Modelling Water Quality In Data Poor Catchments: A Combined Modelling Approach

R.A.Letcher[1], W.S. Merritt[1], A.J. Jakeman[1], B.Baginska[2]
Centre for Resource and Environmental Studies (CRES) the Australian National University,
Canberra ACT 0200, [2] NSW EPA, Sydney.

This paper develops an approach for modelling water quality which provides useful information in situations where there is little available data, as is the case in many Australian catchments. Different types of erosion and sediment/nutrient transport models that are currently available are discussed, with a view to identifying the most appropriate modelling approach for data poor situations. It is suggested that a combined empirical and conceptual approach would provide useful information on the data poor catchment, while utilising the limited available data.

1. INTRODUCTION

Water quality issues have become increasingly important to catchment stakeholders, such as management groups, land owners and government departments, in recent years. Information on the sources of pollutants in catchments and on the response of water quality to changing land use practices is required by such stakeholders. However this information is limited in many catchments. Whilst streamflow data may be available for several sites in a large catchment (and often on a short time step, of the order of minutes, for decadal periods), measurements of sediment and nutrient concentrations may be taken at only a relatively small proportion of sites within the catchment (often with time steps of a month or more). Erosion and sediment/nutrient transport models are increasingly being called upon to provide valuable information on water quality in such data poor situations.

This paper reviews the techniques currently available for modelling erosion and sediment/nutrient transport, focusing on the applicability of available models in data poor circumstances. It is suggested that a combined empirical and conceptual approach would be most appropriate for utilising the limited information available in data poor catchments.

2. AVAILABLE MODELS

A wide range of models exist for use in sediment transport and water quality modelling. These models differ in terms of complexity, the pollutants and processes considered, and the data required for model use. In general there is no 'best' model for all applications, rather the most appropriate model will depend on the intended use, data availability and the characteristics of the catchment being considered.

In general, models fall into three main categories, depending on the physical processes simulated by the model and the data dependence of the model:

- empirical;
- conceptual; and
- physics based.

All three model types have inherent limitations and advantages in their application. Classification of models as empirical, conceptual or physics based is subjective. Most models do not fit neatly into these categories, rather they are likely to contain a mix of modules from each of these categories. Models may also be described as hybrids between two or more of these classes.

2.1 Empirical Models

Empirical models are generally the simplest of all three model types. Empirical models are based primarily on the analysis of observations and seek to characterise response from these data (Wheater *et al.* 1993). The computational and data requirements for such models are usually less than for conceptual and physics based models, often being capable of being supported by coarse measurements. Jakeman *et al.* (1999) state that 'the feature of this class of models is their high level of spatial and temporal aggregation and their incorporation of a small number of causal variables'. Many empirical models are based on the analysis of catchment data using stochastic techniques, and as such are ideal tools for the analysis of data within catchments (Wheater *et al.* 1993). They are particularly useful as a first step in identifying sources of sediment and nutrient generation.

Most empirical models do not attempt to represent the physical processes involved in sediment generation. For this reason empirical models tend to be regarded as more catchment specific than the other two types. Consequently the ability of empirical models to predict the effects of changes in catchment characteristics, such as land use, on water quality and sediment yields can be limited. Empirical models also tend not to be event responsive, ignoring the processes of rainfall-runoff in the catchment being modelled.

Empirical models are often criticised for employing unrealistic assumptions about the physics of the catchment system, ignoring the heterogeneity of catchment inputs and characteristics, such as rainfall and soil types, as well as ignoring the inherent nonlinearities in the catchment system (Wheater *et al.* 1993). Such models are generally based on the assumption of stationarity; that is, it is assumed that underlying conditions remain unchanged for the duration of the study period. This assumption limits the potential for such models to be applied for predicting the effects of catchment change.

