26th International Congress on Modelling and Simulation, Adelaide, SA, Australia, 30 Nov–4 Dec 2025
mssanz.org.au/modsim2025

Extended uncertainty analysis for the simple catchment load model SPARROW and implications for water management

AH Elliott[a] and AF Semadeni-Davies[a]

AH Elliott[a] and A Semadeni-Davies[a]
[a]Earth Sciences New Zealand. Email: sandy.elliott@niwa.co.nz

Abstract: Uncertainty analysis is increasingly recognised as essential in water quality modelling to inform planning processes. In New Zealand, the hybrid mechanistic–statistical SPARROW catchment model, has been widely applied through its incorporation into the CLUES model to estimate annual average contaminant loads. but previous applications have relied on restrictive assumptions of parameter normality and have not addressed prediction uncertainty or cross-validation. Here we present extensions to the SPARROW uncertainty framework by developing methods to capture parameter and prediction uncertainty under less restrictive assumptions. We applied conventional methods such as leave-one-out cross-validation, bootstrap resampling of observations to generate parameter ensembles, and three parameter uncertainty estimates, including a Gaussian copula approach that preserves empirical parameter distributions while representing correlation structures. Prediction uncertainty was further assessed by convolution of parametric uncertainty with a log-normal residual error distribution. Application to total nitrogen load estimation in three New Zealand regions demonstrated that bootstrap parameter distributions were approximately normal for well-fitting models, but skewed and sometimes multimodal for poorer fits, particularly where parameter bounds were active. For well-fitting models, prediction intervals from multinormal, raw bootstrap, and copula approaches were comparable, suggesting that the standard multinormal approximation is adequate. However, in poorly-fitting models, the multinormal assumption biased median predictions while bootstrap and copula approaches better captured skewed distributions. Across regions, prediction uncertainty was large: the median inter-decile ratio of predicted loads ranged from ~2 in the best-fitting region to over 12 in the worst, with residual error dominating in some cases and parameter uncertainty in others. These results highlight both the feasibility of extended uncertainty analysis in SPARROW and the substantial prediction uncertainty inherent in load prediction. The findings emphasise the importance of recognising structural, data, and model input uncertainties alongside quantified prediction uncertainty. Implications for decision making are profound, as precautionary approaches based on conservative uncertainty bounds can impose costs of a magnitude that challenges technical and economic feasibility. Focussing on relative rather than absolute changes in load provide a way to circumvent these difficulties, but uncertainty with such relative approaches remains to be assessed.

Keywords: SPARROW, uncertainty, CLUES, catchment model

1. INTRODUCTION

Analysis of uncertainty is increasingly being required for water quality models used in planning processes in New Zealand (Ministry for the Environment 2023) and internationally. One simple model applied in New Zealand is the United States Geological Survey (USGS) SPARROW model (Alexander and Gorman Sanisaca 2019; Schwarz et al. 2006), this is a hybrid mechanistic-statistical catchment model used for contaminant loads. A version of SPARROW has been incorporated into the GIS-based Catchment Land Use for Environmental Sustainability (CLUES) catchment model that has been used widely in New Zealand (Elliott et al. 2016). Although model parameter uncertainty and model performance statistics have been routinely reported for CLUES (e.g., Semadeni-Davies et al. 2019), along with discussion of broader aspects of uncertainty, these analyses have not been taken forward into prediction uncertainty or cross-validation. The original SPARROW code does perform some uncertainty based on normal approximations to parameter distributions using Monte Carlo analysis, but they do not investigate parameter normality assumptions. Therefore, we have conducted analysis of CLUES uncertainty using code similar to that for SPARROW but extending the uncertainty analysis to allow for non-normal parameter distributions. We also modified the method for modifying the prediction distribution for residual errors using a convolution approach. We provide results for an example application to total nitrogen (TN) loads in three regions in New Zealand, examining parameter and prediction uncertainty for these examples. We also discuss the implications of this uncertainty, and other forms of uncertainty, for decision making.

