Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-314
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 6 June 2017

Hydrology and Earth System Sciences
Open Access
Discussions

Data-based mechanistic model of catchment phosphorus load improves predictions of storm transfers and annual loads in surface waters

Mary C.Ockenden[1], Wlodek Tych[1], Keith J.Beven[1], Adrian L.Collins[2], Robert Evans[3], Peter D.Falloon[4], Kirsty J.Forber[1], Kevin M.Hiscock[5], Michael J.Hollaway[1], Ron Kahana[4], Kit J.A.Macleod[6], Martha L.Villamizar[7], Catherine Wearing[1], Paul J.A.Withers[8], Jian G.Zhou[9], Sean Burke[10], Richard J.Cooper[5], Jim E.Freer[11], Philip M.Haygarth[1]

[1]Lancaster Environment Centre, Lancaster University, Bailrigg, Lancaster LA1 4YQ, England, UK
[2]Rothamsted Research North Wyke, Okehampton, Devon EX20 2SB, England, UK
[3]Global Sustainability Institute, Anglia Ruskin University, Cambridge CB1 1PT, England, UK
[4]Met Office Hadley Centre, Exeter, Devon EX1 3PB, England, UK
[5]School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, England, UK
[6]James Hutton Institute, Aberdeen AB15 8QH, Scotland, UK
[7]School of Engineering, Liverpool University, Liverpool L69 3GQ, England, UK
[8]Bangor University, Bangor, Gwynedd LL57 2UW, Wales, UK
[9]School of Computing, Mathematics & Digital Technology, Manchester Metropolitan University, Manchester M1 5GD, UK
[10]British Geological Survey, Keyworth, Nottingham NG12 5GG, England, UK
[11]University of Bristol, Bristol BS8 1SS, UK

Correspondence to: Philip M. Haygarth (p.haygarth@lancaster.ac.uk)

Abstract. Excess nutrients in surface waters, such as phosphorus (P) from agriculture, result in poor water quality, with adverse effects on ecological health and costs for remediation. However, understanding and prediction of P transfers in catchments have been limited by inadequate data and over-parameterised models with high uncertainty. We show that, with high temporal resolution data, we are able to identify simple dynamic models that capture the P load dynamics in three contrasting agricultural catchments in the UK. For a flashy catchment, a linear, second-order (two pathways) model for discharge gave high simulation efficiencies for short-term storm sequences and was useful in highlighting uncertainties in out-of-bank flows. A model with non-linear rainfall input was appropriate for predicting seasonal or annual cumulative P loads where antecedent conditions affected the catchment response. For second-order models, the time constant for the fast pathway varied between 2 and 15 hours for all three catchments and for both discharge and P, confirming that high temporal resolution (hourly) data are necessary to capture the dynamic responses in small catchments (10-50 km^2). The models led to a better understanding of the dominant nutrient transfer modes, which will, in-turn, help in planning appropriate pollution mitigation measures.

1 Introduction

The quality of both surface waters and groundwater is under increasing pressure from numerous sources, including intensive agricultural practices, water abstraction, climate change and changes in food production and housing provision to cope with population growth (Carpenter and Bennett, 2011). Sediment and nutrient concentrations and loads are of concern to water utility companies and to environmental regulators who are striving to meet stringent water quality standards. However, accurate estimation of loads requires accurate, high temporal resolution measurements of both discharge and nutrient concentrations (Johnes, 2007) and should include quantification of observational uncertainties (McMillan et al., 2012). Sediment and nitrogen are frequently and relatively easily measured in-situ. In contrast, phosphorus (P) concentrations for water quality assessments are typically measured by manual or automatic sampling followed by laboratory analysis, often at monthly resolution, which do not capture the dynamic nature of P concentrations, and result in biased estimates of P load (Cassidy and Jordan, 2011). Phosphorus concentration in rivers and streams is controlled by many factors, including rainfall, runoff, point sources, diffuse inputs and in-stream P retention and processing. Some of these factors, particularly for small catchments, change at timescales of minutes to hours, and thus the dynamics of P concentration and load need to be studied at similar time scales. In this study, hourly time series of rainfall, runoff and P concentrations are used to help understand hydrological transport pathways of P for three contrasting agricultural catchments across the UK.

There is a wide range of complexity in hydrological and water quality models, applicable at a range of scales and for different purposes. In most models there is a balance between practical simplifications and model complexity, which depends on catchment size, knowledge (or lack of) of the hydrological processes, data availability and computing power. Some of the less complex models for diffuse pollution include export coefficient models (Johnes, 1996) and the Phosphorus Indicators Tool (PIT) (Heathwaite et al., 2003; Liu et al., 2005). The most complex water quality models are idealised, process-based representations of our best understanding of reality, with a highly complex, fixed structure and many parameters, for which there is often little or no site specific data (Dean et al., 2009). These models often include a component for sediment-bound P, where the sediment transfer is based on a form of the Universal Soil Loss Equation (USLE), which is a process-based model known to perform poorly (Evans and Boardman, 2016). Results generated by such process-based models are often highly uncertain, due to the uncertainty in both the model parameters and the model structure (Parker et al., 2013; Jackson-Blake et al., 2015). A review of pollutant loss studies using one process-based model, the Soil Water Assessment Tool (SWAT), revealed that most applications used a monthly time step for calibration, with few applications using a daily time step and none using a sub-daily time step. Model fit for total P (TP) concentration, measured by Nash Sutcliffe Efficiency, often exceeded 0.5 but could be as low as -0.08 for daily calibration. Depending on the calibration criteria, there may be many different parameter sets that fit the calibration data equally well, but because of a lack of data on internal variables, the models do not necessarily fit for the right reasons. Moriasi et al. (2007) advised using several different criteria for assessment of model fit, including a graphical assessment as well as quantitative metrics.

However, process-based models still often fail to meet the acceptance criteria (Jackson-Blake et al., 2015), even when these are relaxed to account for additional uncertainties in the measured input data (Harmel et al., 2006) such as those due to sampling method, sample storage or fractionation (Jarvie et al., 2002).

As a more simple alternative, Young et al. (1996) recommended constructing models that capture the dominant modes of a system, with as few tuneable parameters as possible. Transfer function models, whose structure and parameters are determined by the information in the data, are considered to be among the most parsimonious for rainfall-flow relationships (McGuire and McDonnell, 2006; Young, 2003). Data-Based Mechanistic (DBM) modelling, which uses time-series data and fits a range of transfer functions, allows the structure of the model to be determined by the information in the monitoring data. There will still be structural errors in a DBM model, as it tries to represent a continuum of flow pathways with just the dominant modes, but this simplification will be determined by the information in the data rather than being pre-selected. An optimal model and associated parameters are identified using statistical measures, but a model is only accepted if it has a plausible physical explanation (Young, 1998, 2003; Young and Beven, 1994; Young et al., 2004). With the increasing availability of high temporal resolution datasets for additional variables alongside stream discharge (Bieroza and Heathwaite, 2015; Bowes et al., 2015; Halliday et al., 2015; Outram et al., 2014), this technique has been used effectively for relating rainfall to hydrogen ion (Jones and Chappell, 2014) concentration in rivers, and rainfall to dissolved organic carbon (Jones et al., 2014).

The aim of this study was to investigate, for the first time, whether simple dynamic models of P load could be identified to help understand the hydrological P processes within three contrasting agricultural catchments in the UK that represent a range of climate, topography, soil and farming types. Specifically, the objectives were:

- To identify rainfall-runoff models for each catchment, from hourly time series data collected over three years
- To develop models of P load exported from each catchment, using hourly time series data of P concentrations measured with in-situ bankside analysers
- To improve understanding of the dominant modes of a catchment through comparison of rainfall-runoff and rainfall-TP load models for each catchment.

If successful, this would be the first time that DBM modelling has been applied to high-resolution phosphorus data in catchment science.

2 Methodology

2.1 Study sites

Three rural catchments with different temperate climate, topography and farm types were monitored at high-temporal resolution as part of the UK Demonstration Test Catchment (DTC) programme (Lloyd et al., 2016a; Lloyd et al., 2016b;

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-314
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 6 June 2017

Outram et al., 2014; McGonigle et al., 2014). These were: Newby Beck at Newby, Eden catchment, Cumbria (54.59° N, 2.62° W; 12.5 km^2); Blackwater at Park Farm, Wensum catchment, Norfolk (52.78° N, 1.15° E; 19.7 km^2); Wylye at Brixton Deverill, Avon catchment, Hampshire (51.16° N, 2.19° W; 50.2 km^2) (Fig. 1). Further details of these catchments are available in SI Table S1.

