Some Current Issues In Using Diffuse Large Datasets For Environmental Modelling In New Zealand

Graham McBride,[1] Ude Shankar,[2] Rich Alexander,[3] Sandy Elliott[1]
[1]*NIWA (National Institute of Water and Atmospheric Research Ltd.), Hamilton, NZ*
[2]*NIWA (National Institute of Water and Atmospheric Research Ltd.), Christchurch, NZ*
[3]*USGS (United States Geological Survey), Reston, Va, USA*

Key words: Spatially-referenced, water quality, nutrients, regression, scale, GIS

Abstract: Issues arising from the application of the SPARROW regional regression modelling technique to fluxes on nutrients in New Zealand streams are discussed. These include model scale dependency, as reflected in changing the DEM scale, and the lack of compatibility and transportability of the associated software. Nevertheless, this spatial referencing technique is demonstrating considerable advantages, because it marries the considerable spatially-referenced datasets with measurements routinely made in waterbodies. This facilitates the explanation of the relative importance of contaminant sources and may even be useful as a predictive tool (e.g., in showing the consequences of changing the land-use).

1. INTRODUCTION

Water quality impacts in lowland streams are receiving considerable attention in environmental research and management in New Zealand. These impacts are manifested as:

- nutrient enrichment and the consequent promotion of beds of introduced aquatic weeds (e.g., oxygen weeds such as *Egeria densa*, thought to have been introduced from aquaria releases—also an issue in lakes);
- consequential night-time oxygen depletion, pH variations, and possible ammonia toxicity;

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: 10.1007/978-0-387-35503-0_29

- pathogen load (New Zealand has an extensive open-pasture grazing system);
- loss of water clarity (New Zealand streams have a typically short travel time from source to sea, and there is a public expectation of clear water).

There is now extensive information on land uses, climate, geology, point sources, water level and flow, water quality and topology. These are scattered through a number of large datasets held by a variety of organisations (research institutes, regional and district councils, Government Departments, industry). The challenge now before us is to use these datasets to explain the relative contributions of all possible sources contributing to the impacts and to do so in a way that captures the dendritic drainage pattern of the landscape.

2. REGIONAL REGRESSION MODEL

In collaboration with the United States Geological Survey, we are applying the SPARROW technique to these data (spatially referenced regressions on watershed attributes—Smith et al., 1997; Preston and Brakebill, 1999). In particular we are attempting to explain the loads of nutrients we find at water quality monitoring sites. This has involved a trial application to 38 sites in the Waikato Region ($\approx$ 14,500 km^2 in the north-central North Island). It is now being extended to all 77 sites in the National River Water Quality Network (whose catchments cover about half the New Zealand land surface area of approximately 250,000 km^2).

SPARROW uses a multiplicative model for nutrient flux (F_i) at a number of reaches (i, where i_{max} = 38 for the Waikato application). The model is:

$$F_i = \sum_{n=1}^{N} \sum_{j \in J(i)} \left[\beta_n S_{n,j} \left(e^{-\alpha' Z_j} \right) \left(e^{-k' H_{i,j}} \right) \right] \varepsilon \qquad (1)$$

where: $J(i)$ is the sum of all monitored and unmonitored reaches upstream of reach i (Fig. 1), $S_{n,j}$ is a measure of nutrient mass flux from source n applied to drainage reach j, β_n is a source-specific coefficient reflecting the amount of nutrient available for transport, α' is a vector of delivery coefficients associated with the reach's land-surface characteristics Z_j, k' is a first-order loss coefficient (indexed to stream flow and the areal water load of impoundments), $H_{i,j}$ refers to hydraulic properties of rivers, impoundments and lakes and ε is an i.i.d. error term. Equation (1) is solved using non-linear

regression estimation (SAS, 1993). A GIS digital terrain and reach network provides the spatial framework for the model.

