National scale mapping of rivers showing a threshold Dissolved Inorganic Nitrogen concentration of 0.8 mg/L

Prepared for Irrigation New Zealand
Mr Andrew Curtis

May 2014

Authors/Contributors:
Martin Unwin

For any information regarding this report please contact:

Clive Howard-Williams
Chief Scientist

+64-3-343 7857
clive.howard-williams@niwa.co.nz

National Institute of Water & Atmospheric Research Ltd
10 Kyle Street
Riccarton
Christchurch 8011
PO Box 8602, Riccarton
Christchurch 8440
New Zealand

Phone +64-3-348 8987
Fax +64-3-348 5548

Report date:	**22 May 2014**
NIWA Project:	**SCJ14501 Task INZ**

Contents

Figures

Reviewed by

....
Clive Howard-Williams

Approved for release by

....
Clive Howard-Williams

Summary

Irrigation New Zealand requested a single mapping exercise of Dissolved Inorganic Nitrogen (DIN, where DIN = NO_3-N + NH_4-N) across New Zealand showing river reaches where median DIN concentrations exceeded 0.8 mg/L. To meet this request, I used a water quality data set previously compiled for the Ministry for the Environment to determine median DIN at 622 sites throughout New Zealand, and hence to fit a predictive model relating median DIN to 28 catchment and landcover descriptors known to be strongly associated with variability in water quality over national and regional scales.

The resulting model explained 70.7% of the observed variance in median DIN, and showed a strong tendency for DIN to be highest in catchments dominated by heavy pastoral landcover. When extrapolated to unmonitored catchments, the model suggests that the 0.8 mg/L threshold is exceeded in approximately 10% of New Zealand waterways.

This result is subject to two caveats. First, the period of record for most sites was Jan 2006 - Dec 2010, so the data do not necessarily accurately reflect the current state. Second, the model fitting process I used is conservative, tending to under-predict at the higher end of the observed range. If so, the model may underestimate the extent to which the 0.8 mg/L threshold is exceeded.

1 Introduction

NIWA was requested by Mr Andrew Curtis of Irrigation New Zealand (INZ; email, 15 May 2014) to generate a map, together with a brief supplementary report, showing river reaches across New Zealand where median Dissolved Inorganic Nitrogen (DIN = NO_3-N + NH_4-N)) exceeds 0.8 mg/L. To create an appropriate data set, INZ required a predictive model of DIN concentrations for all rivers, developed using the same methodology favoured by the Ministry for the Environment (MfE) for presenting national scale information on water quality variables used for the National Objectives Framework (e.g., Figure 1 in MfE's *Freshwater Reform: 2013 and beyond* [1]).

This report describes the available data sets on DIN in New Zealand rivers and the methods used to predict DIN at national scale, and discusses the findings. Only minimal detail is given; for further background I refer readers to Unwin *et al*. (2010) and references therein.

2 Methods

Source data for the model were extracted from a database created by merging water quality data from 17 sources, comprising 16 councils and unitary authorities, and the National River Water Quality Network (NRWQN) maintained by NIWA[2]. The database was compiled in 2010 and has been intermittently updated since then, but the period of record varies among regions depending on the currency of their most recent available data set. For the purposes of this report I limited the data to records collected between 1 January 2006 and 31 December 2012, with at least 16 records over at least five consecutive years.

The resulting data set included 622 sites for which DIN data had either been reported directly, or could be synthesised as the sum of Nitrate Nitrogen (NO_3-N) and Ammonium Nitrogen (NH_4-N). Each site was characterised based on median DIN over the available period of record. Data for each site were recorded monthly, bimonthly, or quarterly; requiring at least 16 records per site ensured that medians for sites with only quarterly data were robust enough to be used for modelling.

I used Random Forest (RF) regression modelling (Breiman 2001) to model log-transformed median DIN as a function of 28 predictor variables representing catchment average values (at a ~700 m spatial scale; seeSnelder & Biggs 2002) of environmental variables such as rainfall, temperature, slope, geological characteristics, and land cover (Wild *et al.* 2005, Woods *et al.* 2006). These variables were the same as those used for the MfE model, and are strongly associated with spatial variability in water quality over national and regional scales (Unwin *et al.* 2010). I used importance scores and partial dependence plots (Breiman, 2001) to characterise model performance. Importance scores indicate how much the predictive performance of the model decreases if a specific predictor is not used, and partial dependence plots show the marginal effect of a variable on the response after accounting for the average effects of the other variables in the model. These plots do not perfectly represent the effects of each variable, but provide useful information for interpretation.