2.2 Data collection

Rainfall was measured at 15 minute resolution at three sites in each of the Newby Beck and Blackwater catchments (Outram et al., 2014; Perks et al., 2015) and summed to give hourly totals. The hourly totals from the different rain gauges were combined by areal weighting to give an hourly time series for the catchment. For the Wylye catchment, only daily rainfall was available for sites within the catchment, so raw tipping bucket data were obtained for several sites outside the catchment and analysed to produce an hourly time series which was considered most representative of the rainfall in the catchment. Further details of the rainfall analysis for the Wylye catchment are given in SI Section S1.

River water level was measured at 15 minute resolution in the three catchments, with rating curves developed for discharge estimation (Outram et al., 2014; Perks et al., 2015; Lloyd et al., 2016b). Total phosphorus (TP) concentration was determined in-situ at 30 minute intervals with a Hach Lange combined Sigmatax sampling module and Phosphax analyser using acid digestion and colorimetry (Jordan et al., 2007; Jordan et al., 2013; Perks et al., 2015). Total P loads for each hour were determined by multiplying discharge (averaged to 30 minute resolution) by TP concentration for each 30 minutes and summing to give hourly totals:

$$TPload(t) = k \sum_j Q_j C_j \qquad (1)$$

where *TPload(t)* is the load (kg) exported during the hourly timestep which ends at time t, Q_j are the discharge observations (m^{3}s^{-1}) within the hourly timestep, C_j are the corresponding TP concentration observations (mg L^{-1}) within the hourly timestep, and k is a constant (= 3.6) for conversion of units to give load in kg. Visual inspection of the data indicated that aggregation of the data from 15 or 30 minute resolution to hourly did not result in a significant loss of information. This would not be the case for very small catchments or those where the dynamics being investigated were very fast. Calculation of the load according to Eq. 1 assumes that the TP is well-mixed in the water and that the Hach Lange sampler is taking a representative sample. It also assumes that the rating curve is appropriate over the full range of stage recordings made, and that the relationship between stage and discharge is stationary. Total phosphorus load, rather than concentration, was modelled because water utility companies are concerned about the total load which may have to be removed and because both water flow and load are fluxes, so comparisons between the two are easier to interpret directly than for concentration, which is a state rather than a flux (Jones et al., 2014).

2.3 Transfer function model identification

Transfer function models relating the input (here, a time series of rainfall, *R*) to the output (here, a time series of either discharge, *Q*, or phosphorus load, *TPload*) were identified using continuous-time models (Young and Garnier, 2006) where possible, or in cases where data were missing or identification was difficult, with discrete time models (Young, 2003), the estimation of which handles missing data more robustly. A full description of continuous-time and discrete-time model structures is given in Ockenden el al. (in press) and repeated here, in part, for clarity. The parameter estimates in both continuous-time models and discrete-time models are formulaically related (SI Table S3).

In continuous-time, a transfer function model with time delay τ has the form:

$$Y(s) = \frac{B(s)}{A(s)} e^{-s\tau} U(s) + E(s) \quad (2)$$

where *Y(s)*, *U(s)* and *E(s)* represent the Laplace transforms of the output, input and noise, respectively. *A(s)* and *B(s)* represent the denominator and numerator polynomials in the derivative operator $s = \frac{\mathrm{d}}{\mathrm{d}t}$ that define the relationship between the input and the output, and τ represents the time delay. In this study, models up to second order and without a noise model were considered, denoted by the triad [n, m, τ], where n and m denote the order of the denominator and numerator polynomials. Second order models were only accepted if they could be decomposed by partial fraction expansion into two parallel, first-order transfer functions, i.e.

$$TPload = \frac{b_\mathrm{f}}{s+a_\mathrm{f}} e^{-s\tau} R + \frac{b_\mathrm{s}}{s+a_\mathrm{s}} e^{-s\tau} R \quad (3)$$

This can be interpreted as two parallel stores, which are depleted at different rates, determined by the time constants (direct reciprocals of a_f and a_s) of the fast and slow components of the response, respectively. b_f and b_s are parameters that determine the gain of the fast and slow components, respectively. The terms 'fast' and 'slow' are used here as qualitative terms, since they are not necessarily related to specific process mechanisms; for a second order model (two stores), one store simply depletes at a slower rate than the other. Time constants are catchment specific; for example, for a first order rainfall-runoff model which identifies just the dominant mode (one pathway), the time constant can vary from less than an hour (e.g. for a small, flashy catchment in Malaysian Borneo (Chappell et al., 2006)) to more than three months (e.g. for a chalk stream in Berkshire, UK (Ockenden and Chappell, 2011)). Models were identified using the *RIVCBJ identification* algorithm (Refined Instrumental Variable Continuous-time Box-Jenkins identification, for continuous-time models), or *RIVBJ identification* (Refined Instrumental Variable Box-Jenkins identification for discrete-time models) that are part of the CAPTAIN toolbox (Taylor et al., 2007) for MATLAB®. The algorithms allow unbiased estimates of the model parameters and their covariance matrices. Monte Carlo sampling within the parameter space determined by the covariance matrices allows for uncertainty in derived quantities, such as time constants, to be calculated. Prediction bounds for the model can be calculated by adding the residual uncertainty and the parameter uncertainty. Note that in the case of transfer function models of the hydrograph, the models do not directly reflect the transport of water in the system since the hydrograph represents the integrated effects of celerities in the system rather than flow velocities (McDonnell and Beven, 2014).

This method of model identification requires high-temporal-resolution data that capture the dynamic response to the driving input; therefore, it cannot work if input data (in this case, rainfall) are missing, and does not perform well if too much output data (in this case, discharge or TPload) are missing or not showing a response. For the Newby Beck catchment, linear models were identified for short storm sequences up to one month, and were considered applicable to periods of similar conditions. Models were not identified for short periods for Blackwater and Wylye, as the presence of a much slower pathway (with a time constant of the same order as the length of the identification period) did not allow model parameter estimates to be sufficiently constrained over such short periods.

For longer time series, when seasonal change and antecedent wetness are expected to have an impact on the response, linear models were improved by inclusion of the rainfall-runoff non-linearity (Beven, 2012) based on the storage state of the catchment, for which the discharge is used as a proxy, i.e.

$$Re(t) = R(t)(Q(t-1))^{\beta} \tag{4}$$

where *Re(t)* is the effective rainfall at time *t*, *R* is the observed rainfall, *Q* is the observed discharge and *β* is a constant exponent that is optimized from the observed data at the same time as model identification. For annual TP loads, the models (still with hourly timestep) were identified based on the data for hydrological years 2011/12 and 2012/13 for Newby Beck, but, because of missing output data, just for hydrological year 2012/13 for the Blackwater and Wylye catchments. Models were validated on the data for all, or part, of the hydrological year 2013/14.

Model fit was assessed according to model bias, to evaluate systematic over- or under-prediction of the model, and to R_t^2 (also known as Nash Sutcliffe Efficiency, NSE):

$$R_t^2 = 1 - \frac{\hat{\sigma}^2}{\sigma_y^2} \tag{5}$$

where $\hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N}[y_i - \hat{y}]^2$; $\sigma_y^2 = \frac{1}{N}\sum_{i=1}^{N}[y_i - \bar{y}]^2$; $\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$ (6)

$\hat{y}$ is the model simulation; $\hat{\sigma}^2$ is the variance of the model residuals and σ_y^2 is the variance of the observations, y_i. A balance of model fit and over-parameterisation was sought using the Young Information Criterion (YIC) and visual inspection of the model fit to the monitoring data. Model assessment criteria are defined in SI Section S2.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-314
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 6 June 2017

3 Results and Discussion

3.1 Observed hydrological response and total phosphorus load in the three catchments

Time series data from each catchment (Fig. 2) indicated large contrasts in the hydrological response of each study catchment, with Newby Beck (Eden) showing a very flashy response to rainfall (Fig. 2a). Although a fast response at certain times was also evident in the Blackwater (Wensum) catchment (Fig. 2c) and the Wylye (Avon) catchment (Fig. 2e), there was also a more pronounced seasonal response, particularly in the Wylye where a large groundwater component could be observed in the winter periods. TP concentrations in all three study catchments revealed peaks that corresponded with runoff, with maximum values of 1.0 mg L^{-1}, 0.9 mg L^{-1} and 1.5 mg L^{-1} in the Newby Beck, Blackwater and Wylye catchments, respectively. Newby Beck showed a very low background concentration of TP at low flow (minimum < 0.01 mg L^{-1}), compared to 0.05 – 0.1 mg L^{-1} in the Blackwater, and around 0.12 mg L^{-1} in the Wylye. The presence of a measurable, background, non-rainfall dependent concentration suggests an additional source of phosphorus to the recently applied agricultural sources. Such non-rainfall dependent sources include legacy stores of agricultural P in the soil, both large and smaller point source discharges, such as sewage treatment works and domestic septic tanks (Zhang et al., 2014), and groundwater, specifically contributions from mineral sources in the Upper Greensand geology of the Hampshire Avon(Allen et al., 2014).