2. METHODS

The SPARROW model represents a catchment with a dendritic drainage network, with mean annual contaminant source loads in each subcatchment and first-order decay in streams and lakes. The source and decay terms can be specified in a flexible way, but in New Zealand we have used an approach within CLUES based on source yields for various land uses, adjusted by landscape variables such as slope, rainfall and drainage. Various model parameters are determined by constrained non-linear least-squares optimisation (NLSE) to measured contaminant yields (load per unit upstream area) at monitoring stations, in log-transformed space, which has been found previously to conform to typical assumptions such as error normality.

In the standard SPARROW code, there is no cross-validation of model fit. Parameter uncertainty is estimated from first-order uncertainty analysis associated with the NLSE, interpreted as a multi-normal distribution. Prediction uncertainty is determined either by combining the residual variance with parameter-related variance from first-order analysis, or by Monte Carlo resampling of the multi-normal parameter distribution and normal residuals distribution.

We investigated extensions to the uncertainty methods. This entailed re-writing the SPARROW core calculations within CLUES in Python (using the least_squares function in SciPy for model fitting), then creating code for: a) leave one out cross-validation (LOO), which repeatedly and systematically fits the model to the observations, each time removing one observation of mean annual load at a monitoring site, and evaluating model fit from the collated set of residuals and predictions for the left-out observations; b) bootstrap resampling of observations (200 bootstrap samples with replacement) to obtain a set of parameter vectors (a parameter ensemble); c) parametric uncertainty prediction by sampling from the parameter ensemble d) and convolution of the parametric prediction distribution with a log-normal residual error distribution derived from the residual root mean square error (RMSE). The parameter sampling in c) was done with three alternatives: i) random sampling of the multinormal parameter distribution as in the standard SPARROW approach (‘multinormal’); ii) applying each of the parameter vectors from the original bootstrap parameter estimation directly (‘raw’ method) or iii) sampling from a Gaussian copula(‘copula’). For the random parameter sampling, 500 samples were used.

The copula method (iii above) takes the empirical distribution of parameters from the parameter vector, maps it into a Gaussian distribution based on cumulative probability values, samples from the multinormal distribution derived from the mapped points and then remaps back to the original parameter space. This has the advantage of keeping key features of the original marginal distribution while representing parameter correlation structure.

The residual convolution method (d above) generates a set of log-normal error values (500) based on uniform quantiles, applying these to each of the predictions from the parametric resampling to derive a larger set of predictions, extracting quantiles for each prediction location. This still assumes independent and homoscedastic residual errors but allows for non-normality of the predictions from the parametric resampling and avoids random variability that would be associated with Monte Carlo resampling.

All calculations were performed in Python, using standard libraries such as numpy and Scipy rather than specialist packages for statistical aspects (for example, numpy randn for multivariate normal sampling and numpy.linalg.cholesky for Cholesky decomposition involved in the Gaussian copula method).

These methods were applied for TN calibration for three different regions in New Zealand, representing a range of model goodness-of-fit. Each region represented about one fifth of the country, with typically 100 to 200 load observations per region. The regions are denoted as Regions 1 to 3, because calibrations are preliminary. For more details on the SPARROW modelling approach as applied in New Zealand, see Elliott et al. (2005) and the CLUES references above.

3. RESULTS

3.1. LOO Cross-validation

Cross validation fit statistics were worse than those obtained from fitting to all the available data (Table 1). For Region 1, where there was an exceptionally good fit, the difference in fit metrics was minimal. However, in Region 3, where there was a worse fit (but still with 'satisfactory' NSE according to Moriasi et al. 2015) there was a moderate difference in fit statistics between cross-validation and all data, although the fit was broadly comparable. An RMSE of 0.762 (from Region 3 with no cross-validation) implies a factor error of 2.14 (from exponentiation of the RMSE which is natural log space), whereas an RMSE of 0.842 (from cross-validation) implies a factor error of 2.32; the difference is not of great practical significance.

Table 1. Summary of model fit metrics. All metrics are based on natural logs of TN yield. NSE is Nash-Sutcliff model efficiency. R^2 values were very close to NSE, so are not shown.

