 

Coefficients: (6 not defined because of singularities)
                                       Estimate Std. Error t value Pr(>|t|)    
(Intercept)                           4.1309754  0.6631237   6.230 2.93e-08 ***
landuseDBeef                          2.1771166  1.0294944   2.115 0.037963 *  
landuseDCropping                     -2.8356831  0.7521016  -3.770 0.000334 ***
landuseDDeer                         -4.9341390  2.6233999  -1.881 0.064098 .  
landuseDHorticulture                 -1.5682734  1.4912306  -1.052 0.296520    
landuseDSheep                        -1.1433546  0.4076033  -2.805 0.006485 ** 
landuseDVegetables                   -6.5553757  1.8196307  -3.603 0.000581 ***
rain_plus_irrig                      -0.0010298  0.0006921  -1.488 0.141211    
processRunoff                        -2.0267257  0.7137892  -2.839 0.005891 ** 
fertiliser_N                          0.0029520  0.0009118   3.238 0.001833 ** 
landuseDBeef:processRunoff                   NA         NA      NA       NA    
landuseDCropping:processRunoff               NA         NA      NA       NA    
landuseDDeer:processRunoff                   NA         NA      NA       NA    
landuseDHorticulture:processRunoff   -4.1318142  0.5619182  -7.353 2.63e-10 ***
landuseDSheep:processRunoff                  NA         NA      NA       NA    
landuseDVegetables:processRunoff     -1.9597288  0.7343243  -2.669 0.009428 ** 
landuseDBeef:rain_plus_irrig                 NA         NA      NA       NA    
landuseDCropping:rain_plus_irrig      0.0024301  0.0008010   3.034 0.003372 ** 
landuseDDeer:rain_plus_irrig          0.0040950  0.0028224   1.451 0.151217    
landuseDHorticulture:rain_plus_irrig  0.0006907  0.0016370   0.422 0.674363    
landuseDSheep:rain_plus_irrig                NA         NA      NA       NA    
landuseDVegetables:rain_plus_irrig    0.0064048  0.0016725   3.830 0.000274 ***