A summary of the observed total rainfall, runoff, mean concentration and TP load is given in Table for the period 1 October 2012 – 30 September 2013 (the hydrological year with the most complete dataset). The highest runoff (per unit area) was observed in the Newby Beck catchment. The annual TP load was lowest in the Blackwater catchment, in spite of this catchment being larger than the Newby Beck catchment.

Investigation of the relationships between TP concentration and streamflow indicated that, for all three catchments, the TP concentration was out of phase with the streamflow; distinct hysteresis loops (SI Figs S1 – S3), also observed by Outram et al. (2014), showed different TP concentrations on the rising stage of a storm hydrograph compared to the same stage on the falling hydrograph. In order to capture the effects of storage, dynamic models are required.

3.2 Identification of linear transfer function models for short storm sequences

For short storm sequences up to about a month, when antecedent flows for events were rather similar, linear models were identified for the Newby Beck catchment. These were useful for infilling missing discharge or TP load data, or for highlighting and estimating uncertainties in discharge and TP load when extrapolation of the stage-discharge relationship was inappropriate. In contrast, whilst it was still possible to identify linear models for short periods for the Blackwater and Wylye catchments, the parameter uncertainty for these models was large; the parameters were not well constrained when the (slow) time constant was of similar order to the period of identification. For this reason, linear models for short periods for the Blackwater and the Wylye were not considered useful.

Table 1 shows results from rainfall-runoff and rainfall-TP load models identified for Newby Beck for a series of contiguous storms in November 2015, immediately preceding Storm Desmond (5 – 6 December 2015), which caused catastrophic flooding in Cumbria and Lancashire, UK. During Storm Desmond, Honister Pass in Cumbria received the highest 24 h rainfall on record (341 mm) and Thirlmere received the highest 48 h rainfall on record (405 mm). The storm was remarkable for the duration of sustained rainfall. At Newby Beck, 156 mm of rainfall was recorded in 36 h. Although the monitoring equipment was recording during Storm Desmond, the peak flows during the storm were out of bank for around 31 h (compared to less than 3.5 h during more typical storms), with anecdotal evidence that the gauging point was significantly bypassed, so these out of bank flows were highly uncertain. Concentrations were assumed to be reasonably accurate, but TP loads were underestimated due to the underestimate of discharge. Storm Desmond was not included in the model identification period. Using the models from the November period to simulate flows and TP load during Storm Desmond (Fig. 3) suggests that both discharge and TP load were underestimated.

Although there are uncertainties associated with whether it is valid to extend the models identified above to an extreme event such as Storm Desmond, we believe that this highlights the possible underestimation in discharge and TP load during Storm Desmond and that the models in Table 2 make more realistic estimations of the true values.

3.3 Identification of transfer function models on annual time series data

Longer term models, based on two years of hourly data, were identified for each catchment. Model fits (R_t^2) for rainfall-runoff models for the identification period (Table 3) were 0.71 for Newby Beck and 0.87 for Wylye, but only 0.37 for the Blackwater. Model bias was less than ± 10% for all three catchments. The non-linearity, which reflects the effect of the antecedent soil moisture conditions in the catchments, was accounted for with the soil moisture surrogate expressed as a power function of discharge (Beven, 2012) with exponent β in Eq. 4, where a value of zero produces a linear response to rainfall and a higher value leads to an increasingly non-linear response. The β values identified for Newby Beck, Blackwater and Wylye were 0.37, 0.65 and 0.59, respectively, indicating the most non-linear response was in the Wensum (Blackwater) catchment, which also gave the lowest model efficiency values. The best identified model for rainfall-runoff in each catchment was a second-order model. In general, models higher than second order gave little improvement in model fit but a large deterioration in YIC, signifying over-parameterisation not warranted by the information in the monitoring data, whereas first order models often gave a reasonable fit to the model peaks (and hence reasonable R_t^2), but poor fit to recession periods.

The dynamic response characteristics of time constant and percentage on each flow pathway (for definitions see SI Table S4), determined after partial fraction decomposition, can be compared between the study catchments for both discrete and continuous time models. The time constants are associated with the dominant pathways and indicate how quickly each impulse response (of water or TP mass) is depleted to 37% (or fraction 1/e) of the peak exported. This measure of the time taken for a fixed proportion of the response to have occurred allows for a piston effect in the movement of water through a catchment and is not the same as tracking individual particles through the catchment. In reality there will be a continuum of runoff pathways with different time constants, but the information in the data indicates that this continuum can be simplified by representation as just two dominant pathways.

The marginal distributions of the time constants and proportion of flow or TP load (Table 3) were determined from 10,000 Monte Carlo realisations using the covariance of the parameter estimates. The parameter uncertainties were small, even for the response characteristics of the TP load models, which had higher uncertainty than rainfall-runoff models; TP load models had coefficients of variation of less than 3% for fast time constants, less than 6% for slow time constants and less than 2% for proportions on pathways. For the rainfall-runoff models, the time constant for the fast pathway was 2.9 h ± 0.1 h for Newby Beck, with 43% ± 0.5% of the water taking this pathway; in the Wylye, the time constant for the fast pathway was 4.1 h ± 0.2 h, but with only 8% ± 0.2% of the water taking this route. This is consistent with the much higher baseflow index in the Hampshire Avon (0.93) than the Eden (0.39) (SI Table S1), which is clearly visible in the data (Fig. 1). For the Blackwater, 25% ± 0.6% of the flow took the fast pathway, which is also consistent with the baseflow index in the Wensum (0.8) being between the Eden and Hampshire Avon. The fast time constant for the Blackwater catchment was much slower, at 14.8 h ± 0.25 h; this may be related to the average slope of the catchment, which is much lower for the Blackwater catchment (less than 2%) compared to 6 – 8% for the Wylye and Newby Beck catchments. The slow time constant for Newby Beck was 147 h ± 5 h, with 57% ± 0.5% of flow taking this pathway; this compared with 441 ± 13 hours (75% ± 0.6 % of flow) for Blackwater and 395 ± 6 hours (92% ± 0.2% of flow) for Wylye.

3.4 Interpretation of TP load dynamics alongside runoff dynamics

For the rainfall-TP load models, at Newby Beck the best identified model was a first order model relating the effective rainfall (from the runoff model) to the TP load (Table 3, Fig. 4). Although it was possible to identify a second order model, this made virtually no difference to model fit, R_t^2, and at the expense of YIC (signifying over-parameterisation), and decomposition of the model revealed time constants for the two pathways that were both less than 8 hours (c.f. 147 hours for the slow pathway for the rainfall-runoff model in Table 3). This indicates that in Newby Beck, all the TP load is transported through a quickflow pathway. This is consistent with most of the load being associated with P mobilised from diffuse agricultural sources, which is transferred by surface runoff or shallow sub-surface flow. This includes particulate P transported in surface runoff or drain flow(Heathwaite et al., 2006), subsurface movement of fine particles and colloids (Heathwaite et al., 2005), and displacement of fast subsurface soluble P sources. Young (2010) recommended a minimum data sampling rate of one-sixth of the time constant, in order to avoid possible temporal aliasing effects. Littlewood and Croke (2013) illustrated the parameter inaccuracy and loss of data when observations were under-sampled for discrete time transfer functions, with inaccuracy decreasing and parameter estimates approaching stable values as the sampling interval decreased from 24 hours (daily sampling) down to hourly sampling. The time constant for the first-order TP load model for Newby Beck was 1.6 ± 0.04 hours. In this study, daily data would not capture the true dynamics of discharge and TP load, and that, ideally, for flashy catchments such as Newby Beck, a sampling interval shorter than hourly would be even more robust. However, for the other catchments in this study, the hourly data frequency was sufficient. The time constant for the TP load model (1.6 ± 0.04 h) was even faster than the fast time constant for the second-order (two pathway) rainfall-runoff model (2.9 h ± 0.1 h), indicating that the TP mass impulse response was depleted at a faster rate than the water response, i.e. that the store was diluted as the storms progressed or that the sources must be readily connected and closer to the stream, since TP depends on transport velocities and we would normally expect velocities to be less than celerities under wet and surface runoff conditions. Those source areas would also be the most readily exhausted so the effects would reinforce each other.

For the Wylye, the best identified TP load model was a second-order model relating effective rainfall to TP load, with 42% ± 1% on a fast pathway (TC = 6.1 ± 0.3 hours) and 58 ± 1% on a slower pathway (570 ± 54 hours) (Table 3, Fig. 5). Compared to the runoff model, this showed a much greater percentage of the TP load on faster pathways such as surface runoff, shallow sub-surface flow or sub-surface drains. Nevertheless, there was still a significant proportion travelling on a slower pathway, which highlights the need for pollution mitigation efforts to include measures that take account of sub-surface and groundwater flows, and also, to recognise that surface runoff from farmland is not the only source of nutrients and sediment (Allen et al., 2014;Evans, 2012).