One example of an empirical erosion model is the landscape-factor approach of Moss *et al.* (1993). This approach was used to derive nitrogen, phosphorus and suspended sediment exports from Queensland coastal catchments according to:

$S(t) = L\ E\ D\ R$

$N(t) = L\ E\ C\ P\ D\ F\ R$

R = storm discharge (ML) / catchment area (km^2)

where
$S(t)$ is suspended sediment
$N(t)$ is nutrient export
L is area of a specified land use
E is erosion rate for a specified land use
D is delivery ratio for a specified land use
R is runoff correction factor for a specified land use
C is soil nitrogen or phosphorus content
P is enrichment ratio (phosphorus only)
F is dissolved nitrogen or phosphorus compensation factor.

Land use is broken down into pristine lands, grazing and cropping.

2.2 Conceptual Models

Conceptual models are typically based on the representation of the catchment as a series of internal storages. They usually incorporate the underlying physical mechanisms of sediment and runoff generation within their structure, representing flow paths within the catchment as a series of storages, each requiring some characterisation of its dynamic behaviour. Conceptual models tend to lump representative processes over the scale at which outputs are simulated (Wheater *et al.* 1993). Parameter values for conceptual models have typically been obtained through calibration against observed data such as stream discharge and concentration measurements (Abbott *et al.* 1986).

Conceptual models tend to include a general description of catchment processes, without including the specific details of process interactions, that would require detailed catchment information (Sorooshian 1991). This allows these models to provide an indication of the qualitative and quantitative effects of land use changes, without requiring large amounts of spatially and temporally distributed input data.

Due to the requirement that parameter values are determined through calibration against observed data, conceptual models tend to suffer from problems associated with the identifiability of their parameter values (Jakeman and Hornberger 1993). Most calibration techniques used for conceptual models of medium complexity (say more than six parameters) are capable of finding only local optima at best. This means that there are many possible 'best' parameter sets available. Spear (1995) identified this problem in large simulation models stating that 'there is not a single point in the parameter space associated with good simulations, indeed there generally is not even a well-defined region in the sense of a compact region interior to the prior parameter space.' In general simpler conceptual models have fewer problems with model identification than more complex models. Thus problems with model identification can be minimised through limiting the number of parameters to be estimated through calibration and possibly identifying additional parameters using *a priori* knowledge of the system (Kleissen *et al.* 1990; Wheater *et al.* 1993). This reduction in problems associated with identifiability through simplification of models may come at the expense of goodness of fit to calibration data. More complex models are more likely to provide a better fit to calibration data, although this does not necessarily extend to providing better predictions of future behaviour, as complex models run the risk of overfitting calibration data (Wheater *et al.* 1993).

The lack of uniqueness in parameter values for conceptual models means that the parameters in such models have limited physical interpretability (Wheater *et al.* 1993). However, this problem can also be associated with empirical and physics based models. Physics based models in particular are often over-parameterised whereas empirical models tend to be naturally much simpler in their level of parameterisation (Beven 1989; Wheater *et al.* 1993).

An example of a simple conceptual model is the IHACRES rainfall-runoff model. The model, using only one storage (as for catchments without significant baseflow) is given by:

$$Q_k = aQ_{k-1} + bU_k$$

where Q_k is modelled streamflow U_k is effective rainfall, which can be calculated in its simplest manner as a product of raw rainfall and a catchment moisture index.

2.3 Physics Based Models

Physics based models are based on the solution of fundamental physical equations describing streamflow and sediment/ nutrient generation within the catchment. Standard equations used in such models are the equations of conservation of mass and momentum for flow and the equation of conservation of mass for sediment (eg. Bennett 1974).

In theory, the parameters used in physics based models are measurable within the catchment and so are 'known'. However, in practice the large number of parameters involved and the heterogeneity of important characteristics within the catchment means that these parameters must often be calibrated against observed data (Beck *et al.* 1995; Wheater *et al.* 1993). This creates additional uncertainty in parameter values. Also, even in situations where parameters can be 'measured' within the catchment, errors in the measurement of important characteristics, and differences between model grid scales and measurement scales will create additional uncertainty as to the veracity of model outcomes (Bloschl and Sivapalan 1995). Another important consideration when using physics based models is that 'small scale parameters used for small scale models may lose physical significance at larger scales' (Seyfried and Wilcox 1995). Where parameters cannot be measured within the catchment they must be determined through calibration against observed data. Given the large number (possibly hundreds) of parameter values needed to be estimated using such a process, problems with the lack of identifiability of model parameters and non-uniqueness of 'best fit' solutions can be expected (Beck 1987; Wheater *et al.* 1993).