[1] http://www.mfe.govt.nz/publications/water/freshwater-reform-2013/
[2] Available on request from MfE

3 Results

Sites with usable DIN data were sparse in the north and west of the South Island, but were abundant and well distributed across the rest of New Zealand (Figure 3-1). This distribution compares very favourably with models for other water quality variables in previous studies (Unwin *et al.* 2010).

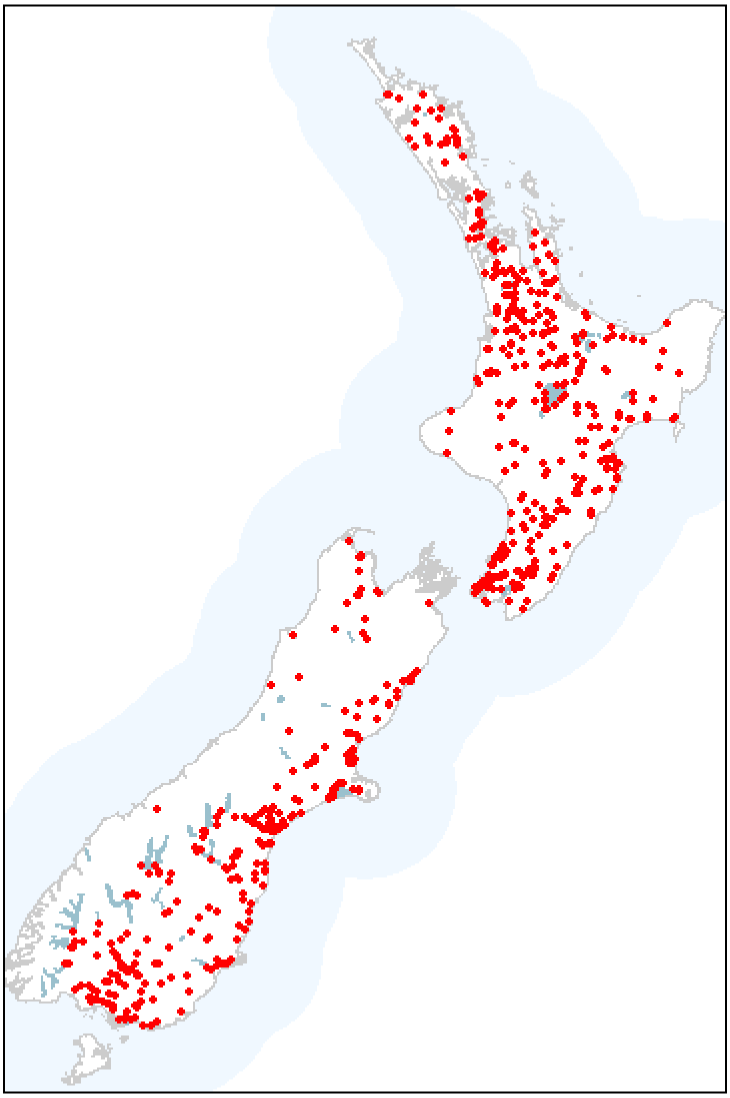

Figure 3-1: Unitary authority and NRWQN water quality monitoring sites used in this study to model median DIN.

The fitted RF model explained 70.7% of the observed variance in median DIN (i.e., r^2 = 0.707). This result compares well with models previously estimated for other water quality variables (Unwin *et al.* 2010), and suggests that the present model is robust.

Predicted DIN exceeded 0.8 mg/L at 58,999 of 574,502 of modelled sub-catchments (Figure 3-2). This suggests that the 0.8 mg/L threshold is exceeded in approximately 10% of New Zealand waterways.