3. WAIKATO CATCHMENT APPLICATION

The digital network was based on raw spot height data defined over a grid of squares with edge length of 500 m. By "burning on" the stream channels from a 1:250,000 scale map series and using an interpolation method (Hutchinson, 1989) that creates hydrologically correct digital elevation models we were able to obtain a grid of squares with edge length of 100 m (Fig. 2). A network climbing algorithm was used to accumulate reach lengths (Smith *et al.*, 1997, White *et al.*, 1992).

Water quality data were obtained from Environment Waikato (the region's resource management agency), consisting of monthly measurements of total nitrogen and phosphorus at 38 sites (representing watersheds of 5 to 8,000 km^2) and continuous flow measurements. Mean flux for 1993-97 was obtained by conventional flux estimation techniques (Cohn *et al.*, 1989; Ferguson 1986).

Explanatory variables evaluated in the exploratory models were:

sources: point source loads; pasture land; pine forest land; scrub lands; soil erosion index; cow populations; apatite minerals;

watershed attributes: rainfall; runoff; slope; soil drainage index; stream density;

channel attenuation: channel length; reservoir channel length; lake & reservoir discharge, volume and surface area.

Successful model calibration was obtained, as indicated on Table 1, enabling some preliminary predictions to be made, as indicated on Fig. 3. The calibrations compare very favourably with those obtained in USA applications, all with more than 38 sites.

4. ISSUES IN FUTURE APPLICATIONS

Three issues are apparent:

First, in applying SPARROW modelling to the national datasets we have the opportunity to refine the DEM. That is, we now have data for grid squares of edge length 30 m, on which we will burn channels from a 1:50,000 scale map series, and for which we will interpolate elevations from 20 m contours. This gives rise to a possibly substantial scale issue. For example, introduction of a finer stream network could alter the importance of stream attenuation.

Second, the software for this modelling consists of a mixture of GIS, DEM and special Fortran routines. Any development of standardised software for these tasks is some considerable way off.

Third, point data is often interpolated to produce a continuous surface of a variable in question. The choice and robustness of a method will depend largely on data density, heterogeneity of the variable and the grid element size for homogeneity (although this is often limited by the hardware or software capability of the modelling system). For example, only a small number of precipitation stations may be required to produce a good rainfall surface for a flat terrain, whereas a much greater number of stations would be required in meteorologically heterogeneous mountainous regions. Also scale differences between different data elements used in model parameterisation must be minimised in order to ensure that modelled processes reflect the scale of the measured data.

5. CONCLUSIONS

There are considerable advantages is using software developments to construct models to marry spatially-referenced data with measurements of fluxes of material in water bodies. It is now beginning to explain the relative importance of sources of nutrients. In the future we hope to extend that explanatory power to water clarity and river pathogen load (for which monitoring is underway). In doing so we have to take careful account of the scale dependency of some features of the model and its calibration process.

6. ACKNOWLEDGEMENTS

7. REFERENCES

Cohn, T.A.; DeLong, L.L.; Gilroy, E.J.; Hirsch, R.M.; Wells, D.K. 1989: Estimating constituent loads. *Water Resources Research 25*: 937–942.

Ferguson, R.I. 1986: River loads underestimated by rating curves. *Water Resources Research 22*: 74–76.

Hutchinson, M.F. 1989: A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. *Journal of Hydrology* 106:211-232.

Preston, S.D.; Brakebill, J.W. 1999: Application of spatially referenced regression modeling for the evaluation of total nitrogen loading in the Chesapeake Bay watershed. U.S. Geological Survey Water-Resources Investigations Report 99-4054, 12 p.

SAS. 1993: SAS/ETS User's Guide, Version 6, 2nd Edition. SAS Inst. Cary, NC, USA: 509–591.

Smith, R.A.; Schwarz, G.E.; Alexander, R.B. 1997: Regional interpretation of water-quality monitoring data. *Water Resources Research 33*: 2781–2798.