The TP load model used for the Blackwater was a linear model relating rainfall directly to TP load. The second-order TP model gave fast and slow time constants of 12.5 ± 0.6 hours and 376 ± 44 hours, respectively (Table 3, Fig. 6). The time constants were similar in magnitude, though both slightly shorter, to the time constants for the runoff model, suggesting a possible exhaustion effect where, as in Newby Beck, the TP mass store was diluted as the response progressed. For the Blackwater, as in the other study catchments, the proportion of TP load transferred on the fast pathway (54 ± 2%) was considerably more than the proportion of water on the fast pathway (25% ± 0.6%).

The proportion of TP load exported on the fast pathway was considerably greater for all catchments than the corresponding proportion of water on the fast pathway, by a factor of approximately two for Newby Beck and Blackwater and approximately five for the Wylye. This suggests that on the fast water pathways, generally associated with shallower pathways such as shallow sub-surface flow, field drains and surface runoff, there is more release of TP than on deeper water pathways. This is consistent with soil profiles in agricultural areas, which generally show P concentrated on the surface and in the near-surface soil layers, with a decrease in P with depth (Heathwaite and Dils, 2000).

Validation of the TP model for Blackwater and Wylye was performed on a shorter period than for Newby Beck (half of the hydrological year 2013/14) because of missing data. Although seasonal non-linearity was still evident in the data from Blackwater, none of the rainfall-runoff models that included the non-linearity validated very well, such that TP models using the effective rainfall simulated by the rainfall-runoff model gave a worse fit to the data than a simple linear model. This may be due to missing data in the discharge and TP time series, particularly over the storm peaks. Alternatively, the power law used to represent the rainfall-runoff non-linearity might not perform very well in the Blackwater catchment because of the large slow baseflow component. Different representations of the rainfall-runoff linearity were also investigated, such as the Bedford Ouse Sub-Model (Chappell et al., 2006), in which the soil storage is related to an antecedent precipitation index. Although changes in the model non-linearity representation made minor differences to model fit, none of the model variants validated well for the Blackwater catchment. This suggests that there may be a different mechanism at work in the Blackwater catchment, in which a fast pathway only becomes active once the soil is fully saturated, or the groundwater level rises to a certain level (Outram et al., 2016). This could be due to the shallow slopes, which encourage infiltration rather than runoff. Alternatively, the response may be more dominated by point sources which are not as rainfall-driven, or sources such as sediment-laden runoff from impervious surfaces (roads/yards), which are rainfall-driven but do not behave in the same non-linear way as the runoff from soil.

In addition, the conditions experienced during the two years used for model identification may not be very similar to the validation period. From the data in Fig. 1c, the winter of 2011 and spring of 2012 showed much lower discharge than the same months in subsequent years. The groundwater recharge, which is shown as in increase in the baseflow in winter, was obvious for winter 2012/13 and winter 2013/14 for both the Blackwater (Fig. 2c) and the Wylye (Fig. 2e), but was not evident for either catchment for the winter of 2011/12. Because of the slow time constants for these catchments, the dataset for model identification needs ideally to be longer than for the Newby Beck catchment, where the dynamics are much faster. This study suggests that the dataset used here was not long enough for the Blackwater catchment to capture an adequate range of conditions.

3.5 Advantages and limitations of the modelling method

The benefits and limitations of the modelling method for TP load are summarised in Table 4. For catchments that exhibit rapidly changing dynamics, such as response to storm events, models calibrated with daily data will have large uncertainties associated with the parameters (and output) because the input data do not capture the high frequency dynamics of processes such as P transfer. This study shows that simple transfer function models using data with sub-daily resolution can simulate the dynamics of TP load, with model fits at least as good as generally achieved with process-based models (Gassman et al., 2007; Moriasi et al., 2007) and with low parameter uncertainty. It is still advisable to validate a fitted model using at least a split record test (KlemeŠ, 1986). This highlights the importance of long and complete datasets with good time resolution for properly representing both flow and TP loads for such catchments. This requirement for adequate datasets is often an obstacle in the use of the DBM modelling method, but as such datasets become more available, the method can be used to improve our understanding of catchments.

The models in Table 3have been identified using a consistent method, as far as possible, to test how well this modelling method copes with the different characteristics of the three catchments. The method has been successfully applied to all the catchments, although less successfully for the Blackwater catchment. It is likely that the models could be improved if catchment-specific adjustments were made or used alongside other models in a hypothetico-inductive manner (Young, 2013). For instance, in the Blackwater catchment, the use of state dependent parameters (Young, 1984) might be more successful to capture the rainfall-runoff non-linearity. This means that, rather than using the form of the non-linearity specified by Eq. 4, the parameters could be allowed to vary according to some other observed state. Models for all catchments could be improved by having a longer dataset, to ensure, as far as possible, that environmental conditions during a future simulation period have already been experienced during the identification period.

For process-based models, the inclusion of the process representations is intended to account for response to changing environmental conditions. This is the basis for arguing that process-based models are better suited for predicting the impacts of future change. However, they also involve a plethora of (often difficult to validate) assumptions in their model structures and parameters. In practice, parameters set during calibration are rarely changed to account for changes in the modelled processes under future conditions, although by calibrating models for conditions similar to the expected future conditions, it may be possible to incorporate non-stationary parameter values (Nijzink et al., 2016). This idea could be integrated into DBM models by choosing identification periods which are most likely to reflect the conditions of the simulation period or through the use of state-dependent parameters. Thus, whilst the data-based assumption of similar conditions may be questioned when limited periods have been used for identification, usually restricted by data availability, we argue that many of the factors contributing to catchment response will not have changed (e.g. catchment topography, soil type and geology) and that this assumption will in many circumstances be no more restrictive than the (different) assumptions made when using process-based models. Clearly, where the factors contributing to catchment response have obviously changed (such as if all septic tanks were upgraded or if farm budgeting reduced the additions of P), then simple transfer function models would not be able to predict the changes over time, whereas, in theory, process-based models might be able to account for such changes, albeit with much uncertainty, (e.g. Dean et al., 2009;Yang et al., 2008). However, for rainfall dominated responses, or responses to changes in rainfall patterns, simple transfer function models can provide valuable understanding of the dominant modes of a catchment, which, in turn, can be used to target management interventions.

4 Summary and Conclusions

High temporal resolution data (hourly) of discharge and TP load have been used to identify simple transfer function models that capture the dynamics of rainfall-runoff and rainfall-phosphorus load in three diverse agricultural catchments. Linear models were identified for short storm sequences in the flashy Newby Beck catchment, when antecedent flows for events were similar. Models identified for November 2015 were used to simulate flows and TP loads in the devastating Storm Desmond (5-6 December 2015), supporting our belief that the discharge and TP load calculated from recorded data during this storm were considerably underestimated. In these circumstances, simple models could be useful to infill missing data or to highlight or estimate uncertainties in the recorded data. Linear models for short periods were not appropriate for the less flashy Blackwater and Wylye catchments when the slow time constant (for a second order model) was similar in length to the time period of identification, making the parameter uncertainty large.

Longer-term models were identified for each of the three catchments on two years of data. Comparison of rainfall-runoff and rainfall-TP load models for each catchment allowed a better understanding of the dominant modes of transport within each catchment, which was based on the times series data alone, rather than other (unmeasured) catchment parameters. In all three catchments, a higher proportion of the TP load was exported via a fast pathway than the corresponding proportion of water on the fast pathway. In agreement with soil profiles in agricultural areas, this suggested that there is more release of TP on fast (generally shallower) water pathways such as shallow sub-surface flow, field drains and surface runoff.

For successful simulations of future conditions, the models require long datasets to ensure that a full range of driving conditions has been included in the identification period. However, this study shows that simple transfer function models can be successful in modelling TP loads and explaining dominant transport modes. Transfer function models make good use of high frequency data, require very few parameters with low uncertainty and allow physical interpretation based solely on the information in the data.

Data availability

The data used in this study are openly available from Lancaster University data archive at https://dx.doi.org/10.17635/lancaster/researchdata/ (reserved until publication).

Supporting Information

Estimation of hourly rainfall time series for the Wylye catchment (Section S1); Model assessment criteria (Section S2); Study catchment characteristics (Table S1); Notation (Table S2); Structure of models and relationship between parameters from discrete-time and continuous-time models (Table S3); Definition of time constants, steady state gains and fraction on each pathway for discrete-time and continuous-time models (Table S4); Model structures and parameters identified (Table S5); Hourly streamflow against total phosphorus concentration for the Newby Beck catchment (Fig. S1), the Blackwater catchment (Fig. S2) and the Wylye catchment (Fig. S3). This material is available online.

