Science of the Total Environment 927 (2024) 171759

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Determining the likelihood and cost of detecting reductions of nitrate-nitrogen concentrations in groundwater across New Zealand

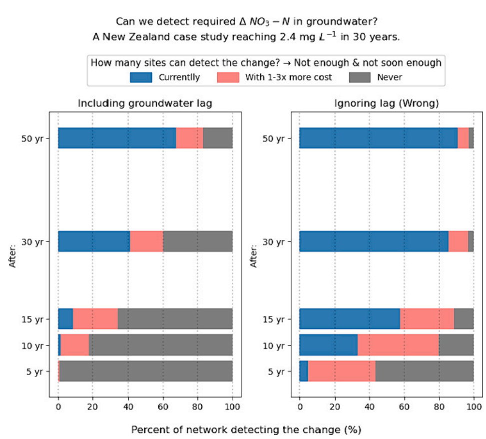
M. Dumont [a], Z. Etheridge [a,b], R.W. McDowell [c,d,*]

[a] Komanawa Solutions Ltd., Christchurch, New Zealand
[b] School of Earth and Environment & Waterways Centre for Freshwater Management, University of Canterbury, New Zealand
[c] Faculty of Agriculture and Life Sciences, Lincoln University, Lincoln, New Zealand
[d] AgResearch, Lincoln Science Centre, Lincoln, New Zealand

HIGHLIGHTS

- Detecting NO3-N reductions likely requires bespoke groundwater monitoring networks.
- Excluding groundwater lag over-estimates detection power.
- 40 % of New Zealand's network can detect policy-relevant changes after 30 years.
- A 1–3-fold increase in funding allows detection in a further 20 % of NZ's network.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Jurgen Mahlknecht

Keywords:
Lag time
Leaching
Land management
Mitigation
Power

ABSTRACT

Nitrate-nitrogen (NO_3-N) is a contaminant of concern in groundwater worldwide. Stakeholders need information on the ability to detect changes in NO_3-N concentrations to prove that land management practices are meeting water quality aims. We created a database of quarterly to monthly NO_3-N measurements in 948 sites across New Zealand; 186 of those sites had mean residence time (MRT) data. New Zealand has set a target of sufficient land use mitigations in the next 30 years to ensure steady state surface water concentrations do not exceed 2.4 mg L^{-1}. Here we assess whether the current monitoring network could identify the impacts of these mitigations, assuming that the mitigations are successfully implemented at the source. Only 41 % of the network could detect statistically significant reductions with the current standard quarterly sampling after 30 years of monitoring. The percentage of sites increased to 60 % with increased monitoring frequency (often weekly) but this required a 100–300 % increase in monitoring costs. However, policy makers and stakeholders typically require information on policy and mitigation effectiveness within 5–10 years. Detection within 5–10 years was very unlikely (0–20 % of sites) regardless of the sampling frequency. Importantly, these analyses include the impacts of groundwater lag

* Corresponding author at: Faculty of Agriculture and Life Sciences, Lincoln University, Lincoln, New Zealand.
E-mail address: richard.mcdowell@lincoln.ac.nz (R.W. McDowell).

https://doi.org/10.1016/j.scitotenv.2024.171759
Received 23 November 2023; Received in revised form 19 February 2024; Accepted 14 March 2024
Available online 22 March 2024

and temporal dispersion on the likelihood of detecting change, ignoring these impacts, incorrectly, yields a much higher likelihood of detecting reductions. We conclude that the current monitoring network is unlikely to be fit for the purpose of detecting NO_3-N reductions within practical timeframes or budgets. Furthermore, we conclude that lag and temporal dispersion effects must be included in detection power calculations; we therefore recommend that MRT data is regularly collected. We also provide a python package to enable easy detection power calculations with lag and temporal dispersion impacts, thereby supporting the development of robust change-detection monitoring networks.

1. Introduction

Groundwater is an important source of potable water for human consumption, being the primary source for around 3 billion people globally. Groundwater also provides base flow for surface streams and rivers and is an important source of inflowing water to lakes and reservoirs. The enrichment of groundwater with nitrate-nitrogen (NO_3-N) has several negative impacts for water quality. At concentrations >0.8 mg L^{-1}, NO_3-N can stimulate the growth of periphyton and phytoplankton (McDowell et al., 2020), lowering the amenity and aesthetic value of streams and lakes. Concentrations >2.4 mg L^{-1} can be toxic to some macroinvertebrates and fish (Camargo et al., 2005; Horak et al., 2019; Wagenhoff et al., 2017), while even greater concentrations (>11.3 mg L^{-1}) can be harmful if consumed by humans (Rahman et al., 2021).

Globally, our policy response has been focused on identifying and remediating the cause of NO_3-N enrichment of groundwater (Carvalho et al., 2019). This has resulted in well-established links between intensive agriculture and NO_3-N enrichment (Larned et al., 2018), but also of time lags between changes in agricultural intensity and changes in NO_3-N concentrations (Howden et al., 2010; Wang et al., 2012). Groundwater lags in both the vadose and saturated zone can provide additional information about the historical patterns in NO_3-N loading and expected future concentrations (MacDonald et al., 2003). Storage of NO_3-N in the vadose zone and, where not monitored, the saturated zone significantly affects NO_3-N concentrations and model predictions at local to global scales (Ascott et al., 2017). Ascott et al. (2017, 2021) make a compelling argument that the effects of NO_3-N storage via lag must be included in NO_3-N analysis, modelling, and management policy for those policies to be effective.

Groundwater lags vary significantly: commonly increasing with increasing depth to groundwater and increasing groundwater path length; and decreasing with increasing hydraulic conductivity of both the vadose zone and aquifer material (Van Meter and Basu, 2015; Vero et al., 2017; Wang et al., 2012). While deep (typically ≥60 m) groundwater systems can report lag times in the order of several decades or more, lags tend to be relatively short in shallow groundwater and surface water systems. For instance, in New Zealand time lags are reported to be 50–100 years in the deep groundwater systems of the North Island's central plateau (Morgenstern and Daughney, 2012), but <5 years in a study of 77 stream catchments representing about half the agricultural land in the country (McDowell et al., 2021). Additional lag components include variations in rainfall, soil storage, and soil nitrogen transformations and mineralisation; however these processes are typically less than one year, so we have not included them in this study as the vadose zone and groundwater transport lags are dominantly the rate limiting step (Trinsoutrot et al., 2000).

Understanding the state and trend of NO_3-N concentrations is a key requirement for resource management and to reduce NO_3-N concentrations to acceptable levels. Without this knowledge it is impossible to know if actions taken on land to address unacceptable nitrate contaminations are having an effect or are justified and there is a risk that groundwater management policies could be developed and implemented based on poor evidence (Ó Dochartaigh et al., 2007). Transient variability affects the likelihood of detecting NO_3-N concentration increases or decreases. Robust monitoring regimes should therefore capture enough samples over a long enough time to adequately characterise the variation in concentrations such that the detection power of the site (the statistical likelihood of detecting change) is high (Burt et al., 2011). Most current sampling regimes will take samples from a site at intervals ranging from annually (most commonly) to quarterly. More frequent samples are deemed unnecessary on the assumption that NO_3-N concentrations do not vary much (Frollini et al., 2021). However, NO_3-N concentrations in many groundwater systems, especially those supplying nearby streams, can vary considerably, often being enriched, or depleted by specific climate or land management events or oxygen status.

Groundwater quality monitoring design and review processes typically focus on the representativeness of sampling in terms of geographic spread and hydrogeological and/or water quality typologies (Morgenstern and Daughney, 2012; Ó Dochartaigh et al., 2007). Detection power analysis assesses the likelihood of detecting a statistically significant change (here a decrease in concentration) at a given site. Unfortunately, it is often neglected in scientific studies and monitoring programs, despite the recommendation that such analysis is essential from scientists, statisticians, and some existing guidance (European Commission, 2007; Weiser et al., 2021). Evaluation of the statistical power of a site within a network provides important information for network optimisation to maximise the chance of detecting an effect. For example, the resources used to sample sites with low detection power can be reallocated to additional monitoring to improve the monitoring regime (Field et al., 2007). However, few, if any, of these recommendations assess the importance of groundwater age on detection power. The outcome can be a network that may not yield robust information to support decisions within the timeframes required by land managers, custodians, and regulators.

