Modelling Report - Nitrate Leaching in Southland

Callum Collier

September 13, 2021

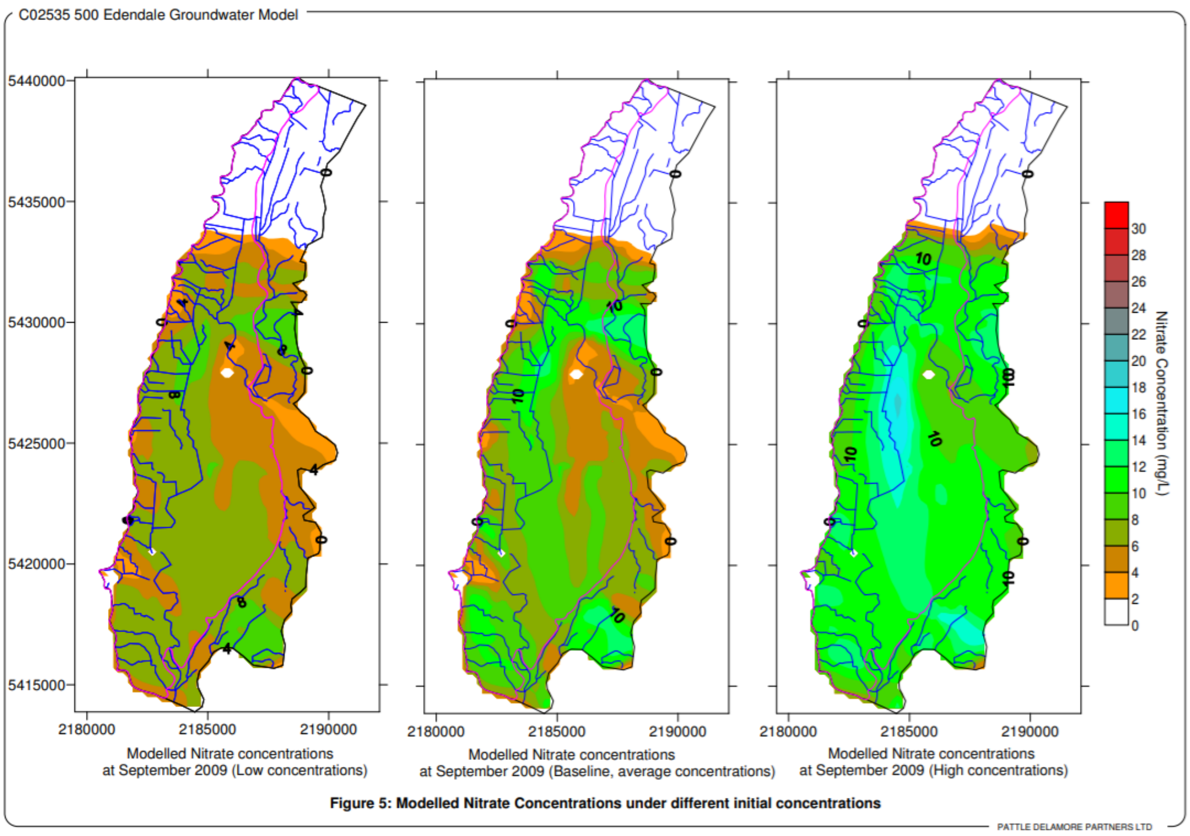

Figure 1: Nitrate Concentrations in Edendale under different initial conditions

Contents

1 Why?

1.1 Problem Description

The groundwater in Southland is of vital importance due to its use as drinking water, as well its potential impact on the environment of rivers and streams within the region. Of particular interest is the steady increase in the concentration of nitrate, which can be heavily linked to increased farming practices in the region. These farming practices are vital for the economy of Southland, however, addressing this increasing concentration in nitrates within the groundwater is also necessary.

1.2 Stakeholders

There are several main stakeholders who are affected by this groundwater pollution or the possible proposals to decrease nitrate concentrations. These include:

- Environment Southland – Primarily concerned with nitrate leaching and the effects this could have on the environment.
- Dairy farmers (Dairy NZ) – Concerned with the financial well being of the farmers and dairy industry.
- Ngāi Tahu – have financial interest. However, due to the harvest of mahinga kai from the Mataura River, also have concerns with nitrate leaching and its environmental impacts.

1.3 Possible Outcomes

In order to address groundwater quality concerns for the review of their resource consent by Environment Southland at a demonstration farm in Edendale, Dairy NZ is proposing to use MAR (managed aquifer recharge) to improve aquifer quality. There are two main points of interest:

- First, to find out whether the installation of an extensive active carbon sink (wood chip layer) at the farm in 2010 has been effective in reducing nitrate concentration levels in groundwater (as Dairy NZ claim).
- Second, to design the proposed MAR program which will be used to improve aquifer quality. (MAR involved injecting water at the high pressure boundary of the aquifer to increase the rate at which nitrate is flushed out of the aquifer).

If both options prove ineffective, other options, such as reducing the scale of dairy cattle farming, may have to be considered, which could have considerable impacts on Dairy Farming in the Southland region.

2 How?

Our modelling study aims to capture the relationship between the pressure in the aquifer and concentration of nitrate in groundwater over time, as well as the effect that increasing cattle numbers has had on the rate of nitrate leaching. Any economic effects resulting from the implementation of schemes to reduce nitrate concentrations will be considered outside the scope of this model.

The Ministry of Health outlines in the DWSNZ (Drinking Water Standard New Zealand) that the MAV (Maximum Allowable Value) of nitrates is 11.3 milligrams per litre and considers water with nitrate concentrations in excess of 75% of the MAV as 'significantly elevated'. [1]

The results from our modelling study will provide insight into the effectiveness that the active carbon sink has had on reducing nitrate levels in groundwater. Additionally, we aim to see how different MAR programs will reduce nitrate concentrations in the groundwater over time, how long it will take to see these reductions and how sustainable these measures are in keeping nitrate levels low enough for future generations. This will help us to determine the best MAR program to be proposed for the resource consent.

[1] Thomas N. (2012). *Edendale Groundwater Model.* Pattle Delamore Partners Ltd, Report prepared for Environment Southland.

3 Given?

One geothermal model that could be of use is a Lumped Parameter Models (LPM). LPMs conceptualise the geothermal reservoir as a single block whose pressure changes in response to fluid extraction, and inflows of adjacent groundwater to fill the pressure low that is created[2]. These can be represented as an ODE and solved analytically or numerically.

While other geothermal models exist, the LPM one is ideal for this model as pressure change within the reservoir can be used to model the change in groundwater nitrate concentration - especially when considering MAR schemes, which involve fluid injection into the aquifer in order to flush out polluted groundwater.

3.1 Data Visualisation

The data available to help model this problem are the cattle stock numbers since 1990 and nitrate concentrations from an Edendale bore since 1980. These data are plotted in Figure 2. Whilst pressure data is not available, the underlying ODE modelling pressure change within the aquifer is known so the initial pressure within the aquifer can simply be treated as an unknown parameter. Additionally, previously conducted studies of aquifers within the Southland region reveal a time delay between two and five years [3]. This time delay represents the response of nitrate concentrations in the groundwater with respect to surface nitrate sources.