rain_plus_irrig:processRunoff         0.0015291  0.0007128   2.145 0.035364 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6405 on 71 degrees of freedom
  (49 observations deleted due to missingness)
Multiple R-squared:  0.8014,	Adjusted R-squared:  0.7567 
F-statistic: 17.91 on 16 and 71 DF,  p-value: < 2.2e-16

                                     Estimate Std. Error t value Pr(>|t|)
(Intercept)                             4.131      0.663   6.230    0.000
landuseDBeef                            2.177      1.029   2.115    0.038
landuseDCropping                       -2.836      0.752  -3.770    0.000
landuseDDeer                           -4.934      2.623  -1.881    0.064
landuseDHorticulture                   -1.568      1.491  -1.052    0.297
landuseDSheep                          -1.143      0.408  -2.805    0.006
landuseDVegetables                     -6.555      1.820  -3.603    0.001
rain_plus_irrig                        -0.001      0.001  -1.488    0.141
processRunoff                          -2.027      0.714  -2.839    0.006
fertiliser_N                            0.003      0.001   3.238    0.002
landuseDHorticulture:processRunoff     -4.132      0.562  -7.353    0.000
landuseDVegetables:processRunoff       -1.960      0.734  -2.669    0.009
landuseDCropping:rain_plus_irrig        0.002      0.001   3.034    0.003
landuseDDeer:rain_plus_irrig            0.004      0.003   1.451    0.151
landuseDHorticulture:rain_plus_irrig    0.001      0.002   0.422    0.674
landuseDVegetables:rain_plus_irrig      0.006      0.002   3.830    0.000
rain_plus_irrig:processRunoff           0.002      0.001   2.145    0.035
```

```
       N_N <NA>
  TN    12   10
  <NA>  66    0
```

Appendix 5 – additional graphs

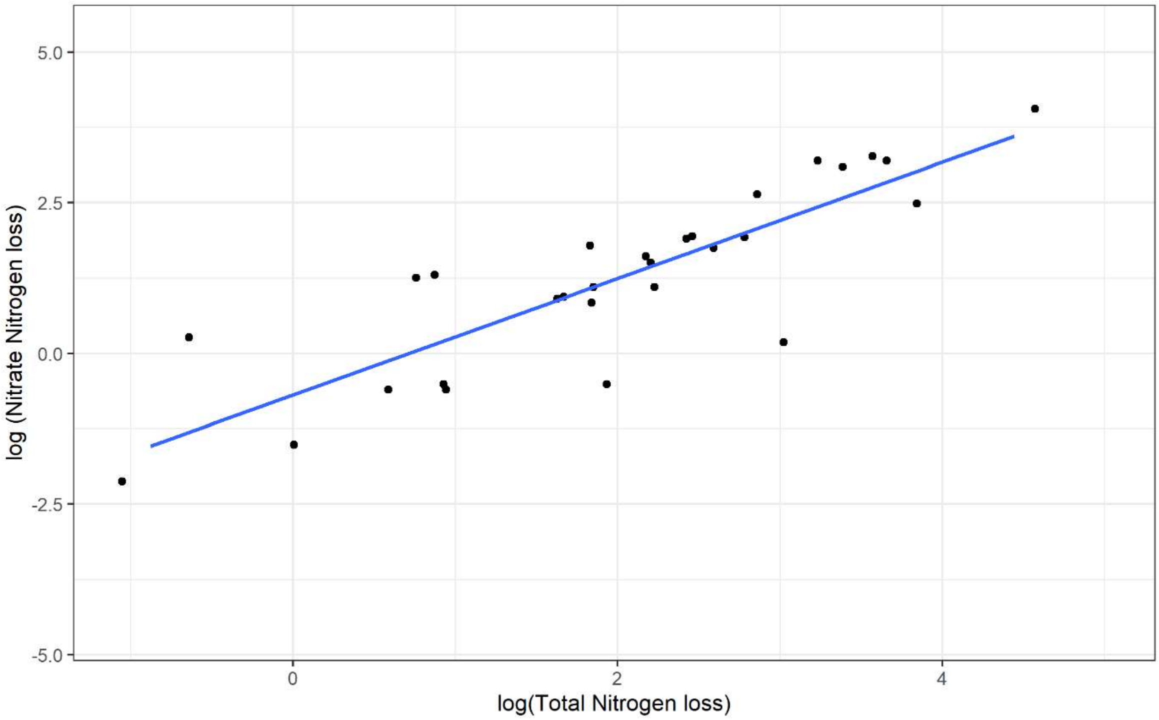

Figure 4. Log (Total Nitrogen loss (TN) (kg N/ha/yr)) versus log (Nitrate-Nitrogen loss (kg N/ha/yr)).

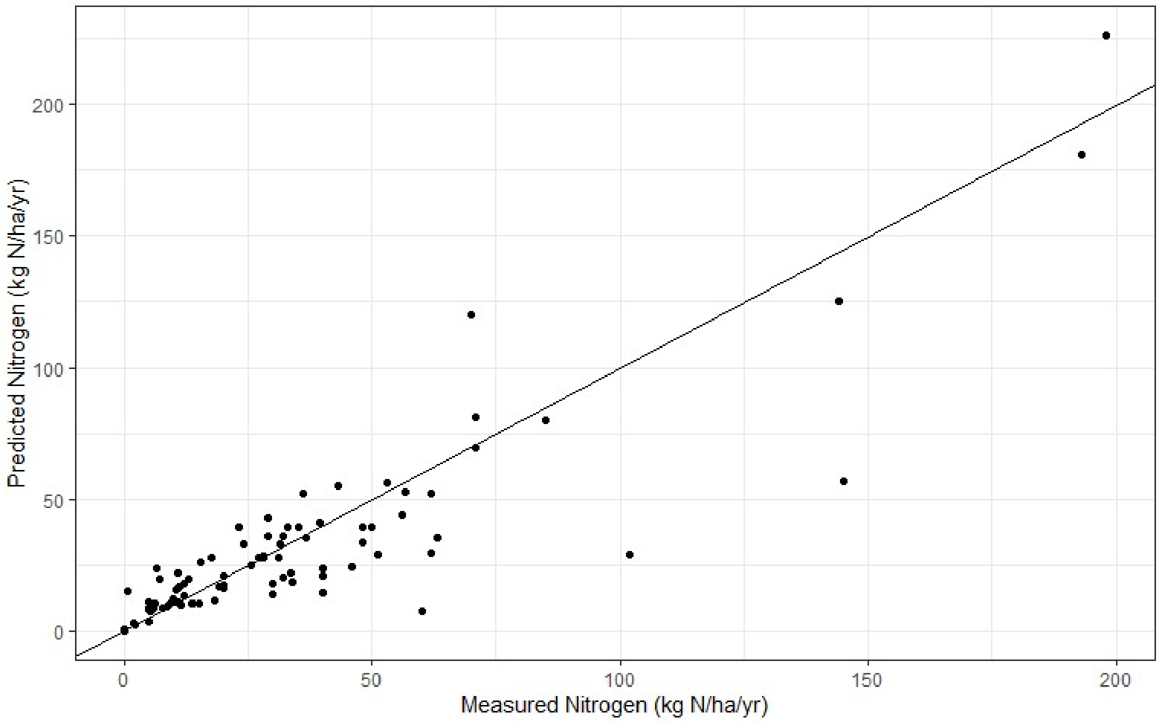

Figure 5. Actual vs predicted values from the final model. Line represents the 1:1 line.