Fundamentally, the aims of this paper is to demonstrate the need to include both detection power and the effects of NO_3-N lag (i.e., storage in the vadose and saturated zones) in NO_3-N mitigation and monitoring network design. We undertook this in three parts: Firstly, we sought to determine the power of the of New Zealand national network of monitored wells to detect reductions in NO_3-N concentrations. Specifically, we test the hypothesis that the current New Zealand network is able to detect the typical range of nitrate leaching reductions required under the New Zealand regulatory framework (e.g. (Ministry for the Environment, 2020) within practical timeframes for management purposes. Secondly, we endeavoured to demonstrate the impact of groundwater age and temporal dispersion on detection power analysis by testing the hypothesis: Excluding groundwater age distributions will significantly overestimate the ability of the New Zealand network to detect NO_3-N mitigations. Finally, to support additional monitoring, we tested the hypothesis that we can predict the detection power at unmonitored sites at a national scale.

2. Methods

2.1. Creating a national dataset of groundwater sites

New Zealand has significant groundwater resources across both the North and South Island. Most groundwater is hosted in relatively young unconsolidated, typically alluvial, sediments (White et al., 2019). We created a national dataset of all regional council groundwater nitrate monitoring sites and a companion dataset of all available groundwater age measurements in New Zealand and processed the data into a final dataset by: 1) cleaning the datasets; 2) removing outliers via an unsupervised Local Outlier Factor (Pedregosa et al., 2011) on the min-max normalised NO_3-N and sampling dates; 3) filtering the sites to those which had at least five measurements, taken over five years, with the mean and median gap between samples of one or less years. A summary of the data and the final dataset is provided in the Supplementary Information.

To estimate the natural variance (noise) of the NO_3-N datasets we first determined whether the site had a trend via seasonal Mann Kendall (SMK) analysis. Where the SMK identified a significant trend ($p < 0.05$), we conducted a simple linear regression and used the standard deviation of the residuals as the NO_3-N noise at that site. Where no significant trend was identified we used the standard deviation of the full dataset as the NO_3-N noise at that site.

2.2. A priori pathways

Detection power assessments estimate the probability that a site or network can statistically identify a change to the groundwater concentrations. National regulations in New Zealand require significant NO_3-N leaching rate reductions in large parts of the country to achieve a national bottom line of 2.4 mg L^{-1} NO_3-N in surface water (Ministry for the Environment, 2020). These regulations are implemented via Regional Plans which specify how, where, and when the reductions should occur. A Regional Plan life is normally 10 years (Resource Management Act, 1991) and stakeholders often wish to identify the effectiveness of mandated reductions after 5 years. Because regional authorities need to weigh environmental and economic factors when defining the rate of nitrate loss mitigation, most NO_3-N reductions to date have been set at 5–20 % reductions over the 10-year plan (Environment Canterbury, 2016, 2021). Therefore, we assessed the detection power of two unique a priori nitrate leaching concentration pathways. Both pathways assume than any changes in nitrate leaching concentrations are fully implemented, successful, and apply proportionally to all land parcels. The first pathway, "*To national bottom line*", is a reduction to 2.4 mg L^{-1} NO_3-N over 30 years, which is based on the national target for surface water. This is likely to be an over estimation of the groundwater reductions needed to achieve the policy goal for surface watercourses with significant runoff contributions from low intensity land use in the upper catchment, which is common in New Zealand. The second pathway, "*Typical reductions*", is a 1.5 % reduction per year from the starting concentration to 2.4 mg L^{-1} NO_3-N. For both pathways, once a site reaches the target 2.4 mg L^{-1} NO_3-N concentration it remains there for the remaining sampling duration. We specified a range of sampling durations (5, 10, 15, 30, and 50 years) and sampling frequencies (annually, quarterly, monthly, and weekly). An example set of a priori pathways is shown in Fig. 1.

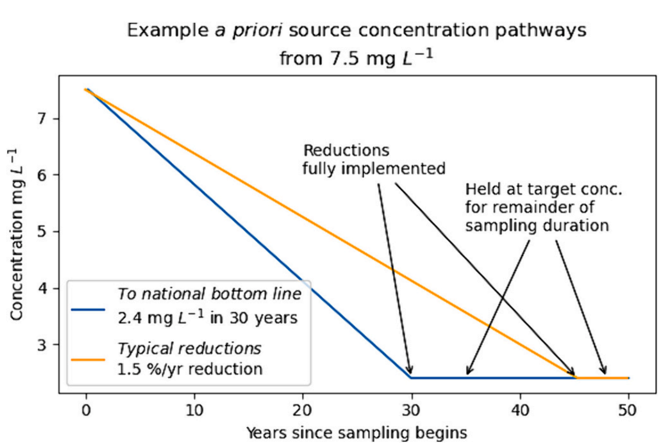


Fig. 1. An example of the a priori pathways from 7.5 to 2.4 mg L-1 NO_3-N.

2.3. Detection power assessment methodology for existing network

We estimated the detection power of all sites without considering groundwater travel processes (i.e., without lag) and including groundwater travel processes (i.e., with lag) for all sites which had recorded groundwater age analysis (e.g., via tritium sampling). Here we define:

- **detection power** as the percent probability that a random noisy time series concentration realisation will reject the null hypothesis for a chosen statistical test below a critical level (here we use the commonly accepted value of $p < 0.05$).
- **noisy concentration realisation** as the modelled noise-free receptor concentration plus a random noise term sampled from a normal distribution with a mean of zero and a scale defined by the observed historical noise.

We undertook all detection power analysis with our open-source detection power calculator (Dumont, 2023) which implements the following methodology. We used 1000 random noisy concentration realisations to estimate the percent probability of the statistical test rejecting the null hypothesis. We used a Mann-Kendall test as the statistical test for: 1) all sites excluding lag and 2) sites with lag and without a significant increasing historical trend. We used a two-part Mann-Kendall test as our statistical test for sites with lag and a significant increasing historical trend (Frollini et al., 2021). Briefly, the two-part Mann-Kendall technique identifies all breakpoints in the data where the expected trend (i.e., increasing, then decreasing) is identified by a Mann-Kendall test and is statistically significant ($p < 0.05$). We chose this methodology as the true receptor concentrations will initially increase (as the lagged effects of higher source concentrations continue to move through the groundwater system) before decreasing in response to a NO_3-N loss mitigation; fitting a simple Mann-Kendall trend would significantly underestimate the detection power as the initial increase would obscure the later reduction. A two-part Mann Kendall identifies this inflection point without a priori information (e.g., the time of the maximum concentration) and is therefore an appropriate analogue to real world detection. Note that our method assumes that NO_3-N behaves as a conservative contaminant; denitrification can reduce NO_3-N concentrations along the groundwater flow path in anoxic conditions (Rivett et al., 2008), which could impact detection power estimates.

For the lag free assessments, we modelled the noise free concentration by directly sampling the a priori pathway (e.g., "*To national bottom line*") at a given sampling frequency (e.g., quarterly) and duration (e.g., 5 years). As noted, the aforementioned methodology does not account for lag times between changes in source concentrations (e.g., nitrate leaching at the base of the root zone), and the associated concentration change in a monitoring well. Accommodation of lag times on detection power analyses requires quantification of two temporal components in groundwater flow processes: 1) the delay in water from the source reaching the well (lag); and 2) the mixing of waters of variable age along the flow path (temporal dispersion). Both components can be estimated from groundwater age tracer sampling and modelling (Małoszewski and Zuber, 1982). The methodology comprises sampling of an age tracer (e.g., tritium) in groundwater and modelling the mean age (aka mean residence time, MRT) and age distribution via a simple 1-D groundwater mixing model. Results are typically presented as the MRT and the mixing model parameters, which describe the age distribution (Małoszewski and Zuber, 1982). Most age interpretations in New Zealand use a single or binary exponential piston flow model ([B]EPFM, equation), which allows for one or more fast flow pathways and one or more slow flow pathways (Stewart, 2012a; Stewart and Thomas, 2008). The Exponential Piston Flow Model (EPFM) combines a flow path section with exponential transit times followed by a piston flow section, to give a model with parameters of mean residence time (τ_m) and exponential fraction (f). The response function is given by:

$$h(\tau) = \begin{cases} 0, & \tau \ < \ \tau_m(1-f) \\ (f\tau_m)^{-1} e^{\frac{\tau}{f\tau_m}+\frac{1}{f}-1}, & \tau \ \geq \ \tau_m(1-f) \end{cases}$$

where τ is the residence time, $h(\tau)$ is the flow model or response function of the hydrological system, f is the ratio of the exponential to the total volumes, and τ_m (1 - f) the time required for water to flow through the piston flow section (Stewart, 2012b).