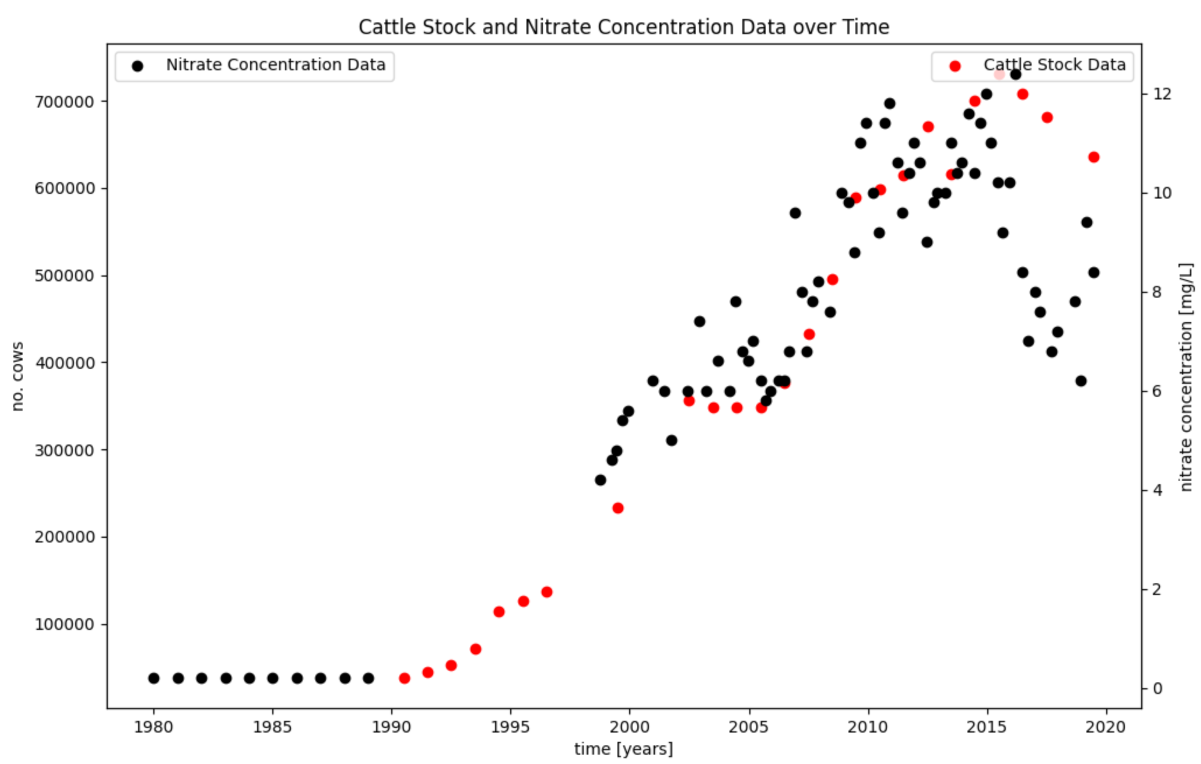

Figure 2: Cattle stock numbers in Southland since 1990 (red dots). Nitrate concentrations at a bore site in Edendale since 1980 (black dots).

3.2 Data Description

There is a roughly steady increase in the cattle numbers in Southland since 1990 and following along this increase are the nitrate concentration measurements (mg/L). Variation in the concentration measurements can be attributed to natural effects such as variations in annual rainfall which affect the overall mass of nitrate that leaches into the system. Prior to 1990 there is no available data for the number of cattle, and the concentration of nitrate is taken as a constant level of 0.2 mg/L. As the number of cattle increases approximately linearly from 40,000 to 360,000, the nitrate concentration also increases approximately linearly to about 6 mg/L, beyond which the nitrate concentration exhibits various peaks and drops. The nitrate concentration experiences a peak of about 12.5 mg/L in 2015, where the annual cattle numbers are experiencing a decrease. This indicates the time delay between changes in the nitrate source and the nitrate concentration within the aquifer.

Pressure effects are assumed to be immediate in the aquifer system, with the driving force for nitrate transport across the boundary being a surface over-pressure of 0.05 MPa. The pressure drop across the local aquifer is roughly 0.1 MPa which we observe as the pressure gradient across the aquifer.

[2]Fradkin, LJ, ML Storey, A McNabb, (1981). "On Identification and Validation of Some Geothermal Models", *Water Resources Research*, 17, 929-936.
[3]Wilson et al. (2014). *Estimating Time Lags for Nitrate Response in Shallow Southland Groundwater.* Lincoln Agritech Tech Rep, 2014-03.

4 Assume?

4.1 Relevant Physical Processes

The physics relevant to a geothermal systems are:

- Conservation of mass, which governs changes in fluid pressure within the aquifer. Fluid flow is by diffusion and is described by Darcy's law.
- Conservation of solute, which governs the concentration of nitrates within the groundwater. The flow of chemical solutes such as nitrates are both advective (carried by moving fluid) and diffusive (moving through porous rock to reach the aquifer).
- Conservation of momentum, which governs changes in stress and deformation. Most rocks are well-described by linear elasticity.

Our main concern here is understanding the fluid pressure in the aquifer, since pressure impacts the rate at which solutes such as nitrates are flushed out of the aquifer and whether any implemented schemes will be sustainable.

While our goal is to model nitrate concentrations within the aquifer, the exact physics underlying the transport of this solute into and throughout the aquifer are assumed to be negligible and will therefore be neglected. This, however, may introduce inaccuracies as it will not take into account local/spatial nitrate concentrations and will instead assume that the nitrate is evenly spread across the aquifer.

Stress and deformation physics will also be neglected, which may also introduce inaccuracies in the pressure model because poroelastic effects can buffer pressure changes.

4.2 Spatial/Temporal Domains and Boundary Conditions

The LPM approach requires that we represent the entire geothermal reservoir as a single control volume. This is the spatial domain, which for our model is the Edendale demonstration farm. While this is a relatively small domain, the results of our model will be used to gauge the effects of carbon sinks and MAR schemes for the whole of Southland, which will obviously introduce inaccuracies due to differing aquifer types, varying pressure gradients and varying intensities of farming. Any spatial variation within the Edendale aquifer will also be considered negligible i.e. the effects of localise temperature, elevation and point/anthropogenic sources will not be considered.

The time domain shall be from the beginning of when we have data (1980) until 30 years after the present day (2050). This time domain allows for the time delay between surface nitrate and infiltration into groundwater to be observed and ensures that any MAR scheme implemented is sustainable for the foreseeable future before other measures can be implemented to control and monitor nitrate concentrations.

The initial condition is the fluid pressure in the aquifer. Since this initial pressure is unknown, it is necessary to treat it as an unknown parameter of the model and to determine its value through calibration.

The effect of water sources that act as inputs and outputs of groundwater in the aquifer are included as an LPM, b, in the pressure ODE. This is effectively a boundary condition and the strength of this boundary condition is unknown so will also need to be determined through calibration. The values of the parameters b_1, α, and τ are known to be 0.0001, 0.3 and 2 respectively.