The derivation of mathematical expressions describing individual processes in physics based models is subject to numerous assumptions that may not be relevant in many real world situations (Dunin 1975). In general, the equations governing the processes in physics based models are derived at the small scale and under very specific physical conditions (Beven 1989). However, in physics based models these equations are regularly used at much greater scales, and under different physical conditions. Generally, the equations are derived for use with continuous spatial and temporal data, however the data used in practice is often point source data taken to represent an entire grid cell within the catchment (Beven 1989). The viability of lumping up small scale physics to the scale of the spatial grid used in many physics based models is questionable (Beven 1989). Lane *et al.* (1995) state that 'model parameters derived in this manner represent nothing more than fitted coefficients distorted beyond any physical significance'. Specifically there is a lack of theoretical justification for assuming that equations apply equally well at the grid scale, at which they are representing the lumped aggregate of heterogeneous subgrid processes (Beven 1989). These distortions would conceivably be exacerbated with the use of hillslope scale models for modelling entire catchments. Many physics based models have been developed on a field or hillslope scale. Extending these models to an entire catchment by summing across a spatially distributed grid will create problems with error accumulation.

An example of a physics based model for sediment and nutrient transport is the CREAMS model. Sediment transport in the CREAMS model is calculated according to the steady-state continuity equation

$$\frac{dG}{dx} = D_f + D_s$$

where G is the sediment load, x is distance, D_f is the detachment or deposition rate by flow and D_s is the rate that sediment is added to the flow from lateral areas. These parameters are calculated on a grid scale using equations representing the key physical processes.

3. CONSIDERATIONS IN MODEL APPLICATION

Jakeman *et al.* (1999) noted that the difficulties in environmental modelling can be characterised as problems of natural complexity, spatial heterogeneity and the lack of available data. The complexity of natural systems is due to differences in dimensions, temporal and spatial scales, and thresholds of water flow and sediment and nutrient transport through and within the media. Natural systems, from plot to catchment scale, tend to show a great deal of variation. Grayson and Moore (1993) noted that the scale at which uniformity is assumed in hydrologic models is generally greater than the scale at which directly measurable parameters are measured in the field, although smaller than shown by the outflow hydrographs. Thus, model predictions are subject to errors as a result of the inconsistency of scale between measured parameters and the way they are used in the model. This problem is particularly evident in data intensive models.

The model complexity is determined by the detail of the catchment processes simulated. Not only do the number of equations requiring solution increase in a model representing a large number of detailed processes, but so do the number of input parameters (Bennett 1974). One common misconception is that model accuracy invariably increases with model complexity. This is not the case. The tradeoff between model complexity and accuracy is not simply that increased model complexity increases model accuracy. Simpler catchment models can perform equally well or at least may not be substantially outperformed by more complex models (eg. Loague and Freeze 1985). Jakeman and Hornberger (1993) confirmed this result for different levels of complexity in conceptual models. Complex models suffer from problems with error accumulation and model identifiability, due to overparameterisation (Beven 1989, 1991, 1996). Beven (1989) argues that the physical nature of model parameters in physics based model does not circumvent problems of overparameterisation unless additional parameter observations are available at an appropriate scale. Beven (1991) states that 'in this sense then, physically-based distributed models are no different from any conceptual model'. The lack of available input data for such models means that many of the model parameters must be determined through calibration. This leads to problems of non-uniqueness and means that the physical interpretability of parameter values is questionable.