A binary EPFM (BEPFM) model can be used to describe short-residence-time and long-residence-time flow components in groundwater system, for example near-surface and deep flows to a river, or shallow lateral and deep flows to a groundwater well (Stewart, 2012b). The BEPFM model is formed by adding the two EPM models:

$$BEPFM = b(EPFM_1) + (1-b)(EPFM_2)$$

New Zealand systems are typically modelled with [B]EPFM models because most aquifers are hosted in tertiary-quaternary alluvium (White et al., 2019) and feature a multi-porosity system with preferential flow pathways (Dann et al., 2008). By applying an appropriate model and parameter set we can re-produce the effects of both lag and temporal dispersion on the timing of changes in the well concentration relative to changes in the source concentration. This is important for detection power analysis for two reasons. Firstly, nitrate concentrations measured in a well are often not at equilibrium with the source leaching concentration and hence may change after a nitrate loss mitigation has been implemented for reasons unrelated to the mitigation action. Secondly, temporal dispersion can spread the effect of mitigations across a longer time spreading the peak concentration and reducing the slope of the observed reductions – effectively reducing the detection power.

We modelled the noise-free receptor concentration for each a priori pathway with lag for all sites with reported MRT data via a [B]EPFM model. Initially we attempted to supplement reported MRT data with a cross correlation lag approach but abandoned this technique due insufficient data (see Supplementary Information). The parameters for the [B]EPFM model are usually reported with the MRT value, but where these values were missing from our database we filled them with nearby groundwater age modelling results (Daughney et al., 2010). More information on this gap filling process is available in the Supplementary Information. We distinguished between sites with a significant historical increasing trend and those with no trend or with a historical decreasing trend. We modelled the noise-free receptor concentration for sites with an increasing trend by:

1) Modelling the historical source concentration (see Fig. 2) by optimising the slope of the lagged source concentration from a simple source concentration model (below) to the observed concentration trend slope in the well.
2) Applying the a priori pathway reductions (from the final historical source concentration modelled in 1 to the target concentration, e.g., 2.4 mg L^{-1} NO_3-N) to produce a future source concentration.
3) Calculating the lagged concentration in the monitoring well via the travel time model (e.g., exponential piston flow).

The simple source concentration model is defined as:

$$C(t) = \begin{cases} c_m, & c_0 \ < \ c_m \\ c_0, & c_0 \ < \ c_m \end{cases}, where\ c_0 \ = \ mt \ + \ b$$

where c_m= minimum concentration limit (mg L^{-1}), m=slope (parameterised, mg $L^{-1}Y^{-1}$), b_i=initial source concentration (parameterised, mg L^{-1}), and t= time (years).

We used a minimum concentration limit of 1 mg L^{-1} because our area of interest is medium to high intensity farmland where leaching concentrations >1 mg L^{-1} are expected. We limited the initial source concentration to 20 mg L^{-1} for all sites unless the maximum observed concentration was >20 mg L^{-1}. Where the maximum observed concentration was >20 mg L^{-1} and <30 mg L^{-1} we limited the maximum initial source concentration to 30 mg L^{-1} and all other sites were limited to the observed maximum concentration.

We modelled the lagged concentration for sites with a significant decreasing trend and those with no significant trend in the same fashion as those with a single increasing trend, but assuming that the source concentration was at steady state with the starting concentration (e.g., all historical concentrations were equivalent to the mean measured concentration for the last year), as per Fig. 2.

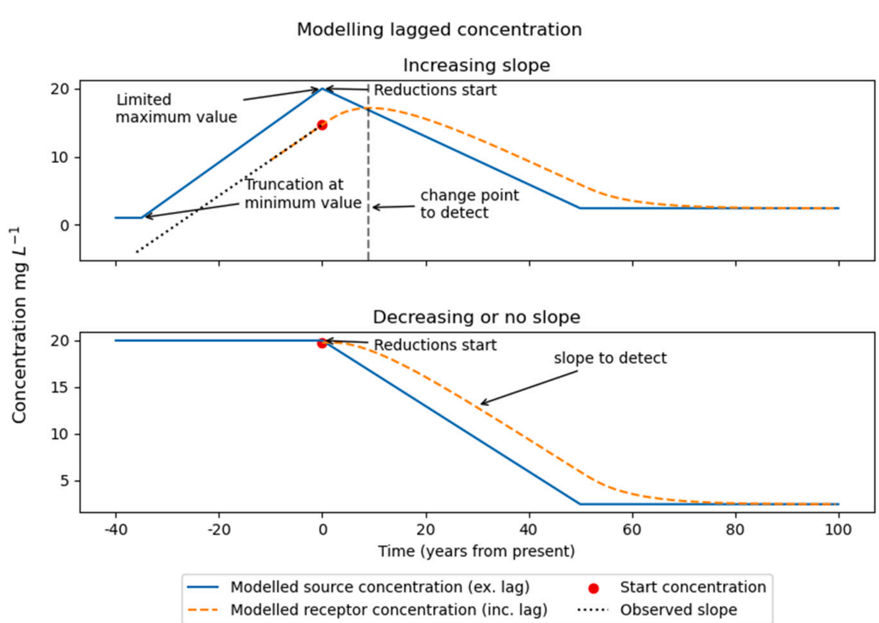


Fig. 2. An example of the methodology to determine source concentration and the lagged receptor concentration for sites with increase vs. decreasing/no slope. The receptor concentration reduction is shifted (lag) and the slope is reduced (temporal dispersion).

2.4. *Identification of new monitoring sites*

Detection of a set of common characteristics of high detection power sites would provide a basis for selection of new monitoring sites. We sought to achieve this by evaluating the extent to which detection power can be predicted at a national scale from conceptually relevant predictor variable data sets with national coverage. These included monitoring well depth, water table depth, hydrogeological unit classifications, modelled nitrate leachate history data, land surface recharge, and land use classifications. We used an agglomerative clustering method (Pedregosa et al., 2011) to develop clusters in the log transformed concentration and NO_3-N noise and then used a boosted gradient classifier (Pedregosa et al., 2011) to predict the cluster value. The median concentration and NO_3-N noise of the relevant cluster was used to calculate the modelled detection power of the site. Full details of our predictors and methodology are provided in the Supplementary Information.

3. Results and discussion

3.1. *Current monitoring network's detection power*

We used our detection power assessment method to evaluate the percentage of sites within the national groundwater quality monitoring network ($n = 465$) that would be likely to successfully detect NO_3-N loss reductions for both a priori pathways and for three sets of lag data: "*lag at lag sites*", comprising analysis which accounts for lag at all sites with a recorded MRT estimate; "*without lag at lag sites*", comprising analysis without lag accounted for at all sites with a recorded MRT estimate; and "*without lag at all sites*", comprising analysis without allowance for lag at all sites in our national dataset. Note that we only included sites with NO_3-N concentrations >2.4 mg L^{-1}. Here we establish a threshold for a statistical power of ≥ 80 %, as a power of this magnitude is typically accepted as having a 'high' likelihood that an effect is real after Di Stefano (2003). In addition, we use the common critical level of $p < 0.05$. Our results a presented in Fig. 3.

