Stochastic modelling of phosphorus transfers from agricultural land to aquatic ecosystems

M.J. Whelan*, E.G. Hope* and K. Fox**

* Department of Environmental Science, University of Stirling, FK9 4LA, UK

** Unilever Research, Port Sunlight, Wirral, CH63 3JW, UK

Abstract This paper describes a simple model of phosphorus (P) transfer from agricultural land to surface waters which incorporates the effects of spatial variability in catchment properties and uncertainty in model parameter values. TOPMODEL concepts are used to estimate water, solute and sediment fluxes to water bodies. The model predicts the spatial distribution of water table depth and saturation-excess overland flow based on topography. Dissolved P (DP) transfer is assumed to occur vertically in the unsaturated zone and laterally in the saturated zone. Readily soluble P is assumed to decrease exponentially with soil depth. Particulate P (PP) transfers are modelled by estimating overland flow discharge and associated sediment transport capacity. Uncertainty in the distribution of soil surface P concentrations and model parameters controlling the mobility of soil P are incorporated using Monte Carlo simulation. Predicted losses of DP are well correlated with discharge and those of PP are episodic. Highest losses of P tended to be predicted near to the stream where the water table is close to the surface. The combination of a deterministic model core with a stochastic generation of model parameters or state variables provides an attractive way of embracing variability and uncertainty in models of this kind.

Keywords Model; phosphorus; stochastic; surface waters; transfer

Water Science and Technology **Vol 45 No 9 pp 167–176** © IWA Publishing 2002

Introduction

Phosphorus (P) is a major limiting nutrient in many freshwater ecosystems (e.g. Foy and Bailey-Watts, 1998). Increases in its availability can result in enhanced primary productivity which can lead to eutrophication. Although simple export coefficient models of P transfers from agricultural land to surface waters (e.g. Johnes and Hodgkinson, 1998) have some value in predicting annual fluxes on the basis of land use, they cannot predict intra-annual seasonality in P losses or provide information on the role of hydrological processes. Conversely, the utility of detailed physically-based models is often hampered by a lack of input data or model parameters at a suitable level of spatial and temporal resolution. Even where suitable input data exist, complex models often give little improvement in predictive capability over much simpler models. Furthermore, despite widespread recognition that some account should be taken of system variability and parameter uncertainty, relatively little attention has been devoted to incorporating these features into models of P transfer. An exception is the recent work of Hession and Storm (2000), who incorporated uncertainty estimates into a catchment-scale model of annual P loads, although their model is rather restricted in terms of temporal and spatial resolution. This paper describes a simple, process-based, parametrically parsimonious model of P transfer from catchments which uses readily available input data and which incorporates the effects of spatial variability in catchment properties and uncertainty in model parameters.

Methods

Hydrological model

The model is based on TOPMODEL concepts (Beven and Kirkby, 1979; Beven, 1997) which attempt to incorporate the influence of topography on runoff response in a simple "semi-lumped" fashion. In our model, the root zone is conceptualised as a single store with a maximum capacity equal to the integrated profile porosity. The water content of the root zone store is manifested in terms of a soil moisture deficit from saturation, δ_R (L), which is depleted by infiltration and augmented by evapotranspiration (ET) and drainage. Surface-atmosphere interactions (interception and ET) are accounted for using a procedure based on a modification of the UK MORECS (Meteorological Office Rainfall and Evaporation Calculation System: Thompson *et al*., 1981). Essentially, the scheme employs the simple "big leaf" concept combined with the Penman-Monteith equation although a simpler ET model (Hargreaves and Samani, 1982) may be employed if meteorological data are limited to mean daily temperatures and daily rainfall totals. Crop growth (height and leaf area index, *LAI*) throughout the year is accounted for by linearly interpolating between marker values typical for crops grown in the area under consideration. Infiltration rates are assumed to always exceed the rate of net precipitation (unless the soil becomes saturated), therefore precluding the generation of infiltration-excess overland flow. This assumption is a direct result of the daily time step chosen for the model and the consequent lack of information on rainfall intensity.