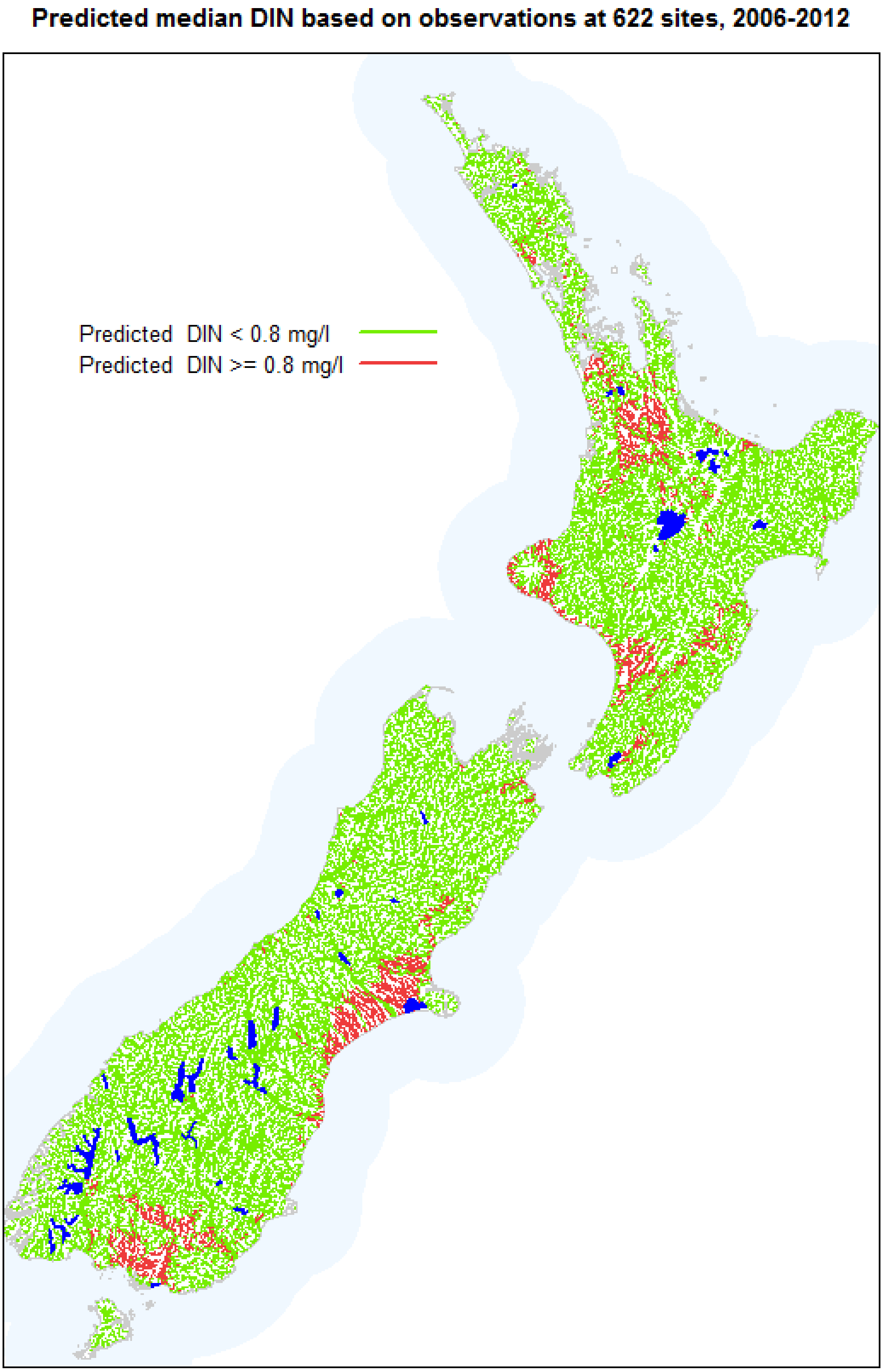

Figure 3-2: Predicted DIN in NZ rivers relative to a threshold concentration of 0.8 mg/L. Catchments smaller than ~10 km² are suppressed to minimise over-plotting at small spatial scales.

Partial plots for the fitted model confirm that elevated DIN levels are strongly associated with the percentage of heavy pastoral landcover in the upstream catchment (Figure 3-3). Again, this result is consistent with RF predictions for the individual components of DIN, NO_3-N and NH_4-N, both of which tend to increase in catchments with a high proportion of pastoral landcover (Unwin *et al.* 2010). In particular, the result for DIN is very similar to predictions for NO_3-N, which is the dominant component of DIN at most sites.