For the "*To national bottom line*" without lag assessment, c. 15 years

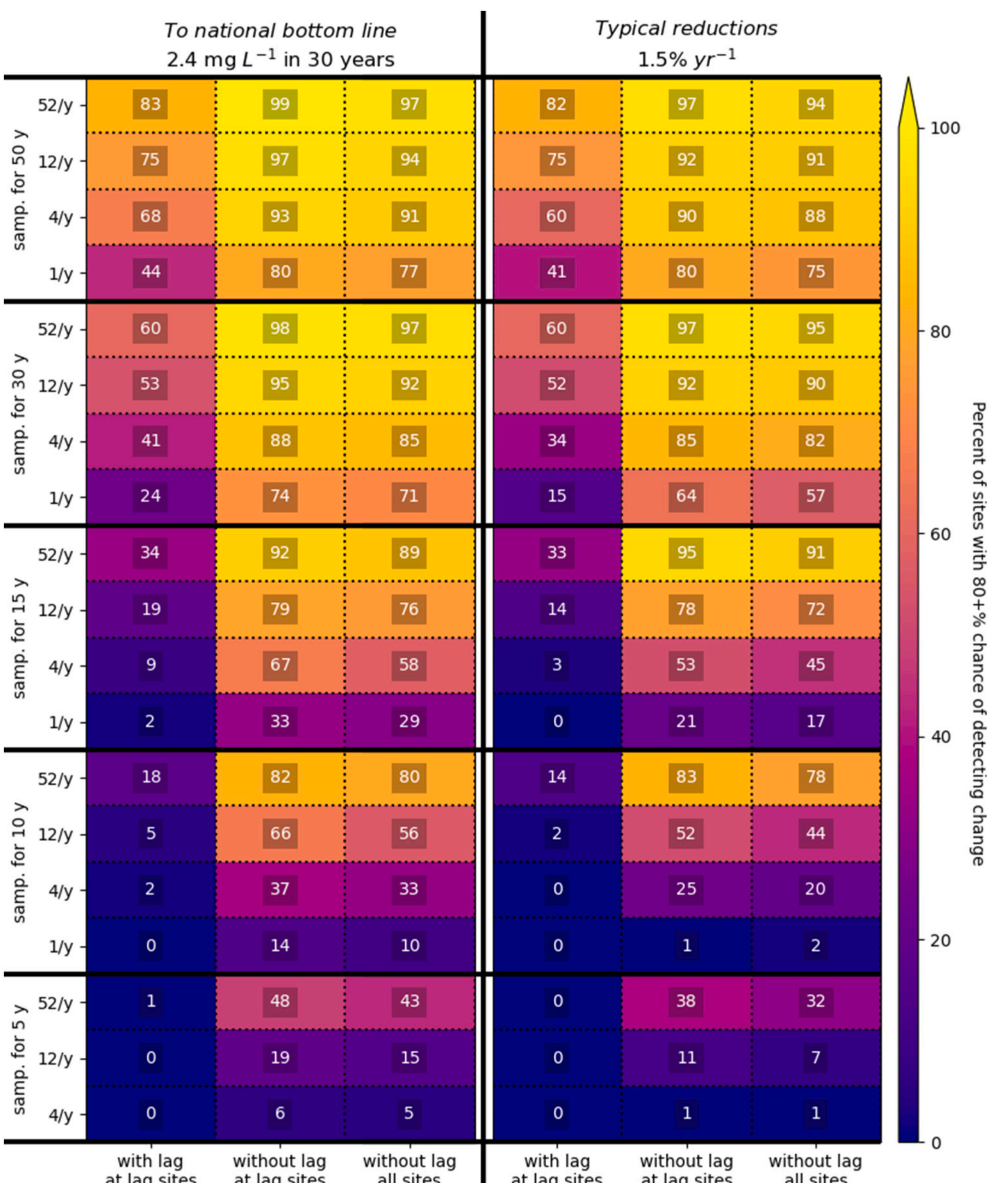


Fig. 3. Percent of sites with a ≥ 80 % chance of detecting a reduction in NO_3-N "*To national bottom line*": a reduction to 2.4 mg L^{-1} NO_3-N in 30 years and "*typical reductions*": a reduction of 1.5 % per year.

of quarterly interval sampling would be required before half (58 %) of the sites could detect the prescribed concentration change. The "*Typical reductions*" scenario typically has lower nitrate loss reductions which yield lower detection powers, particularly at lower sampling duration and sampling frequencies. Only 45 % of sites would achieve an 80 % or greater chance of detecting the reduction after 15 years of quarterly sampling. The detection power differences between the two concentration reduction pathways diminish at high sampling frequencies and durations (e.g., at 50, 30, 10, and 10 years for annual, quarterly, monthly, and weekly sampling, respectively).

Including groundwater travel processes in the "*To national bottom line*" scenario increases the sampling time required for an 80 % detection confidence to over 30 years. Increasing the sampling frequency increases the detection power of the network (including lag), but only after 10–15 years of sampling. This suggests that for the first 5–10 years the lack of detection power is due to groundwater travel processes. After 10–15 years the detection power is impacted both by NO_3-N noise and groundwater travel processes. This is further reinforced by the comparison of the lag free and lagged detection powers. After 5 years the lag free detection power increases markedly with increasing sampling frequency. Similar changes in the lagged detection power only occurs at 15 years. The detection power difference between the "*To national bottom line*" and "*Typical reductions*" observed in the lag-free assessment is less pronounced in the lagged assessments. This is likely because the difference between the two concentration pathways is outweighed by the significant sampling duration required for a high detection power. That said, a fixed or staggered reduction (e.g., 20 % in 10 years, and then 5 % per year afterwards) would likely have a more pronounced detection power difference.

3.2. Predicting detection power for new monitoring sites

We attempted to predict the detection power of unsampled sites using a boosted gradient classifier (Pedregosa et al., 2011). One advantage of this approach is that the analysis calculates the relative importance of the predictor variables. Our work identified land surface recharge and well depth as the most significant predictors of mean nitrate concentrations and noise. Distances to surface water bodies, and to a lesser extent the modelled nitrate leaching history, were also relatively important. Hydrogeological units, land use classifications, modelled redox status, and geological unit age were of low relative importance. Further results from our predictions are available in the Supplementary Information.

To understand the utility of the statistical model predictions we propagated the predicted NO_3-N noise and median concentration values from the test component of 100 model test:train realisations to detection power across all suites of a priori pathways, implementation times, sampling duration, and sampling frequencies. Fig. 4 shows the modelled (via gradient boost) vs observed (from Section 3.1) detection power binned by the observed detection power. The median-median values of each bin fall below the 1:1 line suggesting that the model has some negative bias, particularly as detection power increases. The spread of the model is large. For instance, the interquartile range for observed detection powers between 50 and 60 % is c. 20–80 %. This means that at

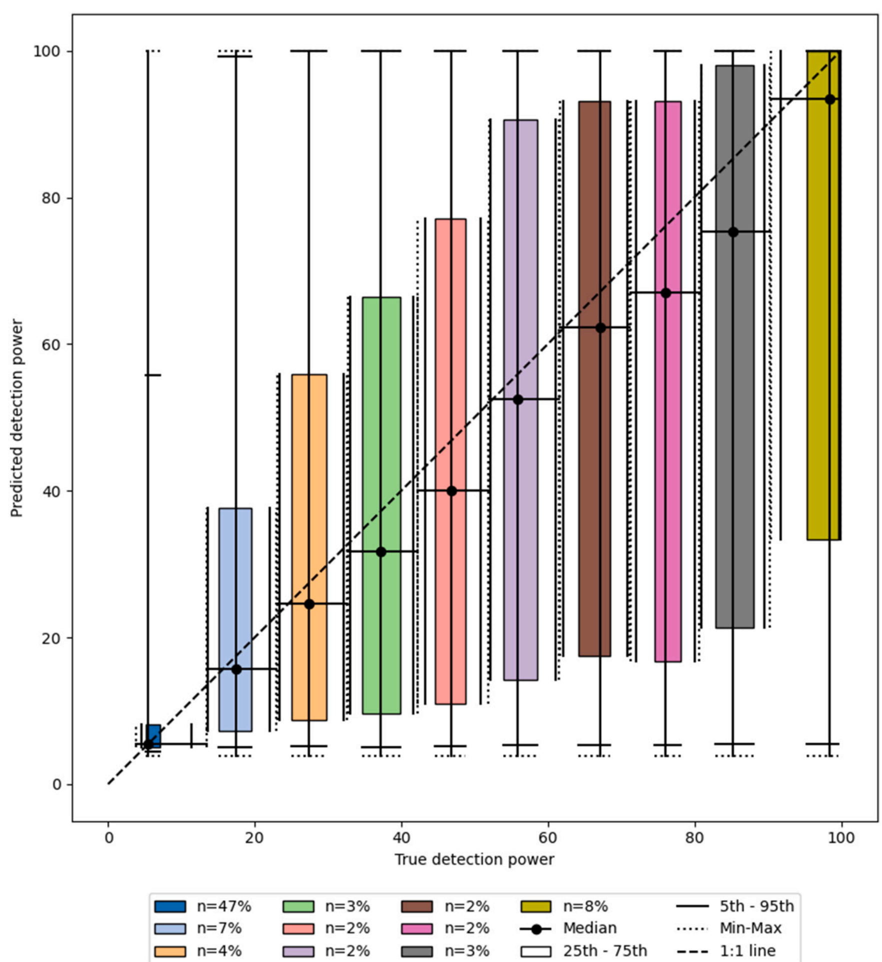


Fig. 4. The modelled vs measured detection power. This is a 2-d boxplot, the data was binned by its true detection power and the boxes and fliers present the distribution of the data on both the x and y axis.

least 50 % of the predictions in this range have a detection power error of c. 30 %. There is minimal disagreement with the model on low observed detection powers; however, this agreement is likely due to the low likelihood of detecting the change regardless of the characteristics of the site (e.g., short sampling duration, low sampling frequency, and minimal annual reductions). These results suggest that this national scale modelling is not able to identify the locations of likely high detection power sites, but application of this method using predictor datasets at regional, or catchment scale may yet yield useful results which would allow existing networks to be optimised for detection power by adding new high-power sites.