It is assumed that root zone drainage occurs entirely in the vertical direction and that matric potential gradients in the root zone approach zero when the root zone moisture content is high enough to allow significant drainage. This allows the use of the one dimensional steady-state gravity flow equation for root zone drainage (Jury *et al*., 1991) in which vertical discharge per unit area, r (LT^{-1}), is equal to the unsaturated hydraulic conductivity (K). The relationship between K and δ_R is approximated using the van Genuchten (1980) function which is based on parameters obtained from the moisture retention curve.

In TOPMODEL, discharge at the catchment outlet is assumed to be inversely proportional to a catchment-wide average deficit from saturation, $\bar{\delta}(t)$. This deficit is regularly updated by mass balance:

$$\frac{d\bar{\delta}}{dt} = Q - \bar{r} \tag{1}$$

where $Q(t)$ (LT^{-1}) is the saturated zone discharge per unit area per unit time and $\bar{r}(t)$ (LT^{-1}) is the average input to the saturated zone per unit time (i.e. vertical soil drainage). The lumped catchment water balance can be related to the spatial distribution of point values of the saturation deficit, $\delta(t,x,y)$, via the following equation (see Beven, 1997)

$$\delta(t,x,y) \;=\; \bar{\delta}(t) \;\; m[\lambda(x,y) - \bar{\lambda}] \tag{2a}$$

where x and y are the co-ordinates of the point in question, m (L) is a curve parameter, which describes the rate of decrease in saturated hydraulic conductivity with depth, and $\lambda(x,y)$ is the topographic index

$$\lambda(x,y) = \ln\left(\frac{a}{\tan\beta}\right) \tag{2b}$$

in which a is the area drained per unit contour length (L) and β is the local slope angle for point (x, y). $\bar{\lambda}$ is the catchment average of $\lambda(x,y)$. Note that the derivation of Eq. (2) assumes a spatially uniform soil transmissivity when δ is zero.

For any value of $\bar{\delta}(t)$ there will be a range of point-specific values for $\delta(t,x,y)$, which will be determined according to values of $\lambda(x,y)$, scaled by the parameter m. Areas in which $\delta(t,x,y) \leq 0$ will generate saturation excess overland flow during storms. Similarly, areas where the root zone is regularly influenced by the saturated zone will have an increased likelihood for lateral transfer of dissolved material in the soil. The spatial distribution of λ

can be derived automatically from a digital elevation model (*DEM*). Areas with similar values are considered to be "hydrologically similar", such that calculations need only be made for classes of λ (obtained from its discrete probability density function [*pdf*]) rather than for every point in the basin.

Different crops have different sowing and harvest dates, and different phenologies which can affect the time course of the root zone water balance and average drainage rates to the saturated zone. Root zone water dynamics are, therefore, calculated separately for each land use in each class of λ. The proportion of each land use type in each class of λ is obtained by superimposing land use and λ layers in a GIS. The average daily (area-weighted) root zone water balance is calculated by

$$\frac{\Delta\bar{\delta}_R}{\Delta t} = \bar{r} + \bar{e} - \bar{n} \tag{3}$$

where $\bar{r}$ is the catchment average drainage from the root zone, $\bar{e}$ is the catchment average rate of ET and $\bar{n}$ is net precipitation (after interception). All terms on the right hand side are rates (LT^{-1}). Values of δ are converted to equivalent water table depths by assuming an effective porosity for the "actively draining fraction" which can be considered equivalent to the pore space between "field capacity" and saturation (Quinn *et al.*, 1995).

Spatial distribution of soil phosphorus

Soil P concentrations will vary across the landscape in response to vegetation or cropping history, soil texture and drainage. In the absence of detailed information on the spatial distribution of P we assume that it is imperfectly correlated with λ. The *pdf* of λ can often be approximated by a log-normal model. In this case, for each point (x,y) the equivalent standard normal deviate, $z(x,y)$ is

$$z = \left(\frac{\ln(\lambda) - \mathrm{E}[(\ln(\lambda)]}{\mathrm{s.d.}[(\ln(\lambda)]} \right) \tag{4}$$

where E[ln(λ)] and s.d.[ln(λ)] are, respectively, the mean and standard deviation of the log-transformed *pdf* of λ. Given the mean and standard deviation for P content at the soil surface, $P(x,y,0)$ ($M\ M^{-1}$), along with the correlation coefficient (ρ) between $P(x,y,0)$ and $\lambda(x,y)$ and assuming a log-normal *pdf* for $\hat{P}(x,y,0)$, then stochastic realisations of P content, $P(x, y, 0)$, can be obtained from:

$$\hat{P} = \exp\{\mathrm{E}[\mathrm{In}(P)] + \mathrm{s.d.}[\mathrm{In}(P)]\mu\} \tag{5}$$

where E[ln(P)] and s.d.[ln(P)] are, respectively, the mean and standard deviation of the log-transformed *pdf* of P and μ is a random normal deviate, correlated with $z(x,y)$ using

$$\mu = z \cdot \rho + \nu\sqrt{(1-\rho^2)} \tag{6}$$

in which ν is an independent random normal deviate.

This procedure attempts to predict the global features ("texture") of P concentrations across the landscape (Deutsch and Journel, 1992) but is not conditional (i.e. it does not attempt to match specific, locally measured, values). The acceptability of the spatial representations produced is based on preserving (contemporaneously) the statistics of P content (mean, variance and shape of the distribution) and its correlation with λ. A number of model iterations will generate frequency distributions of P concentration for each cell which can then be used to indicate uncertainty (Deutsch and Journel, 1992).

Depth distribution of soil phosphorus

Soil P is assumed to decrease with depth (after Haygarth *et al.*, 1998) according to a negative exponential function i.e.

$$P(x,y,z) = P(x,y,0)\exp(-k_p z) \tag{7}$$

where $P(x,y,z)$ is the concentration of P at depth z (L) and k_p is a curve parameter which describes the rate of decrease in P concentration with depth. If we assume a constant bulk density (ρ_B) with depth in the entire soil profile, then we can express $P(x,y,z)$ and $P(x,y,0)$ in terms of relative density (ML^{-3}) by multiplying by ρ_B. The total P content, $P_{TOT}(x,y)$ (ML^{-2}), in a soil profile with depth z_{MAX} is thus:

$$P_{TOT}(x,y) = \int_{z=0}^{z\max} P(x,y,z)\rho_B dz = \frac{P(x,y,0)\rho_B}{k_p}\left(1-\exp(-k_p z_{MAX})\right) \tag{8}$$

If we know the depth to the water table (z_{SAT}) we can calculate the fraction (F_s) of total P which is beneath the water table:

$$F_s = 1 - \left(\frac{1-\exp(-k_p z_{SAT})}{1-\exp(-k_p z_{MAX})}\right) \tag{9}$$

The fraction of P_{TOT} which is in the unsaturated zone is thus $(1-F_s)$.

P losses

Dissolved P (DP) losses ($ML^{-2}T^{-1}$) from both the saturated (lateral) and unsaturated (vertical) zones (FDP_S and FDP_U respectively) for each class of λ are calculated from:

$$FDP_S = F_S \cdot P_{TOT} \cdot k_A \cdot \exp(-\delta(t,\lambda)/m) \tag{10}$$

$$FDP_U = (1-F_S) \cdot k_A \cdot P_{TOT} \bar{r} \tag{11}$$

where k_A is the proportion of soil P which is in solution. Note that FDP_S is weighted to account for the exponential decrease in soil transmissivity with increasing depth and the consequent decrease in subsurface discharge. The assumption that DP is a constant proportion of total P is rather unrealistic and fails to account for the influence of important DP sinks such as plant uptake (which will be seasonal) and the occurrence of washout/exhaustion phenomena. In its current state the model can, therefore, be considered static in terms of soil P and variations in model output will largely be the result of hydrological processes. Concentrations of DP (C_{DP}) in the stream are simply:

$$C_{DP} = (FDP_S + FDP_U)/Q \tag{12}$$

Particulate P losses ($ML^{-2}T^{-1}$) are calculated from an estimate of wash erosion in overland flow. Neglecting the effects of rain spash and soil creep, we adopt the equation of Kirkby and Cox (1995)