	NSE		RMSE	
Region	No cross validation	With Cross-validation	No cross validation	With Cross-validation
Region 1	0.893	0.874	0.239	0.260
Region 2	0.712	0.633	0.342	0.386
Region 3	0.601	0.514	0.762	0.842

3.2. Parameter distribution

The bootstrap parameter distribution for Region 1 was approximately normal (see the distribution for an example parameter in Figure 1, and the variance of the parameter estimates from the bootstrap parameter distribution were close to the values of parameter variances from the multinormal distribution from the least squares analysis with all data (results not show). Parameter normality was not evaluated formally, but implications of the normality assumption for predictions are considered in the next section. For Region 3, however there were larger variances and a more skewed distribution, which departs from the normality assumption of first-order uncertainty analysis. For some other parameters in Region 3, the departure was more pronounced; for example, the stream decay parameter was multimodal and influenced by the imposed parameter bounds. In general, parameter bounds, when active, induced non-normality.

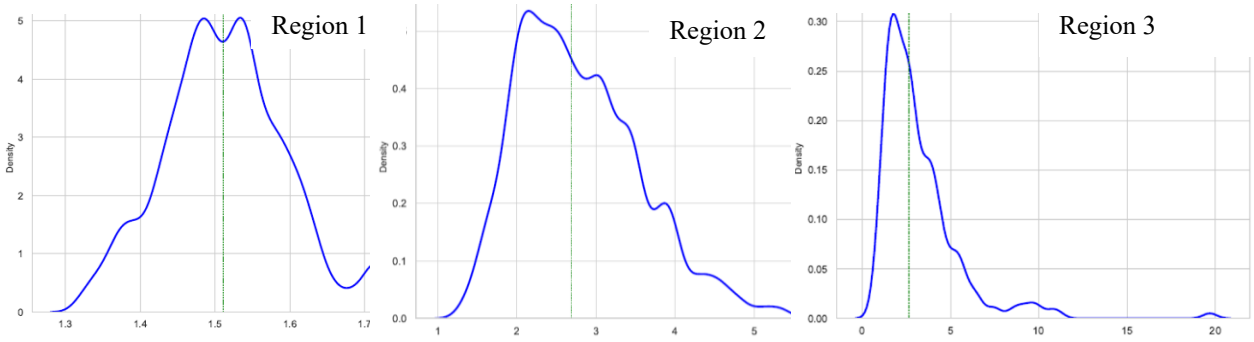

Figure 1. Example empirical parameter distributions for one of the land use yield coefficients for three example regions. The curves are from kernel density estimation of the probability density function and the green line is the median value. Values on the horizontal axis are natural log of TN yield (t/km^2). Note that the parameters are different in the two different regions, reflecting differences in model form. The meaning of the parameters is slightly different between runs, so the primary point to illustrate is the difference in shape and spread between regions.

3.3. Prediction uncertainty

For Region 1, the multi-normal sampling, prediction with the raw bootstrap parameter set, and Gaussian copula resampling gave similar prediction intervals (see Table 2 for an example prediction location). For Region 3, the raw and copula method gave similar results to each other. However, the multinormal method gave different predictions. For example, the median prediction was 2.08 for the multinormal and the Copula median value was 3.53. The multinormal median value for Region 3 was close to the value predicted from the central parameter

estimates from the model fit with all data (2.11). This shows that for models with large errors and skewed distributions, applying the multinormal distribution method could biased the median predicted value significantly.

The effect of adding residual error is shown in the last two columns of Table 3. Adding residuals did not change the median prediction greatly. However, it did affect the predictions for other quantiles in some cases. Even for the Region 1, the interdecile ratio (90th decile divided by 10% decile) was changed from 1.13 without residuals to 1.87 with residuals, showing that residual error dominated the overall prediction error in that case. For Region 3, parameter error dominated overall (interdecile ratio of 16.8 without residuals (parametric error), and 28.1 with residuals), but residual error also contributed. For the Region 2 example, there was approximately an even contribution from residual and parameter uncertainty. For Region 3, the residual errors may have interacted with departures in non-lognormality in the parametric errors, because the median with residuals is different from the median without residuals (whereas for log-normal distributions, the median should not be affected by multiplying by the log-normal error distribution with a log-mean of zero).