3.3. Is the current network capable of detecting likely nitrate reductions?

The current monitoring network includes the monitoring points (e.g., wells), the knowledge about these locations, and the current sampling frequency. From our national dataset, most sites are sampled quarterly or less frequently with a 5th, 50th, and 95th percentile of 1.1, 3, 6.3 samples per year, respectively. The network has groundwater travel time assessments at <20 % of sites and 58 % of those sites have a MRT of 10 years or more. Finally, approximately 30 % of monitoring points have increasing nitrate concentrations. Given these limitations we can only assess the monitoring network detection capacity under the assumption that the population of monitoring sites with MRT information provides an unbiased sample of the full monitoring network.

Table 1 highlights the proportion of the network capable of detecting changes within practical timeframes for water resource management purposes. Our results indicate that most of the network would be incapable of detecting "*Typical reductions*" within required timeframes (5–10 years), with very few sites likely to conclusively show whether NO_3-N loss mitigations have been implemented effectively after 5 years. The larger reductions of "*To national bottom line*" are more likely to be detected, but still only 9 % of sites would detect change halfway through the assumed implementation period and only 41 % of sites would detect the change after full implementation (30 years).

Compliance with NO_3-N reduction requirements in New Zealand is typically assessed through some combination of Farm Environment Plans and associated audits and/or nutrient budget modelling (Environment Canterbury, 2016, 2021). These modelling approaches are prone to bias and uncertainty due to the complex nature of the processes, missing information, and/or the nature of the assumptions used to model the processes (e.g., (Etheridge et al., 2018)). These uncertainties underscore the need for unambiguous monitoring-based information so that a) regulatory authorities can evaluate compliance; b) stakeholders and communities can be assured that the water quality targets will be achieved, and c) landowners can determine whether potentially significant investments in nitrate loss mitigation are effective. Due to the lack of detection power, the current monitoring network does not meet these needs.

Table 1
The percentage of the network capable of detecting a change after n years of quarterly sampling (current sampling frequency).

Change	Years of sampling	Percent of network with ≥80 % probability of detecting a reduction	
		Excluding lag	Including lag
"*Typical reductions*"	5	1	<1
1.5 % NO_3-N reduction	10	20	<1
per year to 2.4 mg L^{-1} NO_3-	15	45	3
N	30	82	34
	50	88	60
"*To national bottom line*"	5	5	<1
2.4 mg L^{-1} NO_3-N in 30	10	33	2
years	15	58	9
	30	85	41
	50	91	68

Fig. 5 demonstrates the impact of lag and NO_3-N loss reduction rates on the likelihood of detecting change in the Canterbury Plains area of New Zealand, where groundwater has NO_3-N concentrations up to 13 mg L^{-1}. The NO_3-N loss reductions needed to a 2.4 mg L^{-1} target would be significant (e.g., up to 80 % in 30 years). If actions were implemented to reduce groundwater concentrations to 2.4 mg L^{-1} within 30 years and we disregard the effects of lag, detection power analysis indicates that the current network will confidently detect reductions within 5–10 years. Accounting for lag processes shows that the network is far less capable of detecting these changes, however, with 30–50 + years of monitoring required.

If instead we focus on the "*Typical reductions*" of NO_3-N, even excluding lag, the current network is essentially incapable of detecting these changes within 30 years. This further reinforces the challenges of using the current network to understand whether land use mitigations are being implemented effectively and delivering the water quality outcomes sought under national policy.

3.4. How could network detection power be improved?

Although this study focuses on groundwater monitoring, surface water sampling is likely to detect nitrate concentration changes more quickly in many instances because NO_3-N discharges to surface water bodies typically follows faster near surface and shallow groundwater flow paths than those conveying nitrate to monitoring wells (c. 70 % of the New Zealand network is deeper than 10 m). The improved change detection associated with faster flow paths may be countered by higher noise, lower average rate of NO_3-N loss reductions in the upstream catchment (due to mixed land use), highly variable flow pathways and hence NO_3-N dilution in response to weather and climate patterns, and changes in the mean transit time of the baseflow component and thus surface water lag. Morgenstern et al. (2010) found that the mean transit time of the baseflow component of a small North Island, NZ stream varied from 1 to 100 years depending on the weather an climatic conditions. Integrated analysis of surface and groundwater monitoring site detection power would be required to support design of a cost and detection power-optimised change detection network, but this is beyond the scope of our study.

Options to improve the detection power of the groundwater network include selection of monitoring sites with short lag times and a high signal to noise ratio, increased sampling frequency, and targeted monitoring under an experimental design framework. Knowledge of groundwater lag and dispersion via age tracer sampling and interpretation are an essential and relatively cost-effective way to assess whether a monitoring well will be useful for detecting groundwater NO_3-N changes. Regular monitoring of sites with no age information is an inefficient use of resources; it may be years or even decades before the effects of land use changes are observed at the site. Data analysis in the absence of groundwater age information lead to a high probability of statistical error. Although our national scale statistical modelling was unable to robustly identify potential locations for high detection power monitoring sites a companion study was able to build such a model for surface water(McDowell et al., 2024), though this study did not include the effects of mean transit times on NO_3-N detection power. Therefore, local or regional scale application of this methodology may yet have the potential to reduce the error margin and yield useful information. This would be valuable given the high cost of installing new monitoring sites and the monitoring investment that is required before the statistical power and groundwater age of each new site is known.

Any new monitoring location and frequency must be chosen based on a combination the likelihood of detecting a change (detection power) and the likelihood of a change happening (e.g., representativeness). In the absence of knowledge about a new site, higher initial sampling frequency combined with regular NO_3-N noise and detection power analysis supports practitioners to evaluate the pros and cons of

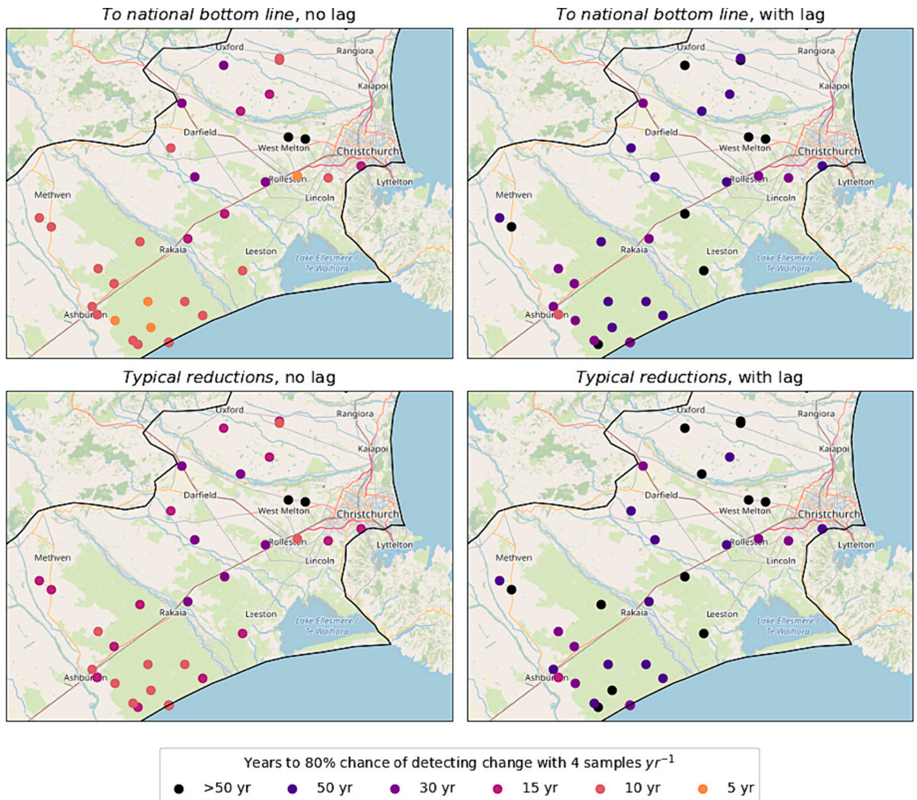


Fig. 5. A spatial example of the years required to have an 80 % chance to detect a reduction in NO_3-N concentrations in the Canterbury Plains (outlined) area of New Zealand with current quarterly sampling. The Canterbury Plains is comprised of thick (>300 m in places) unconsolidated deposits of Quaternary alluvium (mostly outwash). In some coastal areas the alluvium is interlaced with marine sediments.