$$S = k_S \tan\beta \left(\frac{q_S}{q_0}\right)^2 \tag{13}$$

where S is sediment transport per unit area ($ML^{-2}T^{-1}$), q_S (LT^{-1}) is the depth equivalent overland flow discharge and q_0 and k_S are fitted constants with dimensions of overland flow discharge (LT^{-1}) and sediment transport per unit area ($ML^{-2}T^{-1}$) respectively. Since hydrological calculations are only made for classes of λ rather than for every cell in the catchment, the mean gradient for each class of λ was used in Eq. (13). Sediment associated P transfers (FPP, $ML^{-2}T^{-1}$) are assumed to be equal to the product of sediment transport rate and the P concentration at the soil surface. This ignores the possibility of P enrichment due to size-selective entrainment or deposition and/or aggregate stripping (e.g. Sharpley and Smith, 1990) but in the absence of catchment-specific data, the use of an enrichment ratio would only introduce additional uncertainty to an already parameter-rich model. Thus

$$FPP = S.P(x,y,0) \tag{14}$$

Application to the Slapton Wood catchment

The model was applied to the Slapton Wood catchment (Figure 1), a 0.93 km^2 instrumented basin in Devon, UK, with a mixed land use (*cf* Burt *et al.*, 1988).

The spatial distribution of λ was derived from a "depitted" 20 m *DEM* using a slight modification of the multiple flow direction algorithm described by Quinn *et al.* (1991). The root zone hydraulic parameters used in the model were derived from measurements (e.g. Ragab and Cooper, 1993). Parameters for the subsurface flow component of the model were calibrated by minimising the error between predicted and observed daily discharge using data from the calendar year 1971. Root zone parameters were not changed during calibration. Once calibrated the model was run continuously for the period 1970–1985 (with no further adjustments) in four scenarios with different assumptions about the distribution of soil P at the soil surface (Table 1). Parameters describing the spatial distribution and availability of soil P and those describing P transfer were approximated from the literature. For those scenarios in which there is a stochastic element, thirty realisations of the model were performed so as to provide information about uncertainty in model output.

Since the hydrological component of the model has been calibrated and validated (and can, therefore, be considered to be a reasonable representation of the system) the P model parameters represent the largest degree of uncertainty in the model. The effect of including

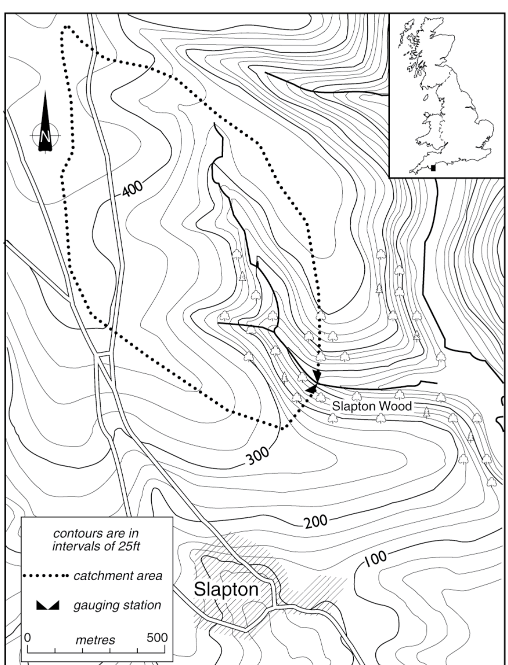

Figure 1 Location map of the Slapton Wood catchment

Table 1 Details of scenarios for which the model was tested. In all cases Mean P(x,y,0) was 15 mg P kg^{-1}

Scenario name	CV for P(x,y,0) (%)	ρ
A	40	0.0
B	0	0.0
C	40	0.5
D	40	1.0

realistic uncertainty in the values of those parameters to which the model was most sensitive was therefore investigated in an additional model run for one year (1972) with 2000 iterations. Values of q_o, k_s, k_p, and k_A were incorporated as uniform probability distributions (see Table 2 for parameters).

Results and discussion

Examples of predicted and measured daily stream hydrographs are shown in Figure 2. The results suggest that the hydrological model structure and most of its assumptions are reasonable for this catchment. On average, total annual predicted runoff was greater than measured runoff (Burt *et al.*, 1988) by just 8.3% (1971–1990). Model results suggest

Table 2 Details of *pdf*s used to examine the impact of parameter uncertainty on model response. In each case a uniform *pdf* was used with the maximum value arbitrarily set at 4 times the minimum value

Parameter	Maximum	Minimum
k_A	0.00002	0.000005
k_s	0.02	0.005
q_o	0.02	0.005
k_p	10	2.5