Author Contributions

M.C.O. ran the DBM model and led the writing of the paper. W.T. assisted with DBM modelling. P.M.H was overall project lead with K.J.B., P.W., and J.Z. also helping manage the project. All authors participated in interpretation of results and the writing and editing process.

Competing interests

Jim Freer is a member of the editorial board of Hydrology and Earth System Sciences.

Acknowledgements

This work was funded by the Natural Environment Research Council (NERC) as part of the NUTCAT 2050 project, grants NE/K002392/1, NE/K002430/1 and NE/K002406/1, and supported by the Joint UK BEIS/Defra Met Office Hadley Centre Climate Programme (GA01101). The authors are grateful to the UK Demonstration Test Catchment (DTC) research platform for provision of the field data (Defra projects WQ02010, WQ0211, WQ0212 and LM0304). The DTC data are available at http://www.environmentdata.org/dtc-archive-project/dtc-archive-project .

References

Allen, D. J., Darling, W. G., Davies, J., Newell, A. J., Gooddy, D. C., and Collins, A. L.: Groundwater conceptual models: implications for evaluating diffuse pollution mitigation measures, Q. J. Eng. Geol. Hydroge., 47, 65-80, 10.1144/qjegh2013-043, 2014.

Beven, K. J.: Rainfall-runoff modelling : the primer, 2nd edition, John Wiley & Sons, Chichester, 2012.

Bieroza, M. Z., and Heathwaite, A. L.: Seasonal variation in phosphorus concentration-discharge hysteresis inferred from high-frequency in situ monitoring, J. Hydrol., 524, 333-347, 10.1016/j.jhydrol.2015.02.036, 2015.

Bowes, M. J., Jarvie, H. P., Halliday, S. J., Skeffington, R. A., Wade, A. J., Loewenthal, M., Gozzard, E., Newman, J. R., and Palmer-Felgate, E. J.: Characterising phosphorus and nitrate inputs to a rural river using high-frequency concentration-flow relationships, Sci. Total Environ., 511, 608-620, 10.1016/j.scitotenv.2014.12.086, 2015.

Carpenter, S. R., and Bennett, E. M.: Reconsideration of the planetary boundary for phosphorus, Environ. Res. Letters, 6, 10.1088/1748-9326/6/1/014009, 2011.

Cassidy, R., and Jordan, P.: Limitations of instantaneous water quality sampling in surface-water catchments: Comparison with near-continuous phosphorus time-series data, J. Hydrol., 405, 182-193, 10.1016/j.jhydrol.2011.05.020, 2011.

Chappell, N. A., Tych, W., Chotai, A., Bidin, K., Sinunc, W., and Chiew, T. H.: BARUMODEL: Combined Data Based Mechanistic models of runoff response in a managed rainforest catchment, Forest Ecol. Manag., 224, 58-80, 2006.

Dean, S., Freer, J., Beven, K., Wade, A. J., and Butterfield, D.: Uncertainty assessment of a process-based integrated catchment model of phosphorus, Stoch. Env. Res. Risk A., 23, 991-1010, 10.1007/s00477-008-0273-z, 2009.

Evans, R.: Reconnaissance surveys to assess sources of diffuse pollution in rural catchments in East Anglia, eastern England - implications for policy, Water Environ. J., 26, 200-211, 10.1111/j.1747-6593.2011.00277.x, 2012.

Evans, R., and Boardman, J.: The new assessment of soil loss by water erosion in Europe. Panagos P. et al., 2015 Environmental Science & Policy 54, 438-447-A response, Environ. Sci. Policy, 58, 11-15, 10.1016/j.envsci.2015.12.013, 2016.

Gassman, P. W., Reyes, M. R., Green, C. H., and Arnold, J. G.: The soil and water assessment tool: Historical development, applications, and future research directions, T. ASABE, 50, 1211-1250, 2007.

Halliday, S. J., Skeffington, R. A., Wade, A. J., Bowes, M. J., Gozzard, E., Newman, J. R., Loewenthal, M., Palmer-Felgate, E. J., and Jarvie, H. P.: High-frequency water quality monitoring in an urban catchment: hydrochemical dynamics, primary production and implications for the Water Framework Directive, Hydrol. Process., 29, 3388-3407, 10.1002/hyp.10453, 2015.

Harmel, R. D., Cooper, R. J., Slade, R. M., Haney, R. L., and Arnold, J. G.: Cumulative uncertainty in measured streamflow and water quality data for small watersheds, T. ASABE, 49, 689-701, 2006.

Heathwaite, A. L., and Dils, R. M.: Characterising phosphorus loss in surface and subsurface hydrological pathways, Sci. Total Environ., 251, 523-538, 2000.

Heathwaite, A. L., Fraser, A. I., Johnes, P. J., Hutchins, M., Lord, E., and Butterfield, D.: The Phosphorus Indicators Tool: a simple model of diffuse P loss from agricultural land to water, Soil Use Manage., 19, 1-11, 2003.

Heathwaite, A. L., Burke, S. P., and Bolton, L.: Field drains as a route of rapid nutrient export from agricultural land receiving biosolids, Sci. Total Environ., 365, 33-46, 2006.

Heathwaite, L., Haygarth, P., Matthews, R., Preedy, N., and Butler, P.: Evaluating colloidal phosphorus delivery to surface waters from diffuse agricultural sources, J. Environ. Qual., 34, 287-298, 2005.

Jackson-Blake, L. A., Dunn, S. M., Helliwell, R. C., Skeffington, R. A., Stutter, M. I., and Wade, A. J.: How well can we model stream phosphorus concentrations in agricultural catchments?, Environ. Modell. Softw., 64, 31-46, 10.1016/j.envsoft.2014.11.002, 2015.

Jarvie, H. P., Withers, P. J. A., and Neal, C.: Review of robust measurement of phosphorus in river water: sampling, storage, fractionation and sensitivity, Hydrol. Earth Syst. Sc., 6, 113-131, 2002.

Johnes, P. J.: Evaluation and management of the impact of land use change on the nitrogen and phosphorus load delivered to surface waters: The export coefficient modelling approach, J. Hydrol., 183, 323-349, 1996.

Johnes, P. J.: Uncertainties in annual riverine phosphorus load estimation: Impact of load estimation methodology, sampling frequency, baseflow index and catchment population density, J. Hydrol., 332, 241-258, 10.1016/j.jhydrol.2006.07.006, 2007.

Jones, T. D., and Chappell, N. A.: Streamflow and hydrogen ion interrelationships identified using data-based mechanistic modelling of high frequency observations through contiguous storms, Hydrol. Res., 45, 868-892, 10.2166/nh.2014.155, 2014.

Jones, T. D., Chappell, N. A., and Tych, W.: First Dynamic Model of Dissolved Organic Carbon Derived Directly from High-Frequency Observations through Contiguous Storms, Environ. Sci. Technol., 48, 13289-13297, 10.1021/es503506m, 2014.

Jordan, P., Arnscheidt, A., McGrogan, H., and McCormick, S.: Characterising phosphorus transfers in rural catchments using a continuous bank-side analyser, Hydrol. Earth Syst. Sci., 11, 372-381, 2007.

Jordan, P., Cassidy, R., Macintosh, K. A., and Arnscheidt, J.: Field and laboratory tests of flow-proportional passive samplers for determining average phosphorus and nitrogen concentrations in rivers, Environ. Sci. Technol., 47, 2331-2338, 2013.

KlemeŠ, V.: Operational testing of hydrological simulation models, Hydrolog. Sci. J., 31, 13-24, 10.1080/02626668609491024, 1986.

Littlewood, I. G., and Croke, B. F. W.: Effects of data time-step on the accuracy of calibrated rainfall-streamflow model parameters: practical aspects of uncertainty reduction, Hydrol. Res., 44, 430-440, 10.2166/nh.2012.099, 2013.

Liu, S. M., Brazier, R., and Heathwaite, L.: An investigation into the inputs controlling predictions from a diffuse phosphorus loss model for the UK; the Phosphorus Indicators Tool (PIT), Sci. Total Environ., 344, 211-223, 2005.

Lloyd, C. E. M., Freer, J. E., Johnes, P. J., and Collins, A. L.: Using hysteresis analysis of high-resolution water quality monitoring data, including uncertainty, to infer controls on nutrient and sediment transfer in catchments, Sci. Total Environ., 543, 388-404, 10.1016/j.scitotenv.2015.11.028, 2016a.