4.3 Conceptual Model

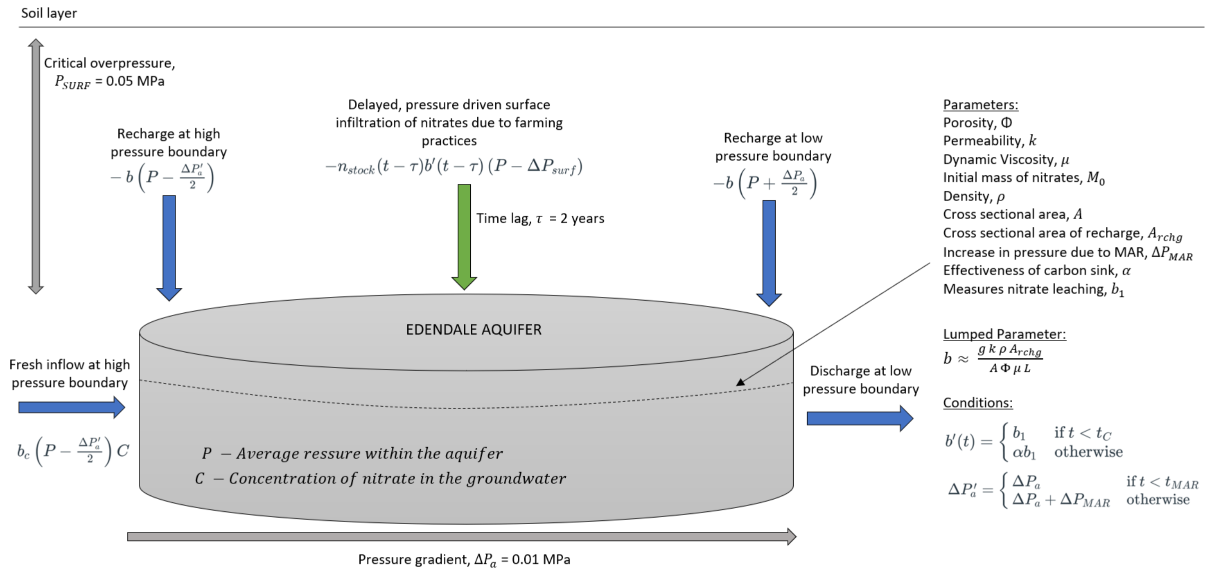

Figure 3: Conceptual model of the Edendale Aquifer

5 Formulate?

5.1 LPM ODE Formulation

The LPM ODE describing pressure changes within the aquifer is

$$\frac{dP}{dt} = -b(P + \frac{\Delta P_a}{2}) - b(P - \frac{\Delta P_a'}{2}) \qquad \Delta P_a' = \begin{cases} \Delta P_a & t < t_{MAR} \\ \Delta P_a + \Delta P_{MAR} & otherwise \end{cases}$$

where P (the independent variable) is the pressure in MPa, t (the dependent variable) is the time in years, ΔP_a is the pressure drop across the aquifer, ΔP_{MAR} is the pressure increase at the high pressure boundary and b is a lumped parameter that depends on physical parameters of the system.

Next, the LPM ODE describing concentration of nitrate within the aquifer groundwater is

$$M_0\frac{dC}{dt} = -n_{stock}(t-\tau)b'(t-\tau)(P - \Delta P_{surf}) + b_c(P - \frac{\Delta P_a'}{2})C \qquad b'(t) = \begin{cases} b_1 & t < t_C \\ \alpha b_1 & otherwise \end{cases}$$

where C (the independent variable) is the concentration of nitrate in mg/L, t (the dependent variable) is the time in years, $n_{stock}(t)$ is the number of cattle stock as a function of time, $b'(t)$ is a piece-wise parameter representing surface nitrate leaching into the aquifer, α is the effect of the carbon sink on the extent of nitrate leaching into the aquifer, ΔP_{surf} is the surface over-pressure, τ is the time delay between changes in the nitrate source at the surface and nitrate input into the aquifer and b_c is a lumped parameter that depends on physical parameters of the system.

5.2 Analytical Solution

We can find a simple analytical solution to the above ODEs by making several simplifications. For the pressure ODE, we can assume that $t < t_{MAR}$ for all time, meaning $\Delta P_a' = \Delta P_a$, giving,

$$\frac{dP}{dt} = -2bP, \qquad P(t) = Ae^{-2bt}$$

where A is a constant of integration. For the concentration ODE, we can once again assume $t < t_{MAR}$ and $t < t_c$, meaning $b'(t) = b_1$. We can further simplify the equation by solving the ODE for a constant number of cattle and setting this to zero i.e. $n_{stock}(t) = n_0 = 0$. Finally, substituting in the analytical pressure ODE solution above gives,

$$\frac{dC}{dt} = -g(Ae^{-2bt} - \Delta P_{surf}) + h(Ae^{-2bt} - \frac{\Delta P_a}{2})C, \qquad g = \frac{b_1 n_0}{M_0}, \quad h = \frac{b_c}{M_0}$$

which can be further simplified by setting $n_0 = 0$, in which case,

$$\frac{dC}{dt} = h(Ae^{-2bt} - \frac{\Delta P_a}{2})C, \qquad C(t) = A_1(e^{-\frac{Ah}{2}e^{-2bt}})(e^{-\frac{\Delta P_a}{2}t})$$

where A_1 is another constant of integration. The solution to this ODE tends exponentially towards the steady-state $C = 0$ at infinite time. This makes sense intuitively as without any nitrate source all the nitrate within the groundwater will eventually be flushed out by the pressure gradient across the aquifer.

6 Working?

We have checked that the Improved Euler implementation is working correctly by running a unit test that compares the output to a by-hand example. This code is given in **main.py**. The above simplified analytical solution can also be used to check that the Improved Euler implementation is working correctly by benchmarking the numerical solution of the LPM against this analytical solution. The results are shown in Figure 4 and indicate that the solver is working as expected. The code for this benchmarking can also be found in **main.py**.

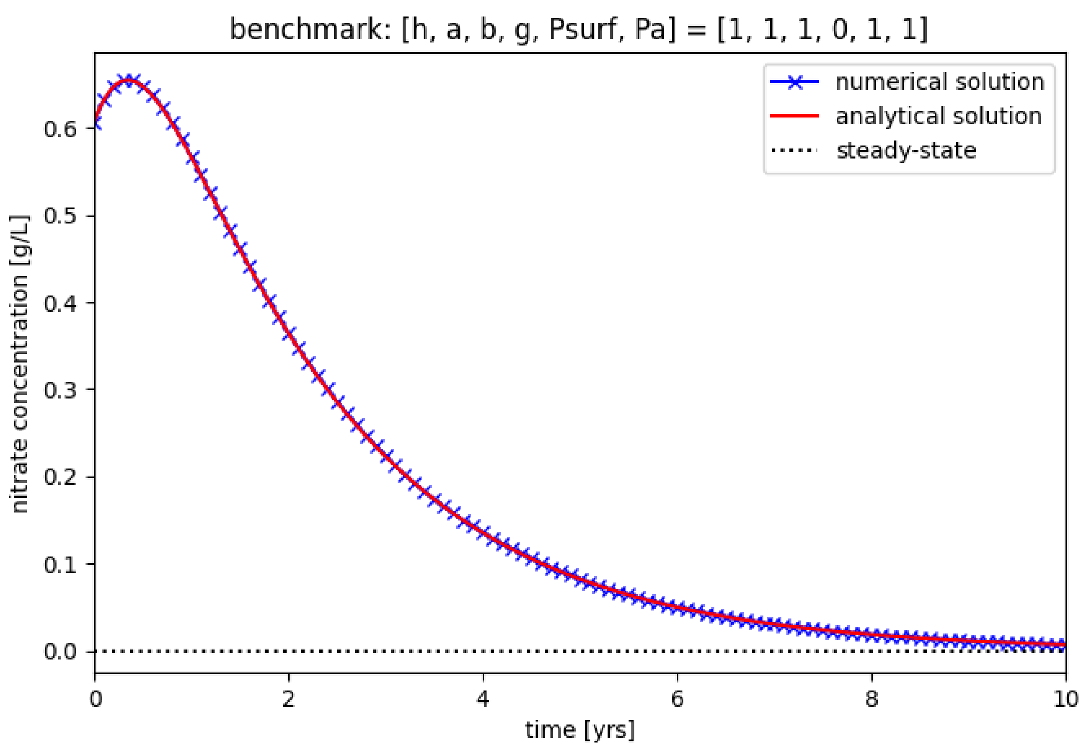

Figure 4: Comparison of numerical solution to analytical benchmark.