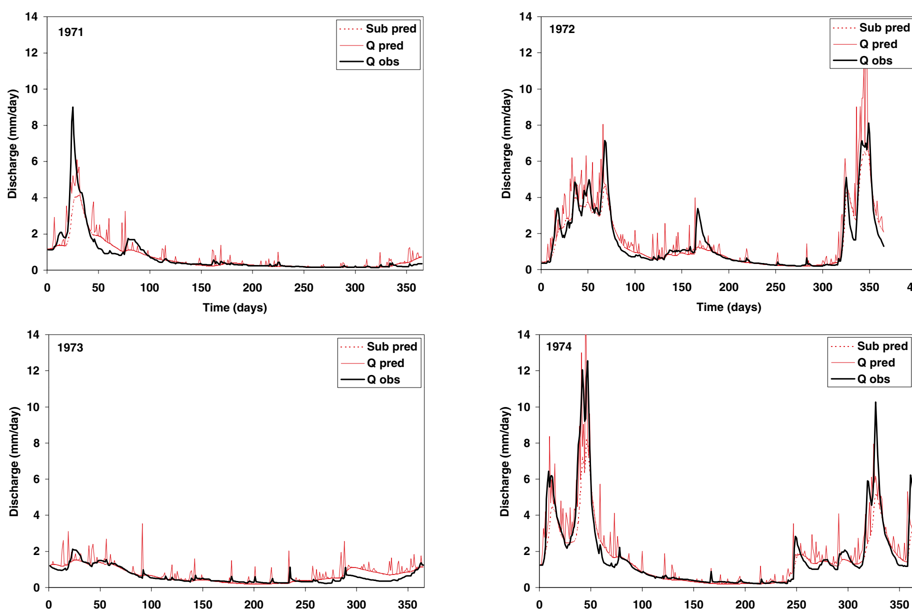

Figure 2 Examples of predicted (*Q pred*) versus observed (*Q obs*) daily runoff for the Slapton Wood catchment for 1971, 1972, 1973 and 1974. *Sub pred* is the predicted sub-surface flow component

that surface saturation rarely occurs in more than 10% of the total catchment area and is usually less than 5%, mostly in and around the stream channels.

Predicted area-weighted DP, PP and TP losses from the Slapton Wood catchment are shown in Figure 3 for scenario A. In general predicted DP losses were greater than those for PP, although the relative importance of *FDP* and *FPP* would be altered during calibration of the P model. As expected, highest losses tended to be predicted in years with high annual precipitation. Average predicted loads for 1972 for each scenario are shown in Figure 4, as an example. Consistently higher values of P were predicted for Scenario D ($\rho = 1$) than for the other scenarios. This is to be expected since high soil P concentrations are always assumed to occur in hydrologically active areas (i.e. areas with high values of λ). In scenarios A and B no relationship was assumed between soil P availability and topography. Even though a coefficient of variation of 0.4 was assumed in scenario A, the fact that soil P was randomly distributed in each stochastic realisation resulted in a very similar predicted average annual flux to scenario B (constant P concentrations throughout the catchment). Results for scenario C were intermediate between those for A/B and those for D. This is consistent with the assumptions of variable soil P and positive correlation, ρ, which tends to enhance P transfers by tending to predict coinciding high soil P concentrations, water table and a high incidence of overland flow.

The predicted pattern of mean daily DP concentration (with estimated uncertainty) and the predicted frequency distributions of DP, PP and TP loads resulting from a full incorporation of P parameter uncertainty are shown in Figure 5. It is clear that incorporating a realistic estimate of uncertainty generates wide uncertainty intervals which can only be reduced by better identification of model parameter values.

The results described in this paper were produced under the implicit assumption that the hydrological component of the model provides an adequate representation of the spatial and temporal variations in flow and water table depth in the Slapton Wood catchment. However, it is now widely recognised that there may be a number of different combinations of parameter values which produce reasonable fits to observed discharge (e.g. Beven, 1997), but relatively few which will also generate good predictions of water table depth variations. It is important to realise that without measurement of water table depths the model will always be poorly constrained and unique calibrations impossible.