Lloyd, C. E. M., Freer, J. E., Johnes, P. J., Coxon, G., and Collins, A. L.: Discharge and nutrient uncertainty: implications for nutrient flux estimation in small streams, Hydrol. Process., 30, 135-152, 10.1002/hyp.10574, 2016b.
McDonnell, J. J., and Beven, K.: Debates-The future of hydrological sciences: A (common) path forward? A call to action aimed at understanding velocities, celerities and residence time distributions of the headwater hydrograph, Water Resour. Res., 50, 5342-5350, 10.1002/2013wr015141, 2014.
McGonigle, D. F., Burke, S. P., Collins, A. L., Gartner, R., Haft, M. R., Harris, R. C., Haygarth, P. M., Hedges, M. C., Hiscock, K. M., and Lovett, A. A.: Developing Demonstration Test Catchments as a platform for transdisciplinary land management research in England and Wales, Environ. Sci. Process. Imp., 16, 1618-1628, 10.1039/c3em00658a, 2014.
McGuire, K. J., and McDonnell, J. J.: A review and evaluation of catchment transit time modeling, J. Hydrol., 330, 543-563, 2006.
McMillan, H., Krueger, T., and Freer, J.: Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality, Hydrol. Process., 26, 4078-4111, 10.1002/hyp.9384, 2012.
Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L.: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, T. ASABE, 50, 885-900, 2007.
Nijzink, R., Hutton, C., Pechlivanidis, I., Capell, R., Arheimer, B., Freer, J., Han, D., Wagener, T., McGuire, K., Savenije, H., and Hrachowitz, M.: The evolution of root-zone moisture capacities after deforestation: a step towards hydrological predictions under change?, Hydrol. Earth Syst. Sci., 20, 4775-4799, 10.5194/hess-20-4775-2016, 2016.
Ockenden, M. C., and Chappell, N. A.: Identification of the dominant runoff pathways from data-based mechanistic modelling of nested catchments in temperate UK, J. Hydrol., 402, 71-79, 10.1016/j.jhydrol.2011.03.001, 2011.
Ockenden, M. C., Hollaway, M. J., Beven, K., Collins, A. L., Evans, R., Falloon, P., Forber, K. J., Hiscock, K. M., Kahana, R., Macleod, C. J. A., Tych, W., Villamizar, M. L., Wearing, C., Withers, P. J. A., Zhou, J. G., Barker, P. A., Burke, S., Freer, J. E., Johnes, P., Snell, M. A., Surridge, B. W. J., and Haygarth, P. M.: Major agricultural changes required to mitigate phosphorus losses under climate change, Nat Commun, in press.
Outram, F. N., Lloyd, C. E. M., Jonczyk, J., Benskin, C. M. H., Grant, F., Perks, M. T., Deasy, C., Burke, S. P., Collins, A. L., Freer, J., Haygarth, P. M., Hiscock, K. M., Johnes, P. J., and Lovett, A. L.: High-frequency monitoring of nitrogen and phosphorus response in three rural catchments to the end of the 2011-2012 drought in England, Hydrol. Earth Syst. Sci., 18, 3429-3448, 10.5194/hess-18-3429-2014, 2014.
Outram, F. N., Cooper, R. J., Sunnenberg, G., Hiscock, K. M., and Lovett, A. A.: Antecedent conditions, hydrological connectivity and anthropogenic inputs: Factors affecting nitrate and phosphorus transfers to agricultural headwater streams, Sci. Total Environ., 545, 184-199, 10.1016/j.scitotenv.2015.12.025, 2016.
Parker, G. T., Droste, R. L., and Rennie, C. D.: Coupling model uncertainty for coupled rainfall/runoff and surface water quality models in river problems, Ecohydrology, 6, 845-851, 10.1002/eco.1308, 2013.
Perks, M. T., Owen, G. J., Benskin, C. M. H., Jonczyk, J., Deasy, C., Burke, S., Reaney, S. M., and Haygarth, P. M.: Dominant mechanisms for the delivery of fine sediment and phosphorus to fluvial networks draining grassland dominated headwater catchments, Sci. Total Environ., 523, 178-190, http://dx.doi.org/10.1016/j.scitotenv.2015.03.008, 2015.
Taylor, C. J., Pedregal, D. J., Young, P. C., and Tych, W.: Environmental time series analysis and forecasting with the Captain toolbox, Environ. Modell. Softw., 22, 797-814, 10.1016/j.envsoft.2006.03.002, 2007.
Yang, J., Reichert, P., Abbaspour, K. C., Xia, J., and Yang, H.: Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China, J. Hydrol., 358, 1-23, 10.1016/j.jhydrol.2008.05.012, 2008.
Young, P., Parkinson, S., and Lees, M.: Simplicity out of complexity in environmental modelling: Occam's razor revisited, J. Appl. Stat., 23, 165-210, 1996.
Young, P.: Data-based mechanistic modelling of environmental, ecological, economic and engineering systems, Environ. Modell. Softw., 13, 105-122, 1998.
Young, P.: Top-down and data-based mechanistic modelling of rainfall-flow dynamics at the catchment scale, Hydrol. Process., 17, 2195-2217, 10.1002/hyp.1328, 2003.
Young, P. C.: Recursive Estimation and Time-Series Analysis, Springer-Verlag, Berlin, 1984.
Young, P. C., and Beven, K. J.: Data-Based Mechanistic Modelling and the Rainfall-Flow Nonlinearity, Environmetrics, 5, 335-363, 1994.

Young, P. C., Chotai, A., and Beven, K. J.: Data-Based Mechanistic Modelling and the Simplification of Environmental Systems, in: Environmental Modelling: Finding Simplicity in Complexity, edited by: Wainwright, J., and Mulligan, M., John Wiley and Sons Ltd, 371-388, 2004.

Young, P. C., and Garnier, H.: Identification and estimation of continuous-time, data-based mechanistic (DBM) models for environmental systems, Environ. Modell. Softw., 21, 1055-1072, 10.1016/j.envsoft.2005.05.007, 2006.

Young, P. C.: The estimation of continuous-time rainfall-flow models for flood risk management, In: Role of Hydrology in Managing Consequences of a Changing Global Environment, 2010, 303-310,

Young, P. C.: Hypothetico-inductive data-based mechanistic modeling of hydrological systems, Water Resour. Res., 49, 915-935, 10.1002/wrcr.20068, 2013.

Zhang, Y., Collins, A. L., Murdoch, N., Lee, D., and Naden, P. S.: Cross sector contributions to river pollution in England and Wales: Updating waterbody scale information to support policy delivery for the Water Framework Directive, Environ. Sci. Policy, 42, 16-32, 10.1016/j.envsci.2014.04.010, 2014.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-314
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 6 June 2017

Table 1 Observed rainfall, discharge, total phosphorus (TP) concentration and load for the period 1 October 2012 -30 September 2013, for the three catchments

Catchment	**Total rainfall (mm)**	**Total runoff (mm)**	**% discharge data missing**	**Mean annual discharge (m^3 s^{-1})**	**Mean annual TPconc (mg L^{-1})**	**Total annual TPload (kg)**	**% TPload data missing**
Newby Beck Eden, Cumbria	1186	776	0.0	0.31	0.080	1577	19.7
Blackwater, Wensum, Norfolk	634	195	13.8	0.14	0.092	277	30.6
Wylye, Avon, Hampshire	850	273	0.3	0.44	0.149	1705	27.4

Table 2 Rainfall-runoff and rainfall-total phosphorus load (TP) models identified for Newby Beck during the period 7 November – 4 December 2015, with estimations of discharge and TP load during Storm Desmond (5/6 December 2015). CT linear = Continuous-time transfer function with linear rainfall input; R_t^2 = model efficiency measure (Eqn. 5); $TC_{fast/slow}$ = time constant for the fast/slow pathway; $\%_{fast/slow}$ = percentage of output taking the fast/slow pathway; Model bias = 100 * $\Sigma(y_i^{model}-y_i^{obs})/ \Sigma(y_i^{obs})$;

Model	**Model structure**	R_t^2	TC_{fast}	TC_{slow}	$\%_{fast}$	$\%_{slow}$	**Model bias %**	**Σobs during Desmond**	**Σmodel during Desmond**	**% diff**
Rainfall-runoff	CT linear [2, 2, 1]	0.91	3.6 ± 0.4	33 ± 8	55 ± 5	45 ± 5	0.7%	86.6	106.5	23%
Rainfall-TP load	CT linear [1, 1, 1]	0.74	2.7 ± 0.3		100		13%	196.5	273.6	39%

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-314
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 6 June 2017

Table 3 Structure, response characteristics and model fit statistics of rainfall-runoff and rainfall-TP load models for each catchment. Models were calibrated on all or part of hydrological years 2012 and 2013 and validated on all or part of hydrological year 2014. β = exponent in the power law used for rainfall-runoff non-linearity (Eqn. 4); R_t^2 = model efficiency measure (Eqn. 5); Model bias = 100 * $\Sigma(y_i^{model}-y_i^{obs})/\ \Sigma(y_i^{obs})$; $TC_{fast/slow}$ = time constant for the fast/slow pathway; $\%_{fast/slow}$ = percentage of output taking the fast/slow pathway;