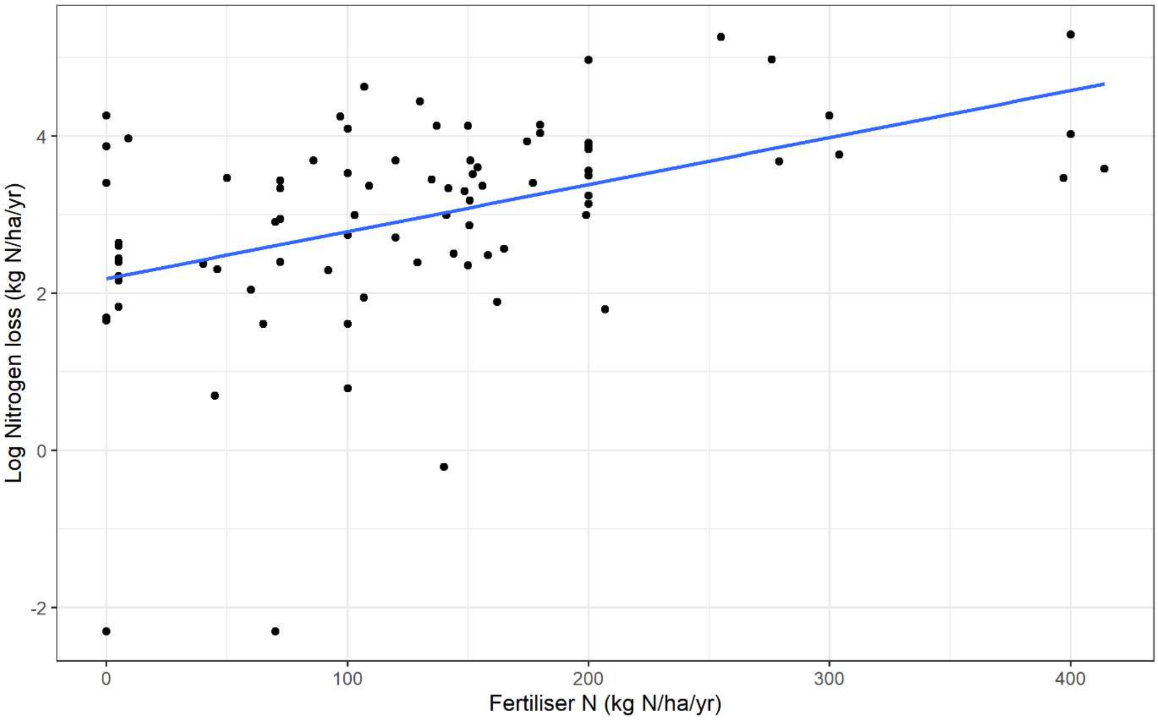

Figure 6. Log ('Nitrogen' loss (kg N/ha/yr))(response variable) predicted by the model for the explanatory variable 'Fertiliser N' (kg N/ha/yr).

Appendix 6 – possible additional use of data for Overseer^{FM}

The data can also be used to assess the accuracy and precision of outputs from other tools, such as Overseer^{FM}. If such an assessment was undertaken for Overseer^{FM}, we recommend the following process be used:

1 From the N loss database, classes would be selected to represent a land use and region.

2 From our assessment we also recommend that sufficient data are available to assess the interaction of land use by fertiliser rate.

3 Estimates of error can be produced for each observation where data were available for multiple years.

4 Overseer files should be produced to match each observation. Note that we recommend that files are created to span a sensible range for significant variables. For example, if a dairy farm did not have stocking rate, files would be created to span the 25th and 75th percentile of rates within the region by land use. This would also include interactions such as stocking rate by observed fertiliser rates (currently 100–400 kg N/ha/yr).

5 To test performance, observations would be plotted against the estimates for each class recording their accuracy and precision, as per Correndo et al. (2021).