7 Suitable?

In order to fit our model, we first had to find a reasonable input for the n_{stock} term. Therefore, we first fit a third order polynomial to the cattle stock data which would then be used as input into the concentration LPM ODE. This polynomial fit can be seen in Figure 5 below.

We then manually calibrated the LPM model to find the unknown parameters that resulted in the best-fit. This model has parameters [b, bc, P0, M0] = [24, 3, 415, 1], with a time delay, τ, of two years. (P0 and M0 represent the initial fluid pressure and initial total mass of nitrate within the aquifer respectively). This best-fit model of the nitrate concentration (black line) can be seen in Figure 6 below.

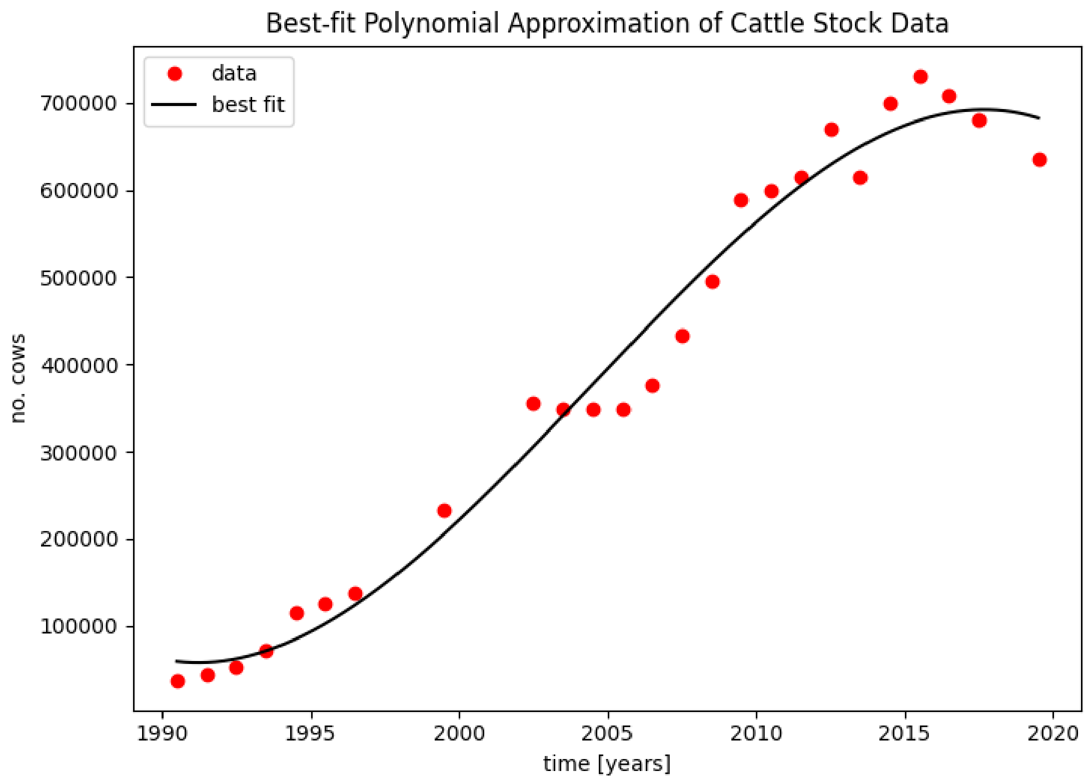

Figure 5: Cubic polynomial fitting to the cattle data for use in the ODE model.

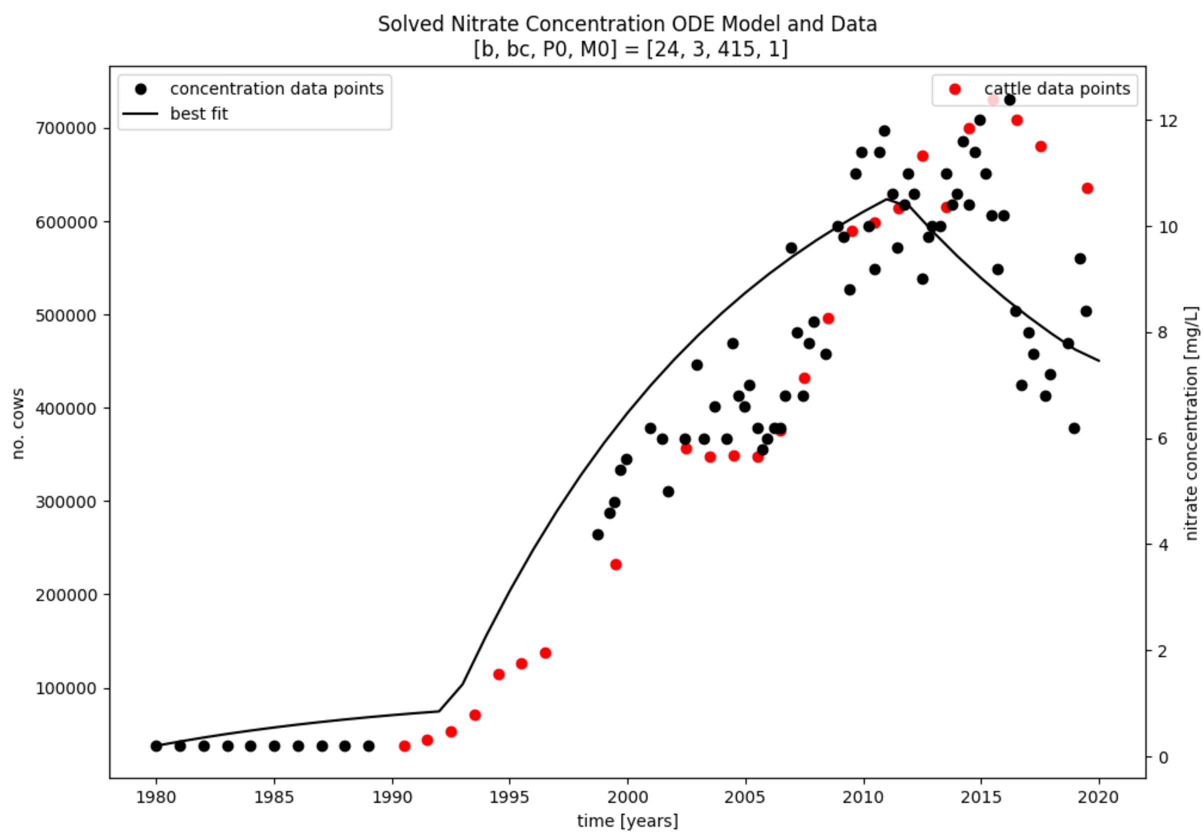

Figure 6: Initial best-fit model.