Table 2. Prediction intervals for different bootstrap prediction methods, for a selected reach in each region (different sites for each region). The Multi-normal is results from parametric resampling from the normal parameter distribution. 'Raw' is predictions using the parameter ensemble from bootstrap resampling of observations. Units are TN load (t/km^2)

Quantile	Muti-normal with residuals	Raw with residuals	Copula with residuals	Copula without residuals
		Region 1		
10%	0.550	0.555	0.554	0.714
50%	0.755	0.759	0.757	0.757
90%	1.036	1.039	1.035	0.809
Interdecile ratio	1.88	1.87	1.87	1.13
		Region 2		
10%	0.342	0.325	0.325	0.419
50%	0.551	0.543	0.541	0.543
90%	0.884	0.900	0.901	0.716
Interdecile ratio	2.58	2.80	2.73	1.71
		Region 3		
10%	0.24	0.74	0.78	1.13
50%	2.08	3.18	3.53	3.25
90%	24.4	19.1	21.9	19.0
Interdecile ratio	101.8	25.8	28.1	16.8

The prediction uncertainty analysis shows that there are large uncertainty bounds. This is shown for the interdecile ratio in Table 2 for example locations, and in Figure 3 for the distribution of the ratio across all locations in the relevant region. The median interdecile ratio was 1.87 for Region 1, with not much variation between sites. For Region 2, the median ratio was 2.84. For Region 3 the median ratio was 12.5, with some very large values. The implications of this uncertainty are discussed in the next section.

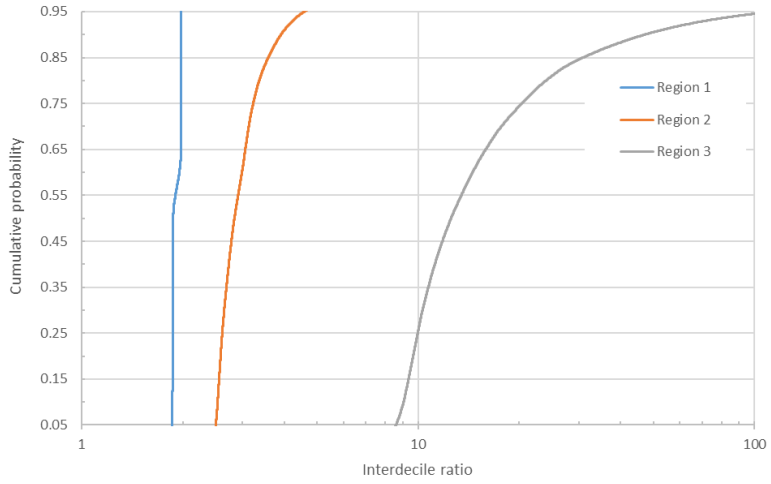

Figure 3. Distribution of interdecile ratio or predicted TN load for the three regions. Only values from 5% to 95% are shown, for compactness of the plot.

4. DISCUSSION

4.1. Comparison of methods and computation considerations

The LOO Cross-validation method involves repeating the least-squares fit for each observed load (typically 100 times) but did not provide significant information beyond fitting to the full dataset even for Region 3, which had the worst fit. A potential benefit is that this method can identify the influence of each observation on the parameters, complementing conventional statistics such as Cooks distance and residual or leverage analysis. Ideally, cross-validation metrics would be used for selecting model variants (for example, lumping different land uses), but that would be computationally cumbersome if many variants are investigated.

Parametric bootstrapping involves repeated model fitting, typically 200 times, introducing computational burden. For example, typically nonlinear optimisation would entail about 20 iterations (epochs) with seven function evaluations per iteration, resulting in about 28000 function evaluations. Since the model is simple, this was not a large burden and was completed in 1 to 8 hours on a PC. Bootstrapping is embarrassingly parallel (runs can be separate without interacting) so could be sped up using high performance computing. Generating multinormal distribution parameters involves only one model fitting, so can be used for initial model exploration, such as outlier identification, initial parameter sensitivity, and investigating different model forms.