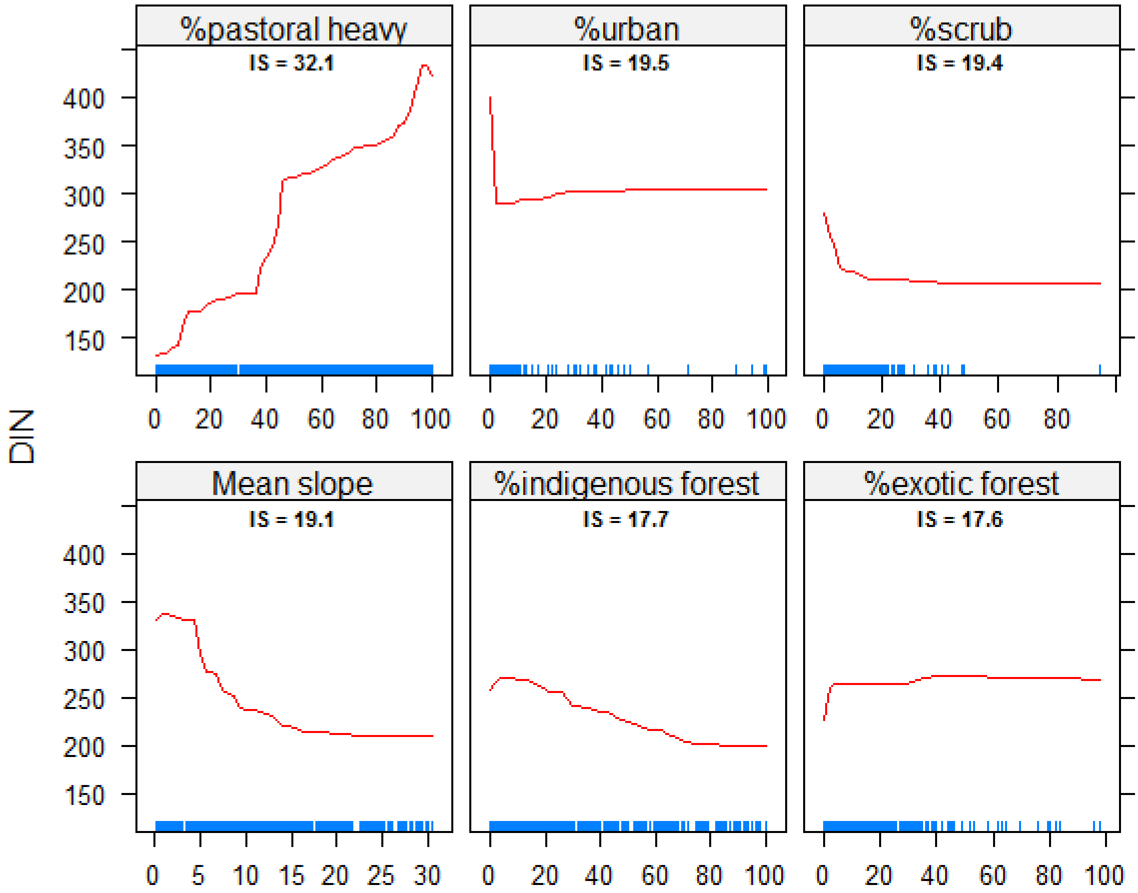

Figure 3-3: Partial plots for the leading six predictors of median DIN in a RF model based on 622 sites. Successive panels show the importance score for each predictor (IS), and the partial dependence of DIN on each predictor (red line). Mean slope is in units of m/km; all other predictors are expressed as percentages .

The results presented here are subject to two caveats. First, as noted earlier (Section 2), the underlying database is slowly becoming outdated. Data for a few sites are current to c. December 2012, but for most sites the period of record is Jan 2006 - Dec 2010. Second, RF models are conservative in that they tend to over-predict at the lower end of the observed range, and under-predict at the higher end. This tendency is relatively weak for the current model, thanks to the high percentage of explained variance, but is readily apparent in a plot of observed vs. predicted medians for each site (Fig. 3.4). In particular, predictions for sites with high observed DIN tend to lie below the diagonal line, consistent with a tendency to under-predict. This suggests that the model may underestimate the extent to which the 0.8 mg/L threshold is exceeded.