Location	Time period (calib)	Model	Model structure	β	R_t^2 for calib (using Qobs)	R_t^2 for calib (using Qsim)	Model bias (calib) %	TC_{fast}	TC_{slow}	$\%_{fast}$	$\%_{slow}$	Time period (valid)	R_t^2 for valid (using Qsim)	Model bias (valid) %
Newby	1.10.11 to 30.9.13	R - Q	CT [2, 2, 1]	0.37	0.86	0.71	-9.7	2.9 ± 0.1	147 ± 5	43 ± 0.5	57 ± 0.5	1.10.13 to 30.9.14	0.78	-14.3
Newby	1.10.11 to 30.9.13	Re – TPload	CT [1, 1, 1]			0.65	2.3	1.6 ± 0.04		100		1.10.13 to 30.9.14	0.62	5.1
Blackwater	1.12.11 to 31.8.13	R - Q	DT [2, 2, 6]	0.65	0.82	0.37	-1.5	14.8 ± 0.5	441 ± 13	25 ± 0.6	75 ± 0.6	1.10.13 to 30.9.14	0.32	-9.4
Blackwater	26.10.12 to 28.7.13	R - TPload	CT [2, 2, 4]			0.62	5.4	12.5 ± 0.6	376 ± 44	54 ± 2	46 ± 2	1.10.13 to 31.3.14	0.31	38.2
Wylye	1.10.12 to 30.9.13	R – Q	DT [2, 2, 6]	0.59	0.94	0.87	3.0	4.1 ± 0.2	395 ± 6	8 ± 0.2	92 ± 0.2	1.12.13 to 20.5.14	0.79	11.0
Wylye	1.10.12 to 30.9.13	Re – TPload	CT [2, 2, 6]			0.55	5.5	6.1 ± 0.3	570 ± 54	42 ± 1	58 ± 1	1.12.13 to 31.3.14	0.50	-19.7

Table 4 Advantages and limitations of the DBM modelling method for rainfall-TP load

Advantages	Limitations
No prior assumption of model structure required	Requires complete, high temporal frequency datasets
Very few parameters required	Requires long datasets to cover a full range of driving conditions
Low parameter uncertainty	Models may not work well for future conditions if the range of conditions has not been included in the identification period
Makes good use of high frequency data	The power law to represent the rainfall-runoff non-linearity may not be the best representation for each catchment
Physical interpretation is made based only on the information in the data	Stationary DBM model will not capture time variable gains

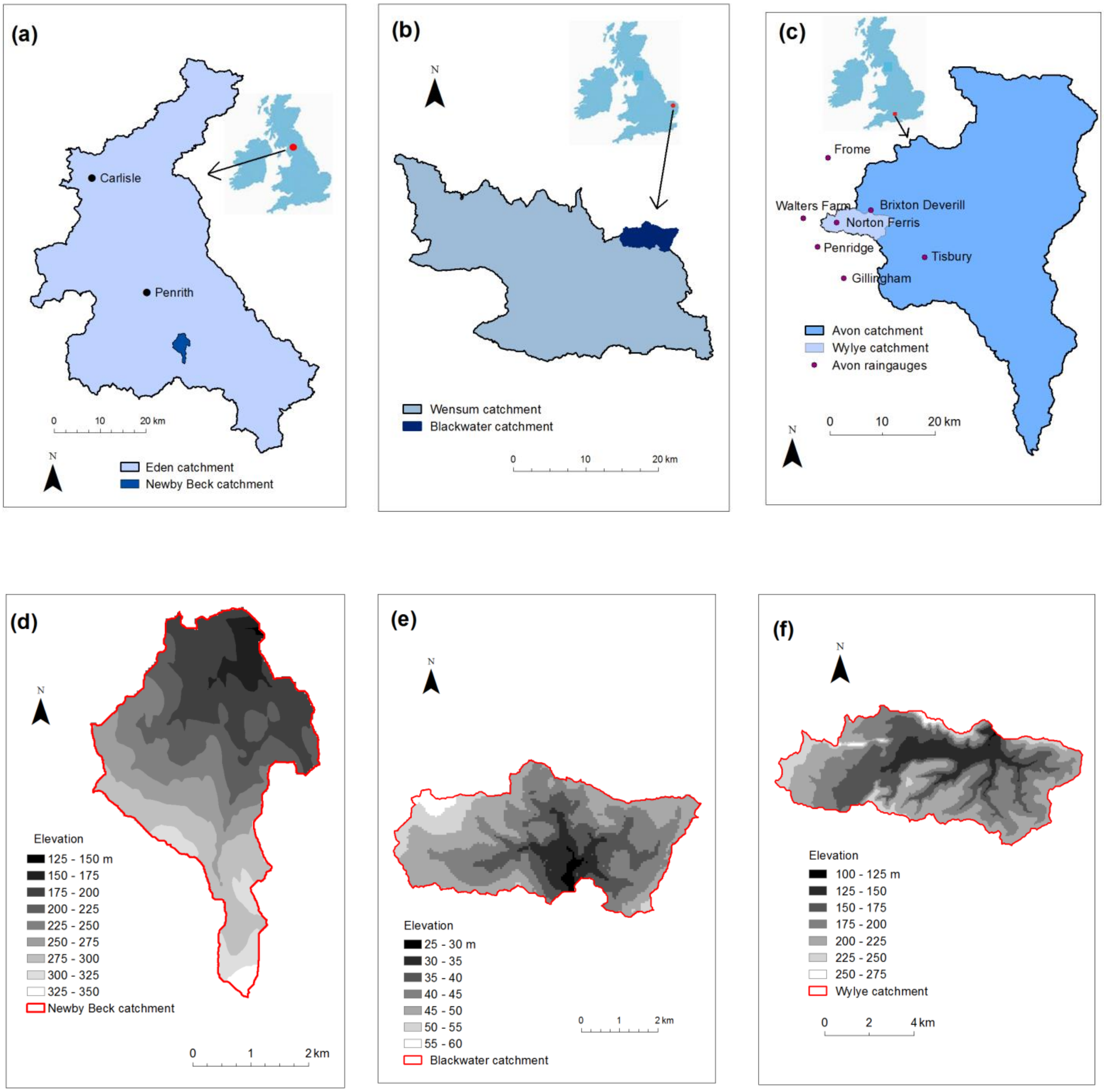

Figure 1 Location and topography of study catchments. Newby Beck, Eden, Cumbria: location (a) and topography(d); Blackwater, Wensum, Norfolk: location (b) and topography (e); Wylye, Avon, Hampshire: location (c) and topography(f). © OS Terrain 50 DTM [ASC geospatial data], Scale 1:50000, Tiles: ny51, ny52, ny61, ny62, Updated: July 2013; Tiles st73, st83, tg02, tg12, Updated: 2 August 2016; Ordnance Survey (GB), Using: EDINA Digimap Ordnance Survey Service, http://digimap.edina.ac.uk; Downloaded: 2017-01-03.