The adoption of a daily time step (which permits model application to catchments with minimal meteorological data and which reduces run times for Monte Carlo simulations) hinders the estimation of infiltration excess overland flow. This will be important in catchments where the generation of surface flow by this mechanism is common, although its impact in model applications in the UK (where infiltration rates are normally greater than

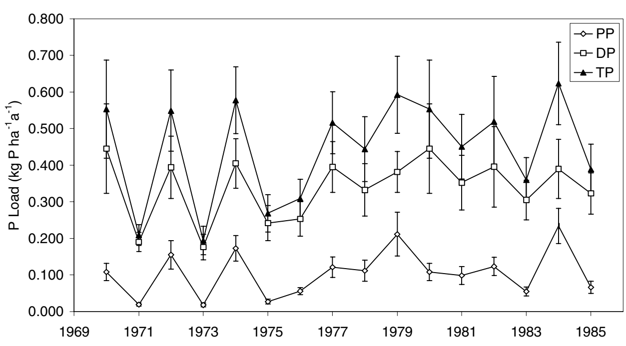

Figure 3 Predicted annual P losses from the Slapton Wood catchment (1970–85) for Scenario A. Error bars show mean ± 1 SD for 30 realisations

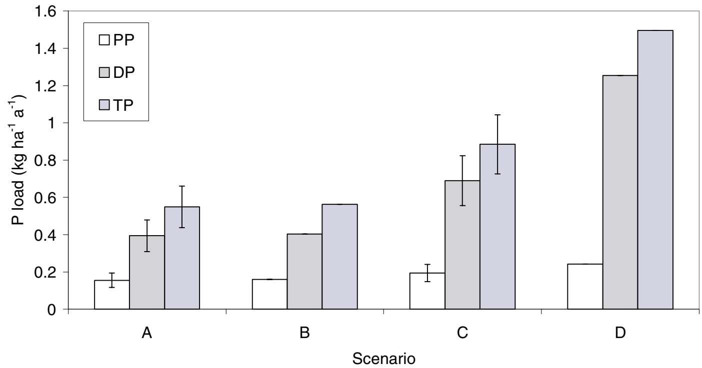

Figure 4 Predicted annual P losses from the Slapton Wood catchment (1972) – Scenarios A–D. Error bars show the mean ± 1 SD for 30 realisations

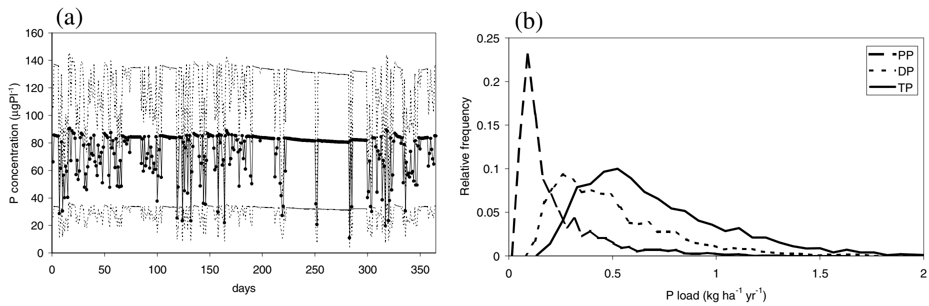

Figure 5 (a) Predicted mean daily DP concentration in the Slapton Wood stream ± 1 SD and (b) frequency distributions of predicted annual DP, PP and TP loads for 1972 produced from 2,000 iterations with values of $P(x,y,0)$ selected according to Scenario A and parameter values selected randomly from uniform *pdf*s (see Table 2)

rainfall intensities) will be limited. At present model output has not yet been compared with observed data on P transfers. The analysis presented here is not intended to evaluate model performance, although calibration and validation of the model will be essential prior to further development. Rather, emphasis is placed on the utility of incorporating state variable and parameter uncertainty using Monte Carlo procedures within a deterministic model framework, which is facilitated by a relatively simple structure and low number of parameters.

A number of simplifying assumptions and omissions have been made in the model which need to be addressed in further developments. The model does not simulate dynamic fluctuations in the sizes of soil P pools and is thus unable to take account of seasonal variations in soil P availability. Considerable seasonal variations in soil P levels have been reported in the literature (e.g. Sharpley, 1985) and will certainly have some control over soil P losses. Other important factors which have been omitted include the effects of land use and P amendments, the potential for P removal by deposition or plant uptake from surface or subsurface flows and the role of in-stream transformations and P input from stream bank erosion.