Empirical and simple conceptual models tend not to require large quantities of data and are computationally simple. In contrast, the physics based models require a large amount of input data and consequently can be difficult to use. This can be a particular problem in Australian catchments where input data is typically sparse. A large number of parameters in these models will have to be determined through calibration in such sparse data situations, raising difficulties with identifiability, model uniqueness and the physical interpretability of calibrated parameters. These problems will also be observed with complex conceptual models.

An important consideration in choosing models is the accuracy and validity of the model. This relates to the issue of suitability of the model to a particular environment. The use of models that are based on a large amount of detailed observations collected under different conditions (eg. different scales, soils, climatic regions) may not be valid or feasible in data poor catchments where physical catchment conditions may be very different.

4. CONCLUSIONS FOR MODELLING DATA POOR CATCHMENTS

Data sets on sediment and nutrient concentrations in data poor catchments are typically only available at large catchment scales of the order of 100 to 1000 square kilometres, as well as for a limited temporal period, often only up to a few years. Such information is inadequate to support the application of complex models which contain large numbers of parameters and/or which make detailed assumptions about the physical processes driving transport. Only the key catchment processes warrant description in such data poor circumstances.

Complex conceptual and physics based models also place high demands on the user, who must be very experienced technically in using models. Even for the experienced, the unique calibration of so many parameters is not possible. Different users will therefore obtain different parameter sets (Wheater *et al.* 1993).

In addition, physics based models are typically only designed to be applied at small scales. Their application to larger scales brings attendant problems of high computational requirements and errors associated with the application of these models in situations where the underlying assumptions are not met.

Therefore, it is only empirical models and simple conceptual models which can be considered as suitable for modelling catchment exports in data poor catchments.

5. ELEMENTS OF A NEW APPROACH

Given the problems with complex models the most practicable approach, particularly in data poor catchments, is one integrating an empirical landscape-factor approach with simple conceptual models (see Figure 1). Landscape factor models will be useful when predicting at ungauged sites, such as when conceptual models require predictions at sub-catchment or even landscape scales where no measurements are available. That is, they will be useful for disaggregating the exports predicted at larger catchment scales and lend themselves to being incorporated into conceptual models as sub-catchment scale predictions. Conceptual models will be useful especially to link sub-catchment exports and route them through catchment and basin networks. As runoff and discharge are the major drivers of catchment exports, a good conceptual model will be one which:

i) predicts runoff from catchments in response to precipitation (and historic catchment conditions) and then routes discharge and pollutants through an instream component; and

ii) incorporates the key processes (eg. quick flow, slow flow, stream advection, suspension and resettling) in a parametrically efficient manner.

With climate being the major determinant of long-term variability of catchment exports, a conceptual model, which allows for the input of rainfall and other climate variables (such as temperature), is essential to help characterise the variability of exports.

An approach combining a landscape factor model with conceptual model components, accounting for processes of rainfall-runoff, runoff-sediment generation, streambank erosion and in-stream routing of flow, nutrients and sediment, should provide a means for improved modelling and increased information on sediment and nutrient exports in data poor catchments.

The steps required to apply such a combined modelling approach may be as follows:

1. Calibrate and apply a rainfall-runoff model at all sites where there is any discharge data. An example of a rainfall-runoff model which could be used for this purpose is IHACRES. This model has been successfully applied for streamflow modelling in Australian catchments and worldwide (eg. Jakeman *et al.* 1993; Post and Jakeman 1996; Ye *et al.* 1997), while having a relatively simple parameterisation.
2. Derive relative indices of erosion and delivery potential based on landscape attributes for use in a landscape-factor model, such as that used by Moss *et al.* (1993).
3. Calibrate the landscape-factor model in similar scale catchments to generate estimates of static loads at subcatchment scales.
4. Construct a model to predict the streamflow dynamics at subcatchment scale to generate climate-sensitive loads. This would involve calibration of a runoff-sediment component on any sediment or nutrient concentration data that is available, using the already calibrated rainfall-runoff models. A runoff-sediment module could be used is the runoff-sediment component of the LASCAM model (Viney and Sivapalan 1999).
5. Apply an instream solute transport model to infer reach sources and sinks over reaches where upstream and downstream concentration data exist. This allows quantification of the significance of bank erosion and tributary inflows versus streamflow. The STARS model (Green *et al.* 1999) could be used for this component.