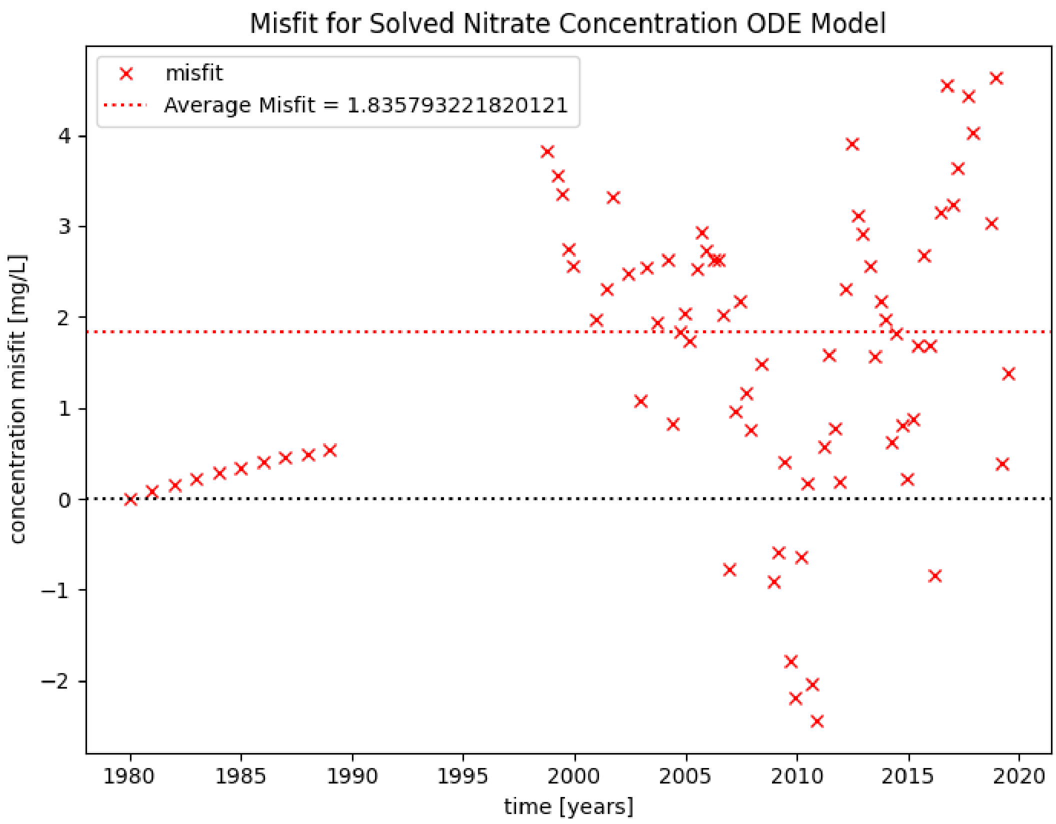

Figure 7: Misfit of initial best-fit model.

The model does a reasonable job capturing the increase in nitrate concentration from 1990 and the decline in the concentration from 2012 (after the implementation of the carbon sink in 2010), however, overestimates the impact of cattle numbers on the nitrate concentrations between 1990 and 2005 whilst failing to capture the extent of the peak nitrate concentrations measured from around 2010 to 2015. The misfit of the model against the data can be seen in Figure 7 above. The misfit is as large as 4.5 mg/L at times but the average misfit is 1.84 mg/L (as seen in Figure 7). The misfit becomes significant past the year 1998 and is mostly positive until around 2010, which highlight the way our model overestimates the concentration within this time frame. Past 2010, the misfit is both positive and negative, which is understandable given the erratic nitrate concentration data points past 2010, which make it difficult to fit an accurate model.

8 Improve?

Given the significance cattle numbers have on nitrate concentrations, it is clear that the polynomial fit does not accurately capture the varying cattle numbers between 1990 and 2020. Therefore, we decided to instead use cubic spline interpolation of the cattle stock data to improve our best-fit model. This interpolation can be seen in Figure 8 below.

On further inspection, it was also clear that we had a stiff ODE, meaning explicit numerical methods such as Improved Euler would result in unstable solutions that don't accurately represent the intricacies within the model. Therefore, we decided to implement the LSODA algorithm from the scipy package, which is an implicit numerical method designed specifically to tackle stiff ODEs.

Finally, we decided to implement a sum-of-squares method within a differential evolution algorithm in order to optimise the values of the unknown parameters, rather than manually calibrating the LPM model.

These changes can all be seen in **main.py**, and produces the improved best-fit LPM model seen in Figure 9 below. This time, the model has parameters [b, bc, P0, M0] = [24.02, 3.12, 414.17, 0.90], which are similar to the manually calibrated parameters in our initial best-fit model, revealing that it was not the manual calibration that resulted in the worse fit.

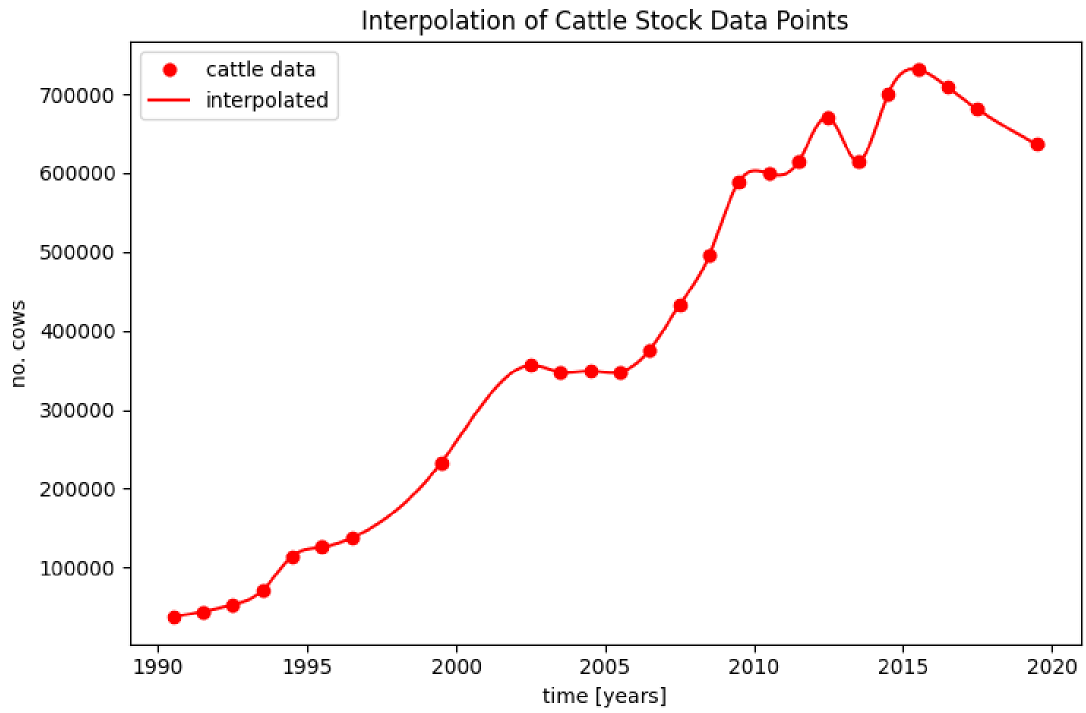

Figure 8: Cattle stock data interpolated using cubic splines in order to improve the model.