The examples presented here also demonstrated that if the model fit is good (as in Regions 1 and 2), the multinormal approach would provide a good approximation of the parameter distribution and prediction uncertainty, so that bootstrapping parameters is not necessary. It is difficult to know whether the fit is good enough for the linear multinormal approach to provide a good approximation. One potential sign is whether the parameter constraints are active or close to the central parameter value (say, within on standard deviation of the parameter marginal variance), in which case the parameter distribution from resampling would likely be skewed. A model efficiency >0.7 and RMSE <0.5, along with separation from parameter constraints, could be used as a tentative rule of thumb for when the multinormal distribution could be used.

Bayesian methods, not presented here, such as the Metropolis-Hastings MCMC (Monte Carlo Markov Chain) method, also entail a large number of function evaluations, and would likely be comparable to the bootstrap parameter method in terms of computation effort. Non-linear least squares could be used to generate a prior distribution for the Bayesian method. Such methods have been applied previously to SPARROW (Wellen et al. 2014) but have not made their way into the standard SPARROW model distributed by the USGS or seen widespread use, perhaps due to their more sophisticated conceptual basis. Genetic algorithms also entail a large number of function evaluations, but have disadvantages in terms of prediction uncertainty, the final population of parameters being dependent on selection criteria. The GLUE method, which is often used for expensive water quality simulation models, could entail fewer function evaluations (especially for well-behaved models such as SPARROW), but there are difficulties with prediction uncertainty due to arbitrary acceptance criteria for 'behavioural' models.

The raw and copula methods for prediction uncertainty gave similar results. Both methods require carrying the ensemble parameter set into the prediction stage (either as the parameter distribution or copula empirical functions). Copulas have a slight advantage in that they would allow for more predictive samples than the raw method, which could capture the parameter distribution and its implications for uncertainty more thoroughly. Although used widely in applied statistics, copulas do not seem to have widespread operational application in water quality catchment modelling, despite their ease of implementation and elegance. The multinormal method does not require as much information to be brought into the prediction stage (just the parameter variance matrix). It still requires repeated predictions in a Monte Carlo sampling approach so is not more computationally efficient than the other methods. It also has the disadvantage of having to introduce methods for dealing with infeasible parameter instances (for example, negative contaminant yields) and, for models with high parameter uncertainty, can give biased estimates of predictions and their uncertainty, as shown for Region 3.

The convolution method used for residual adjustments residual method has advantages over the Monte Carlo method used in the standard SPARROW software, because it does not rely on random sampling of a parametric distribution. It does not require additional model evaluations beyond that needed for parameter uncertainty but does involve some numerical expense in calculating the convolution numerically, but with less than doubling the overall computation time for our examples. The method could potentially be extended for empirical and heteroscedastic residual distributions.

4.2. Other sources of uncertainty

The assessment of cross-validation, parameter, residual, and prediction uncertainty only gives a partial view of the overall uncertainty. There are considerable areas of structural uncertainty in SPARROW, such as assumptions about equations to represent the influence of landscape delivery factors and stream decay and spatial discretisation. As one example, we found a regionally calibrated version of the CLUES *E.* coli model applied to the Taranaki

Region, gave unreasonably high load estimates for the central part of the region which is dominated by a mountain (Mount Taranaki) (Semadeni-Davies and Elliott 2024). We found that the high load predictions were because the calibrated rainfall parameter (which has an exponential form) in the model did not account for higher orographic rainfalls on the mountain's slopes. Subsequent water quality data from that area, which were not available for our calibration showed near pristine water quality. To counter this issue, we restricted the use of the rainfall factor to mean annual rainfalls of less than 3 m/y. This example also highlights difficulties with biased input data sampling. Load estimation from monthly spot water quality measurements entails high imprecision and differences between estimation methods (Snelder et al. 2017). While some of this measurement uncertainty is included in the unexplained error term, there could be systematic biases and structural uncertainty in load estimation that is not captured by the residual errors. Additional errors will occur due to model input uncertainty, such as imprecise or outdated land use, point source, soil characteristics, and rainfall, which are not accounted for in our analysis. Thus, while provision of numeric uncertainty values gives the impression of precise uncertainty estimation, that distracts from a more honest presentation other sources of uncertainty.