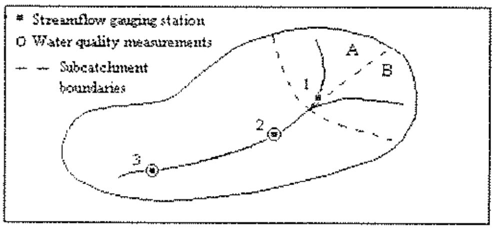

Figure 1 - Idealised catchment

Figure 1 illustrates some potential uses for such a combined modelling approach. The first of these is applied over the catchment scale, where load estimates are of interest at a number of points or nodes in the catchment. If discharge data is available at nodes 1, 2, and 3 (see Figure 1) and some concentration data is available at points 2 and 3, then it would be possible to calibrate the rainfall-runoff model at all three points. The STARS model, for example, could be implemented between nodes 2 an 3 to consider instream components, and to quantify the significance of streambank erosion. IHACRES and the runoff-sediment component of LASCAM, for example, could be used over the catchment scale to accounts for rainfall/ runoff and runoff/ sediment processes. A landscape factor model could be applied for small scale predictions to identify relative land use contributions. A second potential use of the approach is where limited sediment/nutrient concentration data is available for subcatchment A. The landscape factor approach could be calibrated to subcatchment B, an area of similar size to A, and could be linked to a rainfall-runoff model to allow for climate-sensitive loads. This model could then be applied to subcatchment A to provide information on pollutant concentrations.

It may be argued that this recommended approach is similar to the approach using (complex) conceptual models, such as AQUALM (Phillips *et al.* 1993). The main difference is the order in which each component is utilised and parameters are calibrated. The tendency with complex conceptual models seems to be to select parameters at subcatchment scales and to calibrate them at the larger catchment scale, generally ignoring the problem of parameter identifiability. The approach recommended here works in reverse, building simple lumped runoff and instream models at the larger catchment scale and disaggregating their predictions using a simple empirical but landscape driven method, that allows for the dynamics of runoff generation. Another advantage of the approach presented here is in the models suggested, which have advantages not only of best accuracy, but also of parameter efficiency.

ACKNOWLEDGMENTS

This work was undertaken as a part of a review for NSW EPA for the National Pollutant Inventory (NPI) Project of Environment Australia. The complete review report was published as a technical report entitled *Review of Techniques for Estimating Catchment Exports* (Letcher *et al.* 1999).

6. REFERENCES

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E. and Rasmussen, J. "An introduction to the European Hydrological System – Systeme Hydrologique Europeen, "SHE". 1. History and philosophy of a physically-based, distributed modelling system" *Journal of Hydrology*, 87, pp.45-59, 1986.

Beck, M.B. "Water quality modelling: a review of uncertainty" *Water Resources Research*, 23(8), pp.1393-1442, 1987.

Beck, M.B., Jakeman, A.J., and McAleer, M.J "Constuction and evaluation of models of environmental systems" in Beck, M.B., and McAleer, M.J. (eds.), *Modelling Change in Environmental Systems*, John Wiley and Sons, pp. 3-35, 1995.

Bennett, J. P. "Concepts of mathematical modelling of sediment yield." *Water Resources Research*, 10, pp.485-492, 1974.

Beven, K. "Changing ideas in hydrology - the case of physically-based models." *Journal of Hydrology*, 105, pp.157-172, 1989.

Beven, K. "Spatially distributed modelling: conceptual approach to runoff prediction" in Bowles, D.S. and O'Connell, P.E. (eds.), *Recent Advances in the Modelling of Hydrological Systems*, Kluwer Academic, Boston, pp.373-387, 1991.

Beven, K., "A discussion of distributed hydrological modelling" in Abbott, M.B., Refgaard, J.C (eds.), *Distributed Hydrological Modelling*, Kluwer Academic, pp.255-278, 1996.