Conclusions

A simple model of P transfer from agricultural land to surface waters was presented. The model is appropriately-scaled for simulating processes in small to medium sized catchments and is parameter-efficient. It represents an improvement on previous approaches to modelling catchment-scale P transfer in providing greater temporal and spatial resolution than annual export-coefficient models and lower parameter and input data requirements than more complex, event-based, models. Although the model requires further development and has yet to be validated, the idea of combining a deterministic process-based model core with a stochastic generation of uncertain state variables and parameter values, along the lines described, is attractive since it embraces variability and uncertainty whilst maintaining a synthesis of our understanding of the system dynamics.

Acknowledgements

The authors would like to thank Unilever for funding and Claudio Gandolfi for comments on the model.

References

Beven, K. (1997). TOPMODEL: A Critique. *Hydrol. Procs.*, **11**, 1069–1085.

Beven, K. and Kirkby, M.J. (1979). A physically-based, variable contributing area model of basin hydrology. *Hydrol. Sci. Bull.*, **24**, 43–69.

Burt, T.P., Arkell, B.P., Trudgill, S.T. and Walling, D.E. (1988). Stream nitrate levels in a small catchment in south west England over a period of 15 years (1970–1985). *Hydrol. Procs.*, **2**, 267–284.

Deutsch, C.V. and Journel, A.G. (1992). *GSLIB. Geostatistical Software Library and User's Guide*. OUP, Oxford, UK.

Foy, R.H. and Bailey-Watts, A.E. (1998). Observations on the spatial and temporal variation in the phosphorus status of lakes in the British Isles. *Soil Use and Man.*, **14**, 131–132.

Hargreaves, G.H. and Samani, Z.A. (1982). Estimating potential evapotranspiration. *J. Irrig. Drainage Engin.*, **108**, 225–230.

Haygarth, P.M., Hepworth, L. and Jarvis, S.C. (1998). Forms of phosphorus transfer in hydrological pathways from soil under grazed grassland. *European J. Soil Sci.*, **49**, 65–72.

Hession, W.C. and Storm, D.E. (2000). Watershed level uncertainties: Implications for phosphorus management and eutrophication. *J. Env. Qual.*, **29**, 1172–1179.

Jury, W.A., Gardner, W.R. and Gardner, W.H. (1991). *Soil Physics*. John Wiley and Sons, New York, USA.

Johnes, P.J. and Hodgkinson, P.A. (1998). Phosphorus losses from agricultural catchments. *Soil Use and Man.*, **14**, 175–185.

Kirkby, M.J. and Cox, N.J. (1995). A climatic index for soil erosion potential (CSEP) including seasonal and vegetation factors. *Catena*, **25**, 333–352.

Quinn, P., Beven, K., Chevallier, P. and Planchon, O. (1991). The prediction of hillslope flow patterns for distributed hydrological modelling using digital terrain models. *Hydrol. Procs.*, **5**, 59–79.

Quinn, P., Beven, K. and Culf, A. (1995). The introduction of macroscale hydrological complexity into land surface-atmosphere transfer models and the effect on planetary boundary layer development. *J. Hydrol.*, **166**, 421–444.

Ragab, R. and Cooper, J.D. (1993). Variability of unsaturated zone water transport parameters: Implications for hydrological modelling. 1. In situ measurements. *J. Hydrol.*, **148**, 109–131.

Sharpley, A.N. (1985). Phosphorus cycling in unfertilised and fertilised agricultural soils. *Soil Sci. Soc. Am. J.*, **49**, 905–911.

Sharpley, A.N. and Smith, S.J. (1990). Phosphorus Transport in Agricultural Runoff: The Role of Soil Erosion. In: *Soil Erosion on Agricultural Land*, Boardman, J, Foster, I.D.L. and Dearing, J.A. (eds.), Wiley, Chichester, UK.

Thompson, N., Barrie, I.A. and Ayles, M. (1981). *The Meteorological Office Rainfall and Evaporation Calculation System: MORECS (July 1981)*, Met. Office Memorandum No.45, Bracknell, UK.

Van Genuchten, M.Th. (1980). Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils. *Soil Sci. Soc. Am. J.*, **44**, 892–898.