monitoring at each site. They can then optimise the sampling frequency and costs to achieve the monitoring goal. To this end, we have produced an open-source Python package and a series of look up tables to allow users to estimate the detection power of their network (Dumont, 2023). Increasing the frequency of sampling could increase the detection power of the current monitoring system significantly. Our assessment indicates that an increase to weekly sampling from quarterly sampling can yield a five-fold increase in the percentage of sites capable of detecting changes within the required timeframe. Finally, targeted monitoring of monitoring wells with identifiable catchment areas under an experimental design framework such as the Before-After Control-Impact (BACI) methodology (see Stewart-Oaten and Bence, 2001) may prove to be a cost effective approach for determination of mitigation and/or policy effectiveness and progress rates.

3.5. Cost implications of higher frequency sampling for improved detection

The absolute cost of NO_3-N sampling is dependent on the geographic spread of sampling sites, human resources, and whether nitrate sampling visit costs are offset by broader sampling programmes. We obtained operational cost data from several regional authorities (Waikato Regional Council, Greater Wellington Regional council, Environment Canterbury, Horizons, and Environment Southland) to provide a general indication of monitoring cost per site. Average costs per site per visit fall within the \$130–250 NZD range, which includes transport, field equipment, and staff time, but excludes lab fees. Nitrate-N analysis costs are typically in the order of \$12 NZD. Other analytes are often tested at the same time raising the total cost but reducing the sampling cost per parameter.

Costs for age tracer lab analysis depends on the tracer, but range between \$400–900 NZD per sample excluding interpretation (GNS, 2023). Groundwater ages are typically considered stationary (e.g., Duvert et al., 2016) and therefore could be considered a one off expense. There is limited information in New Zealand about the temporal variation of tracers (e.g., tritium) and their derived ages. Stewart et al. (2011) reports mean residence times estimate over 30 years, which largely remain static, while Stewart (2012a) identified increasing groundwater ages in the Christchurch Aquifer system, but these were largely attributed to the ingress of deep old waters driven by increased abstraction associated with urbanisation. Manning et al. (2012) investigated the variation in MRT in springs wells of an alpine, snowfall dominated, headwater catchment. They found that mean residence time ages varied by <1.5–7 years depending on age methodology used. They conclude that these age variations appear to be controlled mainly by variations in the young fraction which was correlated with annual snow water equivalent. New Zealand groundwater is often situated lower in the catchment with less variable recharge therefore the variation observed by Manning et al. (2012) may be higher than what we should expect in the New Zealand context. Regardless any age estimate significantly constrains the detection power. Therefore, we suggest that single age estimates are essential. Repeated estimates would likely reduce the uncertainty of the detection power and may be worthwhile in younger or more sensitive sites. We also encourage further research on the variability of age estimates in these systems.

Table 2 provides an indicative assessment of the cost of increased sampling. These values are based on monitoring at all 948 sites of our national network. Analysing lag time at all sites would require additional resource comparable to the current national NO_3-N monitoring spend, but this would be a one-off investment rather than an ongoing cost. Upgrading the network from quarterly to monthly or weekly sampling would require a 200 % and 1200 % increase in funding, respectively. It is likely that not all sites would need to be monitored, but it is equally likely that additional sites would be required.

Table 2
Indicative additional costs (NZD) to increase the network's NO_3-N detection power.

Age tracer	Assumed cost per sample	Investment to date (age tracer at 20 % of sites)	Age tracer 50 % of sites	Age tracer 100 % of sites
	$1,190*	$221,340	$342,720	$906,780
Percent of current	*N/A*	100 %	155 %	410 %

NO_3-N	Annual costs Current#	Monthly sampling	Weekly sampling
	$815,280	$1,630,560	$9,783,360
Percent of current	100 %	200 %	1200 %

* $900 for analysis + $190 for sampling + $100 for interpretation.
\# Assumes quarterly sampling at a cost of $215/sample.

We can also estimate the minimum cost required to detect a change in the 465 monitoring points where NO_3-N is >2.4 mg L^{-1} (Table 3). An investment of c. $550,000 would be required to obtain groundwater age data for these sites without it and a further $1–3 million NZD would be required each year for 30 years for sampling. Approximately 40 % of sites will not have a detectable change after 30 years (with a maximum of weekly sampling). The wide cost range depicts quarterly or weekly sampling. Note we assumed the lowest cost approach (e.g., sampling at the minimum frequency to detect a change in 30 years).

Recent advances in in situ-logging NO_3-N sensors have potential to increase sampling frequency at lower cost. Assuming a nitrate logger cost of NZD $15,000 depreciated over 10 years, coupled with bimonthly site visits for logger maintenance and validation sampling, the annual cost per site would be in the order of NZ $3000 per year per site. This would equate to NZ $1.3 M per year for 465 sites. There may be other options, such as well owner sampling, citizen science, etc. to reduce the cost of sampling; however, these approaches require good regulator-citizen relations and are difficult to cost.

3.6. Limitations of the approach

Our methodology treats observed NO_3-N variance as noise under the assumption that observed variation is normal, independent of time, and contains no autocorrelation. The outcome of this assumption is an unconstrained increase in statistical power predictions with sampling frequency, regardless of the fact that higher frequency sampling will, at some point, have diminishing returns due to autocorrelation (Close, 1989). Constraining sampling frequency based solely on autocorrelation analysis results could be counterproductive. Sampling at a frequency that is impacted by autocorrelation could provide the data necessary for signal decomposition to remove what might otherwise appear to be random noise. Examples include rainfall event-based NO_3-N flushes from the soil and vadose zone to the water table, seasonal patterns in NO_3-N leaching rates from the soil profile (see (Trolove et al., 2019) and the effects of climate variability in response to interdecadal cycles on nutrient concentrations in surface waters (Snelder et al., 2022). Characterising these patterns and removing them from the monitoring data would improve the statistical power of the monitoring site. Our results also exclude any effects of denitrification, which could further complicate change detection (Boyer et al., 2006; Clague et al., 2019; Rivas et al., 2017). Finally, our analysis was undertaken on a national scale, and we therefore did not manually review our estimate of NO_3-N variance for each site. When we estimated our NO_3-N variance we assumed that the NO_3-N concentrations were either static, monotonically increasing, or monotonically decreasing. NO_3-N concentrations can vary due to a range of factors including rainfall, changes in land use and N leaching, and storage of NO_3-N in the soil profile. These impacts will all be considered noise for our analysis. We did not include these factors due to the national scale of the investigation, the paucity of moderate-high frequency NO_3-N data, and the fact that these methods are not widely used by regulators. Finally, our approach tested whether or not a decreasing NO_3-N trend of any scale could be detected. Stakeholders will also want to understand whether the monitoring results is consistent with the goal (e.g., the final target NO_3-N concentration), which requires a counterfactual approach and will likely affect the detection power of the site.

Table 3
Cost (NZD) to detect 2.4 mg L^{-1} NO_3-N in 30 years.