Bloschl, G. and Sivapalan, M. "Scale issues in hydrological modelling" in Kalma, J.D. and Sivapalan, M. (eds.) *Scale Issues in Hydrological Modelling*, John Wiley and Sons, England, pp.9-48, 1995.

Dunin, F. "The use of physical process models" in Chapman, T. and Dunin, F. (eds.), *Prediction in Catchment Hydrology - A National Symposium on Hydrology*, Australian Academy of Science, Canberra, pp.277-291, 1975.

Grayson, R., and Moore, I. "Effect of land surface configuration on catchment hydrology" in Parsons, A. and Abrahams, A., (eds.), *Overland flow: hydraulics and erosion mechanics*, University College London Press Ltd, London, pp.147-175, 1993.

Green, T.R., Beavis, S.G., Dietrich, C.R., and Jakeman, A.J. "Relating stream-bank erosion to in-stream transport of suspended sediment" *Hydrological Processes*, 13(5), pp. 777-787, 1999.

Jakeman, A.J., Green, T.R., Beavis, S.G., Zhang, L., Dietrich, C.R., and Crapper, P.F. "Modelling upland and in-stream erosion, sediment and phosphorus transport in a large catchment" *Hydrological Processes*, 13(5), pp.745-752, 1999.

Jakeman, A.J. and Hornberger, G.M. "How much complexity is warranted in a rainfall-runoff model?" *Water Resources Research*, 29(8), pp. 2637-2649, 1993.

Kleissen, F., Beck, M., and Wheater, H. "The identifiability of conceptual hydrochemical models" *Water Resources Research*, 26(12), pp.2979-2992, 1990.

Lane, L. J., Renard, K. G., Foster, G. R., and Laflen, J. M. "Development and application of modern soil erosion prediction technology - the USDA experience" *Australian Journal of Soil Research*, 30, pp.893-912, 1992.

Lane, L., Nichols, M., and Paige, G. "Modeling erosion on hillslopes: concepts, theory and data" in Binning, P., Bridgman, H. and Williams, B. (eds.) *MODSIM95*, 1, pp.1-7, Newcastle, November 27-30, 1995.

Letcher, R.A., Jakeman, A.J. , McKee, L.J., Merritt, W.S. Eyre B.D. and Baginska, B. "Review of Techniques to Estimate Catchment Exports", Technical Report, NSW EPA, 1999.

Loague, K., and Freeze, R. "A comparison of rainfall-runoff modelling techniques on small upland catchments." *Water Resources Research*, 21, pp.229-248, 1985.

Phillips, B. C., Blaik S, and Tilley, J. "Modelling the effects of land-use change on receiving waters." *Hydrology and Water Resources Symposium*, Newcastle, Australia, pp. 39-44, 1993.

Seyfried, M.S., and Wilcox, B.P. (1995) "Scale and the nature of spatial variability: field examples having implications for hydrologic modeling", *Water Resources Research*, 31(1), pp.173-184.

Sorooshian, S. "Parameter estimation, model identification, and model validation: conceptual type models", in Bowles, D.S. and O'Connell, P.E. (Eds.) *Recent Advances in the Modelling of Hydrological Systems*, Kluwer Academic, pp.443-467, 1991.

Spear, R. C. "Large simulation models: calibration, uniqueness and goodness of fit" in Binning, P., Bridgman, H. and Williams, B. (eds.) *MODSIM95*, 1, pp.8-15, Newcastle, November 27-30, 1995.

Viney, N.R., and Sivapalan, M. "A conceptual model of sediment transport: application to the Avon River Basin in Western Australia", *Hydrological Processes*, 13, pp.727-743, 1999.

Wheater, H.S., Jakeman, A.J., and Beven, K.J. "Progress and directions in rainfall-runoff modelling", in Jakeman, A.J., Beck, M.B. and McAleer, M.J. (eds.) *Modelling Change in Environmental Systems*, Chichester, John Wiley and Sons, pp.101-132, 1993.