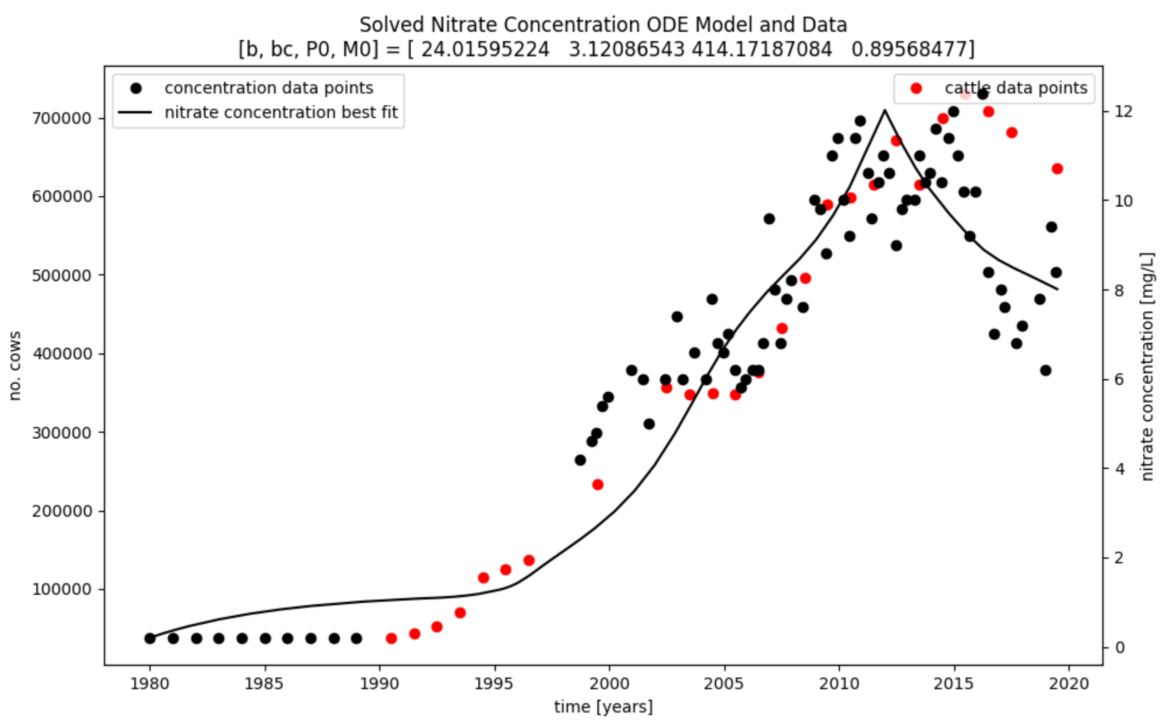

Figure 9: Improved best-fit model.

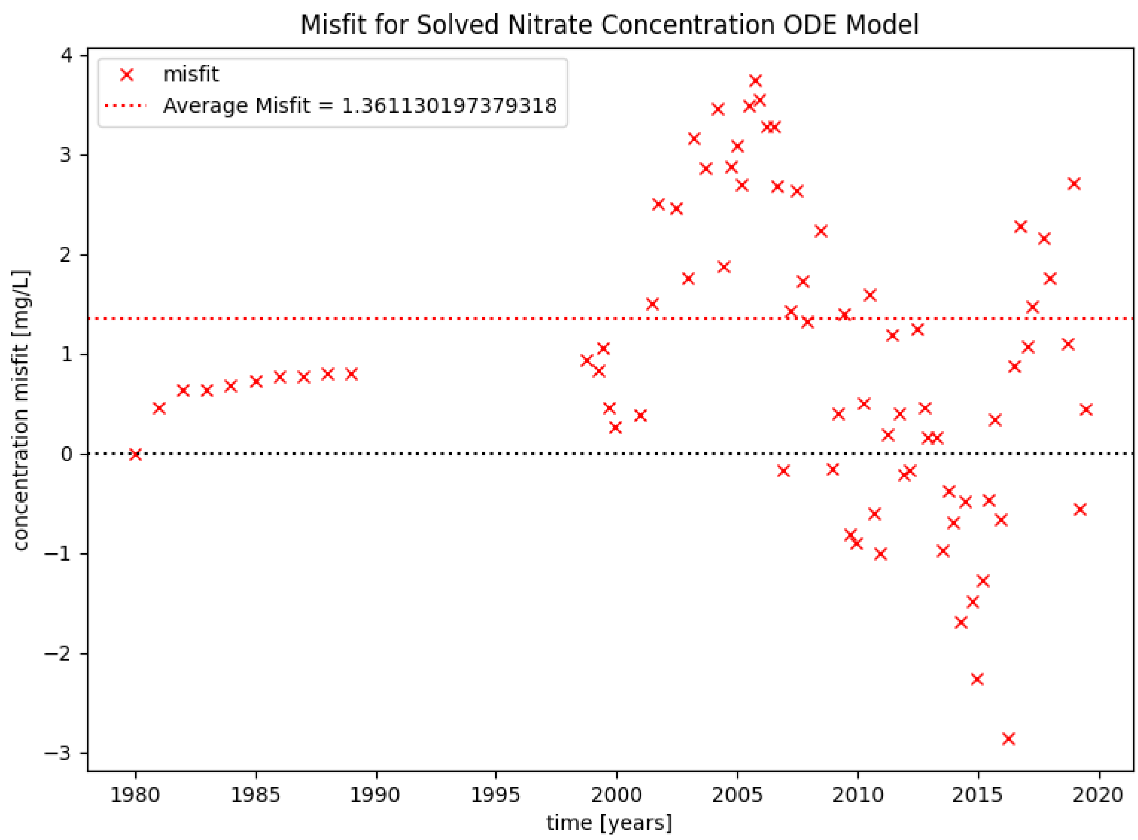

Figure 10: Misfit of improved best-fit model.

The improved model does a better job at capturing the increase in nitrate concentrations over time and the high peak in concentration in around 2012. The misfit for this model (seen in Figure 10 above) also reveals that the improved model is a better fit than our original model. While the misfit is still up to 4 mg/L due to the erratic nitrate data points, the average misfit has reduced from 1.84 mg/L to 1.36 mg/L, showing a marked improvement in the model fit.

9 Use?

We will use the improved best-fit model from Section 8 to consider five different MAR implementation schemes over the next 30 years. These are:

- $\Delta P_{MAR} = 0$ i.e. the increase in pressure at the high pressure boundary is 0 MPa.
- $\Delta P_{MAR} = 0.02$ i.e. the increase in pressure at the high pressure boundary is 0.02 MPa.
- $\Delta P_{MAR} = 0.05$ i.e. the increase in pressure at the high pressure boundary is 0.05 MPa.
- $\Delta P_{MAR} = 0.075$ i.e. the increase in pressure at the high pressure boundary is 0.075 MPa.
- $\Delta P_{MAR} = 0.1$ i.e. the increase in pressure at the high pressure boundary is 0.1 MPa.