	Number Lag sites	Percent lag sites	Full Network	Cost	
NO_3-*N* > 2.4 mg L^{-1}	118	N/A	465	N/A	
Need MRT	0	0	465	$553,350	
No detection after 30 years	47	40 %	186	Quarterly	$159,960
				Weekly	$2,079,480
Sites with ≥80 % probability of detecting a reduction with a minimum of:					
Quarterly sampling	48	40 %	180	$154,800	
Monthly sampling	15	13 %	61	$157,380	
Weekly sampling	8	7 %	38	$424,840	
Total Cost	N/A	N/A	N/A	$896,980 - $2,816,500 per year + $553,350 (one off)	

4. Conclusion

Groundwater in New Zealand requires significant NO_3-N reductions to meet national policy objectives. Considering detection power, groundwater age, and temporal dispersion; only 41 % of the network can detect any reduction for a pathway to 2.4 mg L^{-1} over thirty years with 30 years of quarterly sampling. Given these results we must reject our hypothesis that the current network is well suited to detect NO_3-N mitigations in a timely fashion. Increasing sampling frequency can allow detection in up to 60 % of the network in the same period; however, 40 % of sites are incapable of detecting the change. The cost of increasing this detection power would be $1–3 million NZD per year or a c. 100–300 % increase over current expenditure. These results contrast with our estimates when we disregard groundwater age and temporal dispersion, which, incorrectly, suggests that these reductions would be widely detected (60–80 % detection) after just 10–15 years with only a modest increase from quarterly to monthly sampling. Therefore, we accept our second hypothesis, Excluding groundwater age will significantly overestimate the networks ability to detect NO_3-N mitigations. We conclude that current monitoring networks, which have not been explicitly designed for change detection (e.g., state of the environment networks), are unlikely to be suitable for change detection. Accurate and timely detection of NO_3-N mitigations will require new and bespoke monitoring networks. Although we produced a model to predict good monitoring locations at a national scale, the resulting imprecision precluded its use except for informing additional analyses; thus, rejecting our final hypothesis (we can predict detection power at a national scale). Hence, our advice to generate a robust network is to collect MRT data to inform us of potential lags and to undertake regular bespoke detection power analysis. We have produced a tool to simplify joint age and detection power analysis. This will help re-visiting the detection power of existing sites and the design of a robust network within financial constraints.

CRediT authorship contribution statement

M. Dumont: Writing, Resources, Data analysis. **Z. Etheridge:** Writing – review & editing, Project administration. **R.W. McDowell:** Writing, Resources, Data analysis, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We are grateful to Regional Authorities and the National Institute for Water and Atmospheric Research for providing the data. Funding to write this manuscript was provided by the Our Land and Water National Science Challenge (contract C10X1507 from the Ministry of Business, Innovation and Employment). We also thank Esther Menken for the cross-correlation analysis (see Supplementary Information).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2024.171759.

References

Ascott, M.J., Gooddy, D.C., Wang, L., Stuart, M.E., Lewis, M.A., Ward, R.S., Binley, A.M., 2017. Global patterns of nitrate storage in the vadose zone. Nat. Commun. 8 (1), 1416. https://doi.org/10.1038/s41467-017-01321-w.

Ascott, M.J., Gooddy, D.C., Fenton, O., Vero, S., Ward, R.S., Basu, N.B., Worrall, F., Van Meter, K., Surridge, B.W.J., 2021. The need to integrate legacy nitrogen storage dynamics and time lags into policy and practice. Sci. Total Environ. 781, 146698 https://doi.org/10.1016/j. scitotenv.2021.146698.

Boyer, E.W., Alexander, R.B., Parton, W.J., Li, C., Butterbach-Bahl, K., Donner, S.D., Skaggs, R.W., Grosso, S.J.D., 2006. Modeling denitrification in terrestrial and aquatic ecosystems at regional scales. Ecol. Appl. 16 (6), 2123–2142. https://doi.org/10.1890/1051-0761(2006)016[2123:mditaa]2.0.co;2.

Burt, T.P., Howden, N.J.K., Worrall, F., McDonnell, J.J., 2011. On the value of long-term, low-frequency water quality sampling: avoiding throwing the baby out with the bathwater. Hydrol. Process. 25 (5), 828–830. https://doi.org/10.1002/hyp.7961.

Camargo, J.A., Alonso, A., Salamanca, A., 2005. Nitrate toxicity to aquatic animals: a review with new data for freshwater invertebrates. Chemosphere 58 (9), 1255–1267. https://doi.org/10.1016/j.chemosphere.2004.10.044.

Carvalho, L., Mackay, E.B., Cardoso, A.C., Baattrup-Pedersen, A., Birk, S., Blackstock, K. L., Borics, G., Borja, A., Feld, C.K., Ferreira, M.T., Globevnik, L., Grizzetti, B., Hendry, S., Hering, D., Kelly, M., Langaas, S., Meissner, K., Panagopoulos, Y., Penning, E., Solheim, A.L., 2019. Protecting and restoring Europe's waters: an analysis of the future development needs of the water framework directive. Sci. Total Environ. 658, 1228–1238. https://doi.org/10.1016/j.scitotenv.2018.12.255.

Clague, J.C., Stenger, R., Morgenstern, U., 2019. The influence of unsaturated zone drainage status on denitrification and the redox succession in shallow groundwater. Sci. Total Environ. 660, 1232–1244. https://doi.org/10.1016/j.scitotenv.2018.12.383.

Close, M.E., 1989. Effect of serial correction on ground water quality sampling frequency. JAWRA J. American Water Resour. Assoc. 25 (3), 507–515. https://doi.org/10.1111/j.1752-1688.1989.tb03086.x.

Dann, R.L., Close, M.E., Pang, L., Flintoft, M.J., Hector, R.P., 2008. Complementary use of tracer and pumping tests to characterize a heterogeneous channelized aquifer system in New Zealand. Hydrgeol. J. 16 (6), 1177–1191. https://doi.org/10.1007/s10040-008-0291-4.

Daughney, C.J., Morgenstern, U., Van Der Raaij, R., Reeves, R.R., 2010. Discriminant analysis for estimation of groundwater age from hydrochemistry and well construction: application to New Zealand aquifers. Hydrgeol. J. 18 (2), 417–428. https://doi.org/10.1007/s10040-009-0479-2.

Di Stefano, J., 2003. How much power is enough? Against the development of an arbitrary convention for statistical power calculations. Funct. Ecol. 17 (5), 707–709. https://doi.org/10.1046/j.1365-2435.2003.00782.x.

Dumont, M., 2023. Komanawa groundwater detection power calculator [computer software]. https://github.com/Komanawa-Solutions-Ltd/komanawa-gw-detect-power.

Duvert, C., Stewart, M.K., Cendón, D.I., Raiber, M., 2016. Time series of tritium, stable isotopes and chloride reveal short-term variations in groundwater contribution to a stream. Hydrol. Earth Syst. Sci. 20 (1), 257–277. https://doi.org/10.5194/hess-20-257-2016.

Environment Canterbury, 2016. Canterbury Land and Water Regional Plan Volume1.

Environment Canterbury, 2021. Decision on provisions of proposed plan change 7 to the Canterbury land and water regional plan. https://www.ecan.govt.nz/your-region/plans-strategies-and-bylaws/canterbury-land-and-water-regional-plan/change-7/.

Etheridge, Z., Fietje, L., Metherell, A., Lilburne, L., Mojsilovich, O., Robson, M., Steel, K., Hanson, M., 2018. Collaborative expert judgement analysis of uncertainty associated with catchment-scale nitrogen load modelling with OVERSEER®. In: Farm Environment Planning—Science, Policy and Practice; Currie, LD, Christensen, CL, Eds.

European Commission, 2007. Common implementation strategy for the water framework directive (2000/60/EC). Guidance document no. 15: guidance on groundwater monitoring. European Commission. https://circabc.europa.eu/sd/a/e409710d-f1c1-4672-9480-e2b9e93f30ad/Groundwater%20Monitoring%20Guidance%20Nov-2006_FINAL-2.pdf.

Field, S.A., O'Connor, P.J., Tyre, A.J., Pissingham, H.P., 2007. Making monitoring meaningful. Austral Ecol. 32 (5), 485–491. https://doi.org/10.1111/j.1442-9993.2007.01715.x.

Frollini, E., Preziosi, E., Calace, N., Guerra, M., Guyennon, N., Marcaccio, M., Menichetti, S., Romano, E., Ghergo, S., 2021. Groundwater quality trend and trend reversal assessment in the European Water Framework Directive context: an example with nitrates in Italy. Environ. Sci. Pollut. Res. 28 (17), 22092–22104. https://doi.org/10.1007/s11356-020-11998-0.

GNS, 2023. The tritium and water dating laboratory price List. https://www.gns.cri.nz/assets/Laboratory-files/Tritium-and-Water-Dating-Laboratory/The-Tritium-and-Water-Dating-laboratory-price-list.pdf.

Horak, C.N., Assef, Y.A., Miserendino, M.L., 2019. Assessing effects of confined animal production systems on water quality, ecological integrity, and macroinvertebrates at small piedmont streams (Patagonia, Argentina). Agric Water Manag 216, 242–253. https://doi.org/10.1016/j.agwat.2019.01.026.