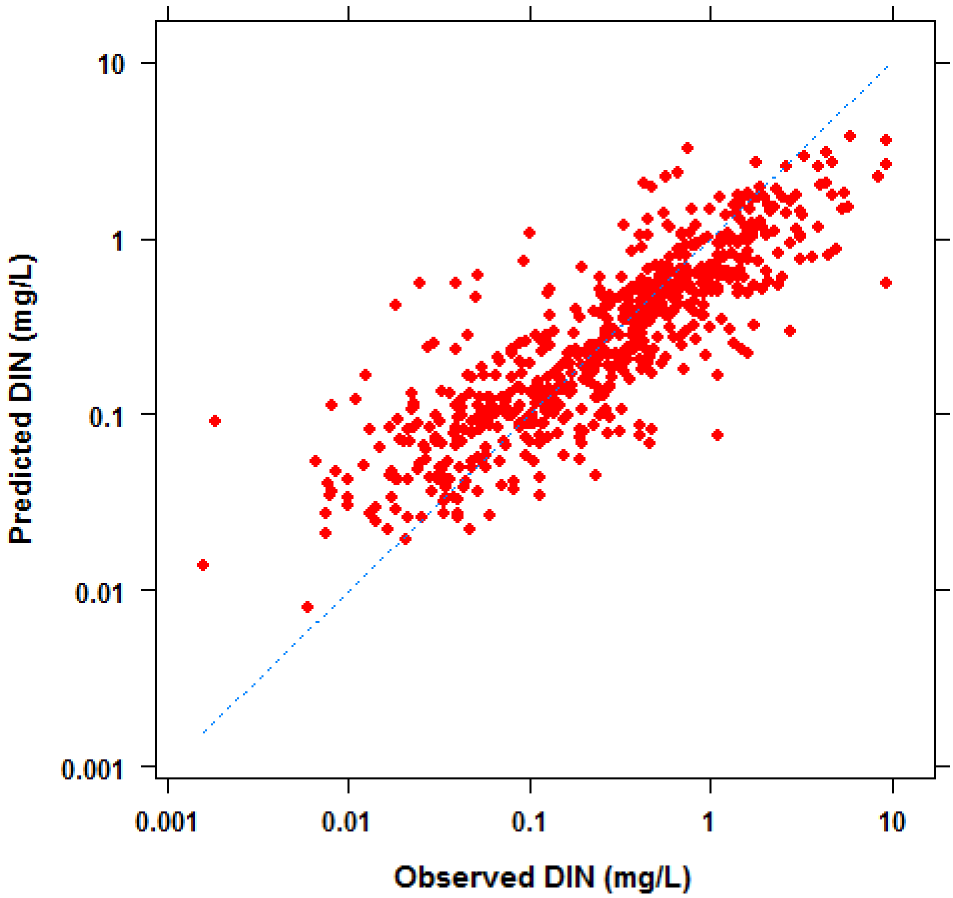

Figure 3-4: Observed vs. predicted DIN (mg/L) at 622 sites used to fit a RF model. The diagonal line represents equality between observations and predictions.

4 References

Breiman, L. (2001). Random Forests. *Machine Learning 45*: 5-32.

Snelder, T.H.; Biggs, B.J.F. (2002). Multi-scale river environment classification for water resources management. *Journal of the American Water Resources Association 38*: 1225-1240.

Unwin, M.J.; Snelder, T.; Booker, D.; Ballantine, D.; Lessard, J. (2010). Predicting water quality in New Zealand rivers from catchment-scale physical, hydrological and land cover descriptors using random forest models. *NIWA Client Report CHC2010-037.* 21 p.

Wild, M.; Snelder, T.; Leathwick, J.; Shankar, U.; Hurren, H. (2005). Environmental variables for the freshwater environments of New Zealand River Classification. *NIWA Client Report CHC2004-086.* 31 p.

Woods, R.; Hendrikx, J.; Henderson, R.; Tait, A. (2006). Estimating mean flow of New Zealand rivers. *Journal of Hydrology (New Zealand) 45*: 95-110.