It is also important to consider how cattle numbers will change in the next 30 years. Given that dairy farming is of great significance in Southland, we have decided to model the growth of cattle in Southland linearly as growth between 1990 and 2020 has also been roughly linear. This is also a conservative estimate that allows for the growth and sustainability of dairy farming in Southland. This future cattle stock model can be seen in Figure 11 below, whilst the future model results for the five scenarios are shown in Figure 12 below.

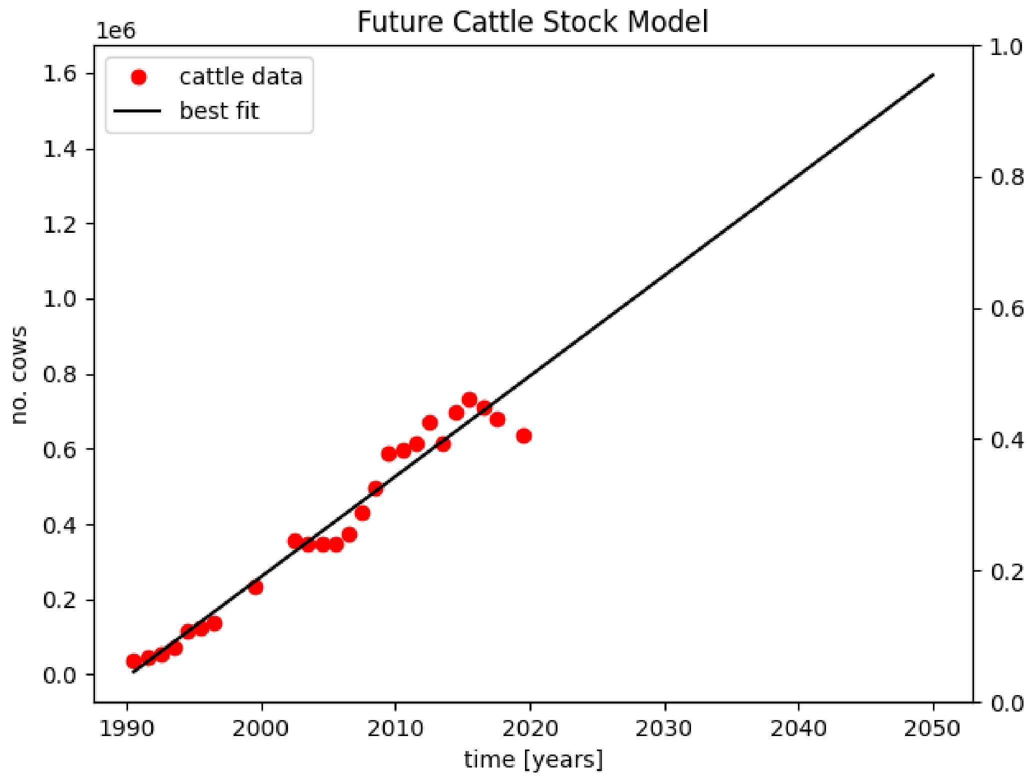

Figure 11: Linear model to approximate future cattle stock growth in Southland.

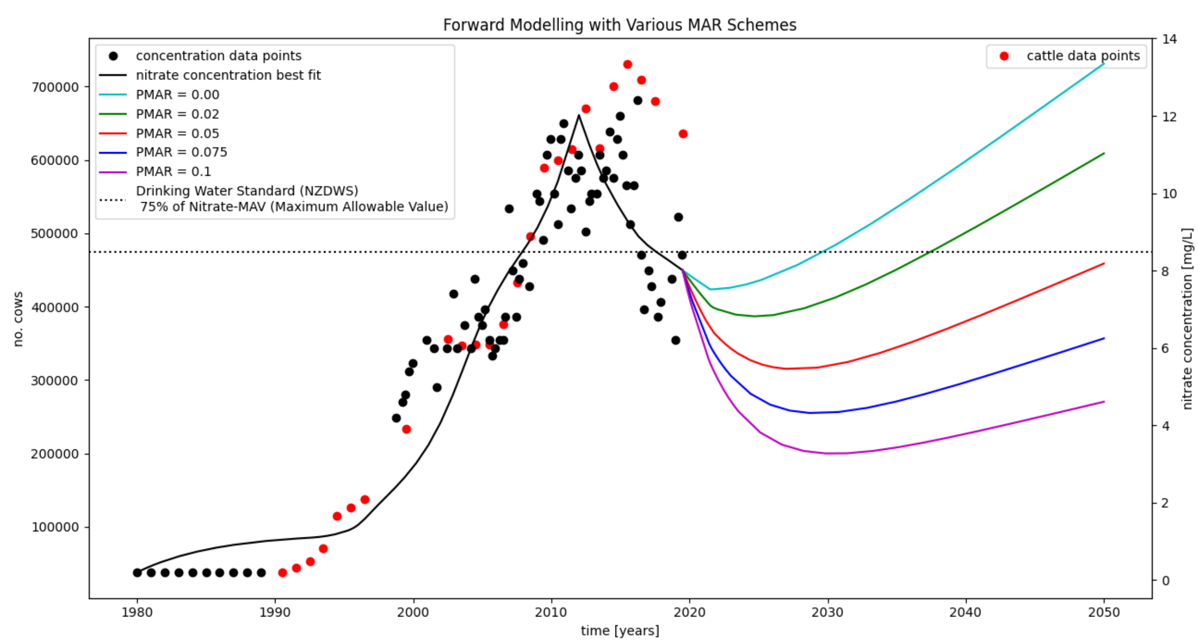

Figure 12: Model predictions for five MAR implementation scenarios: $\Delta P_{MAR} = 0$ MPa (cyan), $\Delta P_{MAR} = 0.02$ MPa (green), $\Delta P_{MAR} = 0.05$ MPa (red), $\Delta P_{MAR} = 0.075$ MPa (blue), $\Delta P_{MAR} = 0.1$ MPa (purple). Best-fit model (black), nitrate data (black dots) and cattle data (red dots) are also shown. The dashed line at 8.475 mg/L indicates the level at which nitrate concentrations are considered 'significantly elevated' (75% of the Nitrate-MAV).

The nitrate concentration after 30 years for each of the five scenarios is 13.4, 11.1, 8.2, 6.3 and 4.6 mg/L, respectively. Therefore, an MAR scheme where $\Delta P_{MAR} > 0.05$ is required for nitrate concentrations to remain below a 'significantly elevated' level over the next 30 years.

10 Unknown?

10.1 Parameter Uncertainty

There is inherent uncertainty in the nitrate concentration data, so we have resolved to explore the set of possible models that result in an acceptable calibration. The concentration data are reported to have a standard deviation of 0.625 mg/L which is then used in the sum-of-squares objective function

$$S(\theta) = \frac{1}{\sigma^2}\sum(\widetilde{p} - p(\theta))^2$$

where $p(\theta)$ is the lumped parameter model for parameters θ, where in this case $\theta = [b, b_c]$ (i.e. b are the parameters for which uncertainty will be modelled. We constructed a posterior distribution by performing a coarse grid search over the $[b, b_c]$ parameter space, centred on the improved best-fit model determined in Section 8, and computing the objective function at each point. We fit the resulting data using a multivariate Gaussian to obtain the best-fit parameters and covariance matrix. Using these as input to np.random.multivariate_normal, we obtain a set of 100 models that have been fitted to the data. These have been plotted in Figure 13 below.