Howden, N.J.K., Burt, T.P., Worrall, F., Whelan, M.J., Bieroza, M., 2010. Nitrate concentrations and fluxes in the River Thames over 140 years (1868–2008): are increases irreversible? Hydrol. Process. 24 (18), 2657–2662. https://doi.org/10.1002/hyp.7835.

Larned, S., Booker, D., Dudley, B., Moores, J.,.R.,.M., Baillie, B., Schallenberg, M., Moriarty, E., Zeldis, J., Short, K., 2018. Land-use impacts on freshwater and marine environments in New Zealand (p. 291). NIWA. https://www.mfe.govt.nz/sites/default/files/media/Fresh%20water/land-use-impacts-on-freshwater-and-marine-environments%20.pdf.

MacDonald, A.M., Darling, W.G., Ball, D.F., Oster, H., 2003. Identifying trends in groundwater quality using residence time indicators: an example from the Permian aquifer of Dumfries, Scotland. Hydrogeol. J. 11 (4), 504–517. https://doi.org/10.1007/s10040-003-0275-3.

Małoszewski, P., Zuber, A., 1982. Determining the turnover time of groundwater systems with the aid of environmental tracers: 1. Models and their applicability. J. Hydrol. 57 (3–4), 207–231.

Manning, A.H., Clark, J.F., Diaz, S.H., Rademacher, L.K., Earman, S., Niel Plummer, L., 2012. Evolution of groundwater age in a mountain watershed over a period of thirteen years. J. Hydrol. 460–461, 13–28. https://doi.org/10.1016/j.jhydrol.2012.06.030.

McDowell, R.W., Noble, A., Pletnyakov, P., Haggard, B.E., Mosley, L.M., 2020. Global mapping of freshwater nutrient enrichment and periphyton growth potential. Sci. Rep. 10 (1), 3568. https://doi.org/10.1038/s41598-020-60279-w.

McDowell, R.W., Simpson, Z.P., Ausseil, A.G., Etheridge, Z., Law, R., 2021. The implications of lag times between nitrate leaching losses and riverine loads for water quality policy. Sci. Rep. 11 (1), 16450 https://doi.org/10.1038/s41598-021-95302-1.

McDowell, R.W., Noble, A., Kittridge, M., Ausseil, O., Doscher, C., Hamilton, D.P., 2024. Monitoring to detect changes in water quality to meet policy objectives. Sci. Rep. 14 (1), 1914. https://doi.org/10.1038/s41598-024-52512-7.

Ministry for the Environment, 2020. National Policy Statement for Freshwater Management 2020.

Morgenstern, U., Daughney, C.J., 2012. Groundwater age for identification of baseline ground- water quality and impacts of land-use intensification – the National Groundwater Monitor- ing Programme of New Zealand. J. Hydrol. 456–457, 79–93. https://doi.org/10.1016/j.jhydrol.2012.06.010.

Morgenstern, U., Stewart, M.K., Stenger, R., 2010. Dating of streamwater using tritium in a post nuclear bomb pulse world: continuous variation of mean transit time with streamflow. Hydrol. Earth Syst. Sci. 14 (11), 2289–2301. https://doi.org/10.5194/hess-14-2289-2010.

Ó Dochartaigh, B.É., MacDonald, A.M., Griffiths, K.J., Lilly, A., DeGroote, J., Chilton, P. J., Hughes, A.G., 2007. Assessing the effectiveness of Scotland's groundwater nitrate monitoring network. Q. J. Eng. Geol. Hydrogeol. 40 (4), 393–406. https://doi.org/10.1144/1470-9236/07-029.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Rahman, A., Mondal, N.C., Tiwari, K.K., 2021. Anthropogenic nitrate in groundwater and its health risks in the view of background concentration in a semi arid area of

Rajasthan, India. Sci. Rep. 11 (1), 9279. https://doi.org/10.1038/s41598-021-88600-1.
Resource Management Act, (1991).
Rivas, A., Singh, R., Horne, D., Roygard, J., Matthews, A., Hedley, M.J., 2017. Denitrification potential in the subsurface environment in the Manawatu River catchment, New Zealand: indications from oxidation-reduction conditions, hydrogeological factors, and implications for nutrient management. J. Environ. Manage. 197, 476–489. https://doi.org/10.1016/j.jenvman.2017.04.015.
Rivett, M.O., Buss, S.R., Morgan, P., Smith, J.W.N., Bemment, C.D., 2008. Nitrate attenuation in groundwater: a review of biogeochemical controlling processes. Water Res. 42 (16), 4215–4232. https://doi.org/10.1016/j.watres.2008.07.020.
Snelder, T.H., Larned, S.T., Fraser, C., De Malmanche, S., 2022. Effect of climate variability on water quality trends in New Zealand rivers. Mar. Freshw. Res. 73, 20–34.
Stewart, M.K., 2012a. A 40-year record of carbon-14 and tritium in the Christchurch groundwater system, New Zealand: dating of young samples with carbon-14. J. Hydrol. 430–431, 50–68. https://doi.org/10.1016/j.jhydrol.2012.01.046.
Stewart, M.K., 2012b. A 40-year record of carbon-14 and tritium in the Christchurch groundwater system, New Zealand: dating of young samples with carbon-14. J. Hydrol. 430–431, 50–68. https://doi.org/10.1016/j.jhydrol.2012.01.046.
Stewart, M.K., Thomas, J.T., 2008. A conceptual model of flow to the Waikoropupu Springs, NW Nelson, New Zealand, based on hydrometric and tracer (18O, Cl,3H and CFC) evidence. Hydrol. Earth Syst. Sci. 12, 1–19.
Stewart, M., Stevens, G., Thomas, J., Raaij, R., Trompetter, V., 2011. Nitrate sources and residence times of groundwater in the Waimea Plains, Nelson. J. Hydrol. N. Z. 50.
Stewart-Oaten, A., Bence, J.R., 2001. Temporal and spatial variation in environmental impact assessment. Ecological monographs 71 (2), 305–339. https://doi.org/10.1890/0012-9615(2001)071[0305:TASVIE]2.0.CO;2.
Trinsoutrot, I., Recous, S., Bentz, B., Linères, M., Chèneby, D., Nicolardot, B., 2000. Biochemical quality of crop residues and carbon and nitrogen mineralization kinetics under nonlimiting nitrogen conditions. Soil Sci. Soc. Am. J. 64 (3), 918–926. https://doi.org/10.2136/sssaj2000.643918x.
Trolove, S., Thomas, S., Van Der Klei, G., Beare, M., Cichota, R., Meenken, E., 2019. Nitrate leaching losses during pasture renewal – effects of treading, urine, forages and tillage. Sci. Total Environ. 651, 1819–1829. https://doi.org/10.1016/j.scitotenv.2018.09.333.
Van Meter, K.J., Basu, N.B., 2015. Catchment legacies and time lags: a parsimonious watershed model to predict the effects of legacy storage on nitrogen export. PloS One 10 (5), e0125971. https://doi.org/10.1371/journal.pone.0125971.
Vero, S.E., Healy, M.G., Henry, T., Creamer, R.E., Ibrahim, T.G., Richards, K.G., Mellander, P.-E., McDonald, N.T., Fenton, O., 2017. A framework for determining unsaturated zone water quality time lags at catchment scale. Agr Ecosyst Environ 236, 234–242. https://doi.org/10.1016/j.agee.2016.12.001.
Wagenhoff, A., Clapcott, J.E., Lau, K.E.M., Lewis, G.D., Young, R.G., 2017. Identifying cong- ruence in stream assemblage thresholds in response to nutrient and sediment gradients for limit setting. Ecol. Appl. 27 (2), 469–484. https://doi.org/10.1002/eap.1457.
Wang, L., Stuart, M.E., Bloomfield, J.P., Butcher, A.S., Gooddy, D.C., McKenzie, A.A., Lewis, M.A., Williams, A.T., 2012. Prediction of the arrival of peak nitrate concentrations at the water table at the regional scale in Great Britain. Hydrol. Process. 26 (2), 226–239. https://doi.org/10.1002/hyp.8164.
Weiser, E.L., Diffendorfer, J.E., Lopez-Hoffman, L., Semmens, D., Thogmartin, W.E., 2021. TrendPowerTool: a lookup tool for estimating the statistical power of a monitoring program to detect population trends. Conserv. Sci. Pract. 3 (7), e445 https://doi.org/10.1111/csp2.445.
White, P., Moreau, M., Mourot, F., Rawlinson, Z., 2019. New Zealand groundwater atlas: hydrogeological-unit map of New Zealand. GNS Consultancy Report 144.