4.3. Importance of uncertainty and mitigation approaches

The predicted load estimates from SPARROW entail considerable uncertainty, with inter-decile ratio of about 2 for the best-fitting models and 12.5 for the worst model. Such prediction uncertainties are not exclusive to our SPARROW modelling approach. For example, using linear quantile regression with a simple export coefficient model, there was approximately a 10-fold difference between 10% and 90% quantile model predictions (Snelder et al. 2024).

One way to deal with uncertainty is to take a precautionary approach, allowing for a tolerance for uncertainty to avoid undesirable outcomes (such as breaching a load target). For example, for a 90% confidence of not exceeding a load limit for Region 1, there might be a requirement that the predicted load is about half the target value. This degree of conservativism would have massive implications for land management. For example, a 50% reduction in nitrogen loads was estimated to reduce farm income by 22% to 28% at a cost of NZD5-9 million per year for a 200 km^2 pasture-dominated catchment (Beukes et al. 2023). For the Southland region, a 30% reduction would cost about NZD 250 million per year loss in gross margin per year (Kaye-Blake et al. 2013). The implications for models with higher uncertainty would be massive, to the point of implicating reductions that are not technically feasible.

Averaging results over a large area will reduce uncertainty, assuming that errors are independent. For example, a regional or national estimate of load will have less uncertainty than for a single site. This averaging effect will depend on the scale of interest (e.g, loading to an estuary is likely to have larger uncertainty that loading to the coast nationally). Often, though, assessments of water quality and required reductions are based on individual representative sites.

Other empirical models of contaminant loads in New Zealand have similarly found typically 10-fold errors in TN assessed through N-fold cross-validation (Snelder et al. 2023).

One typical way to ameliorate this uncertainty is to focus on relative changes in load for future scenarios (such as changes in land use or land management) rather than absolute changes, with the assumption that the bias at a site will be consistent over time (errors are fully correlated for the current and future conditions). Typically a concentration target is chosen, the current concentration is estimated from monitoring or a separate statistical model, and the required relative load reduction is assumed to be the same as the relative concentration reduction, expressed as proportion reduction (Elliott et al. 2020). Analysis of uncertainty analysis of this relative reduction approach has not been conducted. A further aspect is that estimates of current concentration are uncertain, and concentration targets based on ecological thresholds are uncertain.

5. CONCLUSIONS

Overall, this study provided some method refinements compared with the standard SPARROW model code which could be used easily in future modelling with this class of model. It pointed out where the more computationally-intensive LOO and bootstrapping methods may be warranted and identified challenges in relation to the practical implications of prediction uncertainty.

The difference in RMSE between fitting the full dataset and LOO cross validation was small for well-fitting models is of little practical importance for the worst-fitting example. Considering the additional computation expense for little gain, LOO could probably be avoided, expecially if the model fit is good.

For models with a good fit, bootstrap parameter estimation from observation resampling is not necessary, and sampling from the multinormal parameter distribution would be suitable in the prediction stage to estimate prediction uncertainty, as in the standard SPARROW software. In contrast, the poorest-fitting model showed

significant departures from normality, suggesting that parameter bootstrap ensembles should be generated in the fitting stage and applied in the prediction stage.

The convolution method for incorporating residual uncertainty proved to be practical and is an improvement on the random resampling from the standard SPARROW model.

The example applications demonstrate that residual and parameter uncertainty both play a role in prediction uncertainty, with different relative contributions depending on the model, so that both sources of uncertainty should be considered (not just relying on RMSE, for example).

The inter-decile prediction ranges were considerable - about a factor of 2 for the best-fitting model, and about 10 for the worst fitting although still 'satisfactory' model. This provides a challenge in terms of practical applications, where only tens of percent difference in source loading can have large implications. While conservative estimates of required load reductions may be warranted to safeguard water quality to counter uncertainty, these may be too costly, or even practically infeasible, to implement. A strategy to mitigation this uncertainty is to work with relative rather than absolute loads, but the uncertainty of such an approach has not been assessed.