For each of the five scenarios, the 90% confidence interval (i.e. [5%, 95%]) for nitrate concentration after 30 years is [7.44, 13.74], [6.69, 11.36], [5.31, 8.43], [4.17, 8.03] and [3.15, 8.01] mg/L respectively.

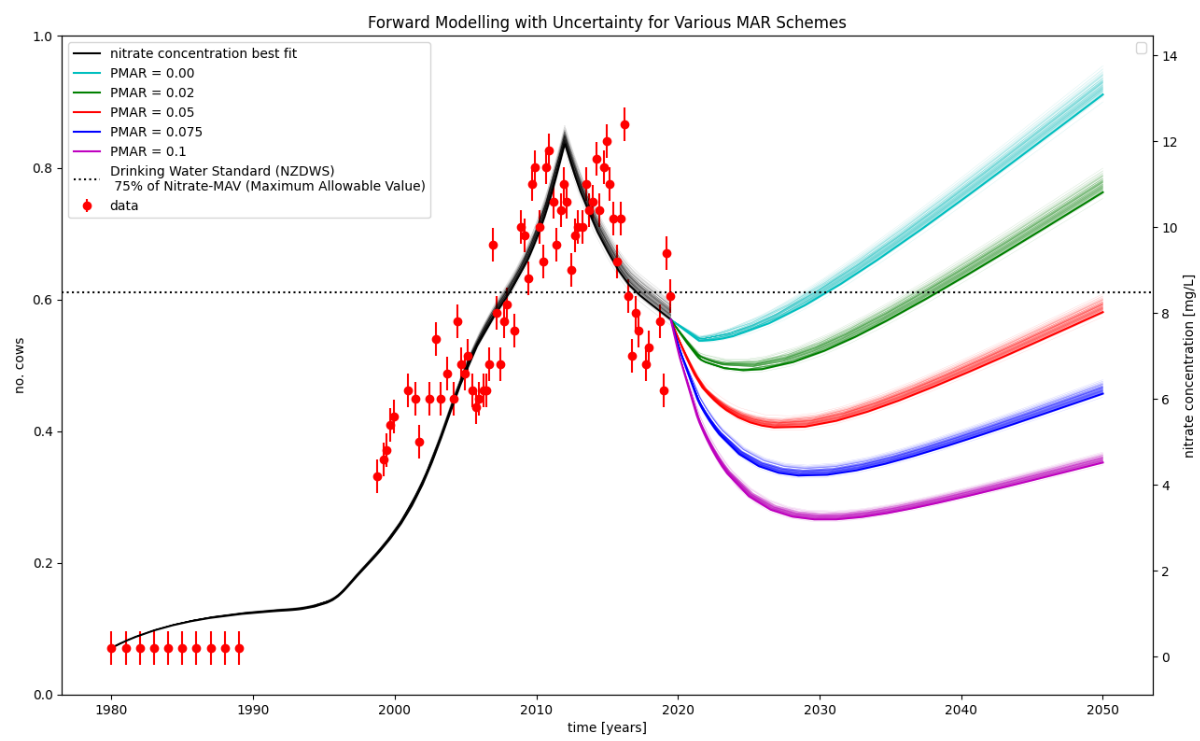

Figure 13: Model predictions for five MAR implementation scenarios: $\Delta P_{MAR} = 0$ (cyan), $\Delta P_{MAR} = 0.02$ (green), $\Delta P_{MAR} = 0.05$ (red), $\Delta P_{MAR} = 0.075$ (blue), $\Delta P_{MAR} = 0.1$ (purple). Historic models sampled from the posterior (black) overlay the nitrate data (red circles) whose error bars show 1 σ deviation. The dashed line at 8.475 mg/L indicates the level at which nitrate concentrations are considered 'significantly elevated' (75% of the Nitrate-MAV).

10.2 Structural Error

Our model has several structural errors that mean it does not fully or accurately depict the full scope of the aquifer. A better geothermal model, such as one that represents the reservoir by dividing it into a series of connected blocks and using partial differential equations (PDEs) to model the flows and pressure changes between them[4], would allow more localised spatial analysis of pressure changes and nitrate concentrations within the aquifer. Our current model cannot capture these localised differences, instead treating the whole aquifer as a single block.

Another shortcoming is the effect that the water injection of MAR has on the nitrate concentration within the groundwater of the aquifer. Whilst injecting water at the high pressure boundary of the aquifer increased the pressure, it also increase the volume of fluid being pushed through the aquifer, reducing the concentration of nitrate in the water further. Our model cannot capture the extent of this concentration reduction as a result of water volume increase within the aquifer as we have not taken the volume of fluid from the MAR schemes into account.

[4] O'Sullivan et al., 2009

11 Recommend?

Changes to pressure and nitrate concentrations within the Edendale Farm aquifer have been modelled over a thirty year period and subject to five different MAR scenarios. Three of these ($\Delta P_{MAR} = 0.05$, $\Delta P_{MAR} = 0.075$ and $\Delta P_{MAR} = 0.1$ MPa) maintained a nitrate concentrations below 8.475 mg/L over the thirty year period, the limit at which nitrate concentrations are considered 'significantly elevated'. If the intent is to avoid this limit (to prevent further implementation of groundwater monitoring and allow the water to be safe for human consumption), then it is recommended to implement an MAR scheme where $\Delta P_{MAR} > 0.05$ MPa. This will minimises the negative impacts that come from increased nitrate concentrations (such as unsafe human consumption and environmental damage) whilst still allowing for the growth of dairy farming in Southland (due to the conservative future cattle stock model assumed in Section 9, Figure 11), thus meeting the needs of all three main stakeholders: Environment Southland, Dairy NZ and Ngāi Tahu.

This recommendation is also a robust one given that at $\Delta P_{MAR} = 0.05$ MPa, there is only a 5% likelihood that the nitrate concentration within the aquifer will be above 8.43 mg/L within 30 years (which is still below the 8.475 mg/L limit).

Our model also shows that the implementation of the carbon sink initiative in 2010 was very effective in reducing the nitrate concentration in the groundwater. This can be seen in Figure 9, which shows a sharp decrease in the nitrate concentration from around 12 mg/L in 2012 to 8 mg/L in 2020. While one could speculate that this was due to other factors, the only significant change in the model was the addition of the carbon sink, while cattle stock numbers continued to increase slightly and other parameters remained constant. This shows that at the very least, introducing carbon sinks into other dairy farms across Southland could be a viable way to quickly reduce nitrate concentrations while other proposals are considered. The effect of carbon sinks in each farm will vary, however, due to a multitude of factors influencing their efficacy, thus they are not a sustainable and watertight solution to this problem.

The forecast above is specific to the scenario modelled, and small deviations from the scenario may invalidate the forecast, for instance, if the MAR scheme implemented is ever so slightly less than $\Delta P_{MAR} = 0.05$ MPa, there is no guarantee that the 8.475 mg/L threshold would not be breached. Furthermore, there are many considerations that have not been taken into account, such as localised spatial nitrate concentrations, stress and deformation physics and volume increase from injected water. It would be advisable to investigate the effect of these processes on the groundwater quality in order to make a more informed decision on the resource consent.