White, D.A.; Smith, R.A.; Price, C.V.; Alexander, R.B.; Robinson, K.W. 1992: A spatial model to aggregate point-source and nonpoint-source water-quality data for large areas, *Computational Geosciences 18*: 1055–1073.

Table 1. Calibration of SPARROW model to Waikato TN data

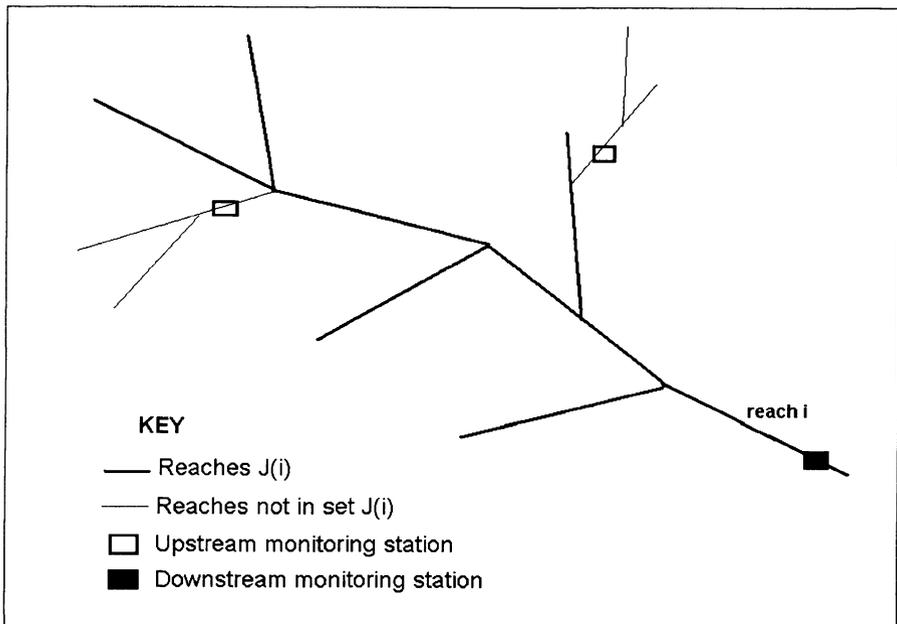| Model Parameter	Mean Estimate	Standard Error	*P*-value	Units
Sources				
Point	1.108	0.885	0.220	dimensionless
Pasture land	67.66	37.83	0.084	kg ha^{-1} yr^{-1}
Other land	7.86	5.57	0.168	kg ha^{-1} yr^{-1}
Soil drainage	0.214	0.098	0.037	index unit^{-1}
Aquatic Loss				
Stream (<1 m^3 s^{-1})	0.1997	0.0923	0.038	km^{-1}
Stream (>1 m^3 s^{-1})	0.0001	0.0059	0.985	km^{-1}
Lakes/reservoirs	0.0034	0.0023	0.150	m day^{-1}

Figure 1. Reach schematic

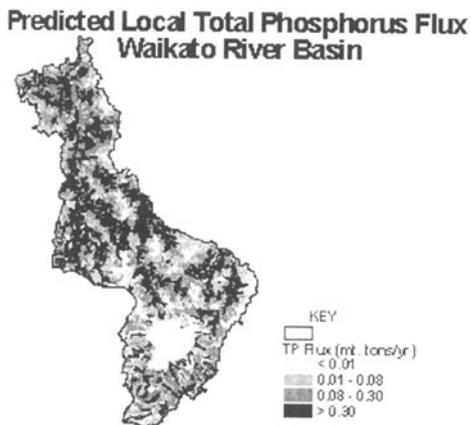*Figure 2.* The Waikato catchment, total area ≈ 14,500 km^2

Figure 3. Predictions of TP flux from interior watersheds as a function of land use and aquatic (stream, reservoir) attenuation. ["Local" refers to phosphorus generated from sources in the immediate watershed, exclusive of upstream sources]