REFERENCES

Alexander RB and Gorman Sanisaca L (2019) RSPARROW: An R system for SPARROW modeling (Software Release). US Geological Survey. https://doi. org/10.5066/P9UAZ6FO.

Beukes PC, Weeber M, Neal M, Thiange C and Burger DF (2023) Evaluating nutrient abatement policies for a pastorally farmed catchment-lagoon ecosystem: A New Zealand case study. Agriculture, ecosystems & environment 342, 108227.

Elliott A, Alexander R, Schwarz G, Shankar U, Sukias J and McBride G (2005) Estimation of nutrient sources and transport for New Zealand using the hybrid mechanistic-statistical model SPARROW. Journal of Hydrology (New Zealand) 44(1), 1-27.

Elliott AH, Semadeni-Davies AF, Shankar U, Zeldis JR, Wheeler DM, Plew DR, Rys GJ and Harris SR (2016) A national-scale GIS-based system for modelling impacts of land use on water quality. Environmental Modelling & Software 86, 131-144. Doi: http://dx.doi.org/10.1016/j.envsoft.2016.09.011.

Elliott AH, Snelder TH, Muirhead RW, Monaghan RM, Whitehead AL, Bermeo-Alvear SA and Howarth CJ (2020) A heuristic method for determining changes of source loads to comply with water quality limits in catchments. Environmental Management 65(2), 272-285.

Kaye-Blake B, Schilling C, Monaghan R, Vibart R, Dennis S and Post E (2013) Potential impacts of water-related policies in Southland: on the agricultural economy and nutrient discharges: NZIER report to the Ministry for the Environment, 29 May 2013. Wellington: NZIER, 2013.

Ministry for the Environment (2023) Developing, adapting, and applying environmental models in a regulatory context in New Zealand. Ministry for the Environment, Wellington, New Zealand, https://environment.govt.nz/publications/developing-adapting-and-applying-environmental-models-in-a-regulatory-context-in-new-zealand/.

Moriasi DN, Gitau MW, Pai N and Daggupati P (2015) Hydrologic and water quality models: Performance measures and evaluation criteria. Transactions of the Asabe 58(6), 1763-1785.

Schwarz GE, Hoos AB, Alexander RB and Smith RA (2006) The SPARROW Surface Water-Quality Model – Theory, Application, and User Documentation. U.S. Geological Survey Techniques and Methods, Book 6, Section B. Chapter 3. USGS, Reston, Virginia.

Semadeni-Davies A and Elliott S (2024) Addendum to Modelling *E. coli* to support implementation of the NPS-FM: modification of CLUES *E. coli* model for Te Maunga. Taranaki Regional Council.

Semadeni-Davies AF, Jones-Todd CM, Srinivasan MS, Muirhead RW, Elliott AH, Shankar U and Tanner CC (2019) CLUES model calibration: residual analysis to investigate potential sources of model error. New Zealand Journal of Agricultural Research. Doi: 10.1080/00288233.2019.1697708.

Snelder T, McDowell R and Fraser C (2017) Estimation of catchment nutrient loads in New Zealand using monthly water quality monitoring data. JAWRA Journal of the American Water Resources Association 53(1), 158-178.

Snelder T, Cox T, Fraser C, Elliott S and Kerr T (2023) Quantifying Catchment Nutrient Modelling Parameters. An analysis using the available New Zealand data. Land Water People, Christchurch, New Zealand.

Snelder T, Elliott S, Muirhead R and Fraser C (2024) Parameters for simple empirical catchment water quality models for simulating *Escherichia coli* in New Zealand rivers. Report to Our Land and Water National Science Challenge. Land Water People, Lyttelton, New Zealand.

Wellen C, Arhonditsis GB, Labencki T and Boyd D (2014) Application of the SPARROW model in watersheds with limited information: a Bayesian assessment of the model uncertainty and the value of additional monitoring. Hydrological Processes 28(3), 1260-1283.