Hydrology and Earth System Sciences
Open Access
Discussions

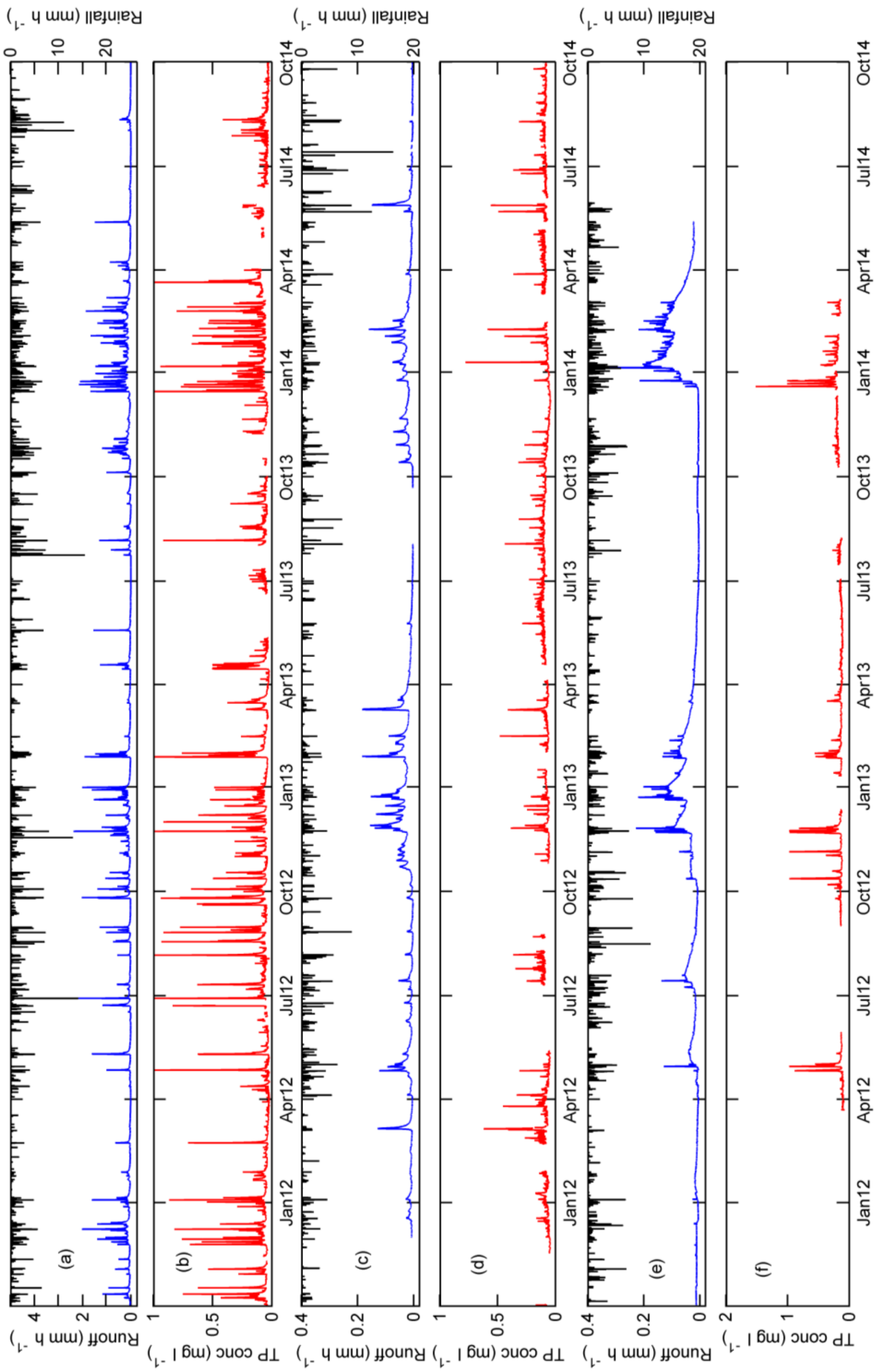

Figure 2 Time series of hourly rainfall, runoff and total phosphorus (TP) concentration at the three Demonstration Test Catchments; rainfall and runoff (a) and TP concentration (b) at Newby Beck, Eden; rainfall and runoff (c) and TP concentration (d) at Park Farm, Blackwater, Wensum; rainfall and runoff (e) and TP concentration (f) at Brixton Deverill, Wylye, Avon.

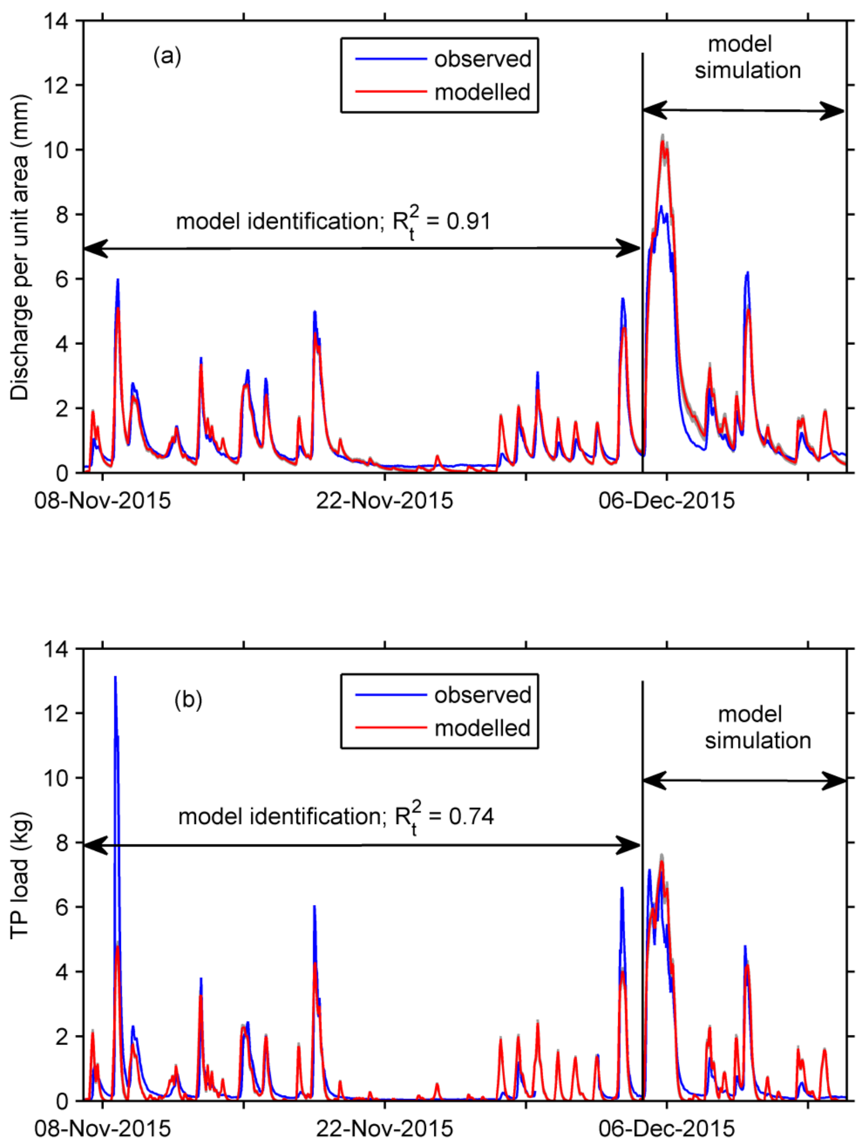

Figure 3 Observed and modelled discharge per unit area (a) and total phosphorus (TP) load (b) in Newby Beck, Eden during November 2015, with the same model used to estimate TP load during Storm Desmond 5/6th December 2015. The grey band indicates the 95% confidence limits on the model prediction due to parameter uncertainty.

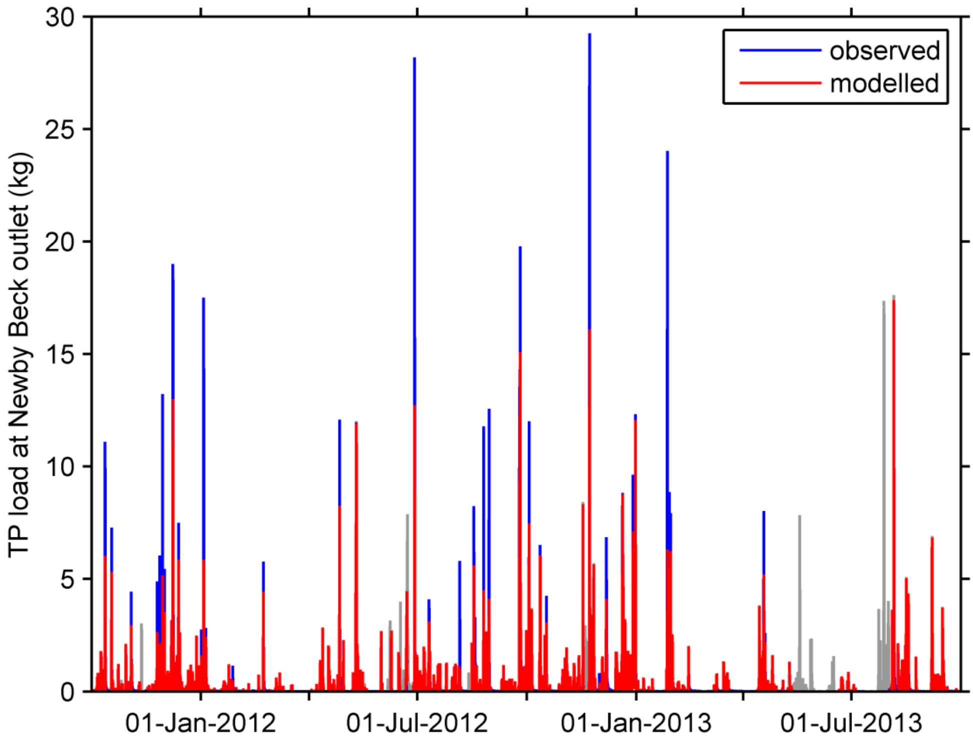

Figure 4 First-order model between effective rainfall and total phosphorus (TP) load at Newby Beck for the identification period 1 October 2011 – 30 September 2013. Continuous-time model with structure [1, 1, 1] (see Table 3); R_t^2 = 0.69. The grey band indicates the 95% confidence limits on the model prediction due to parameter uncertainty (at this scale, only visible during periods where TP data are missing).

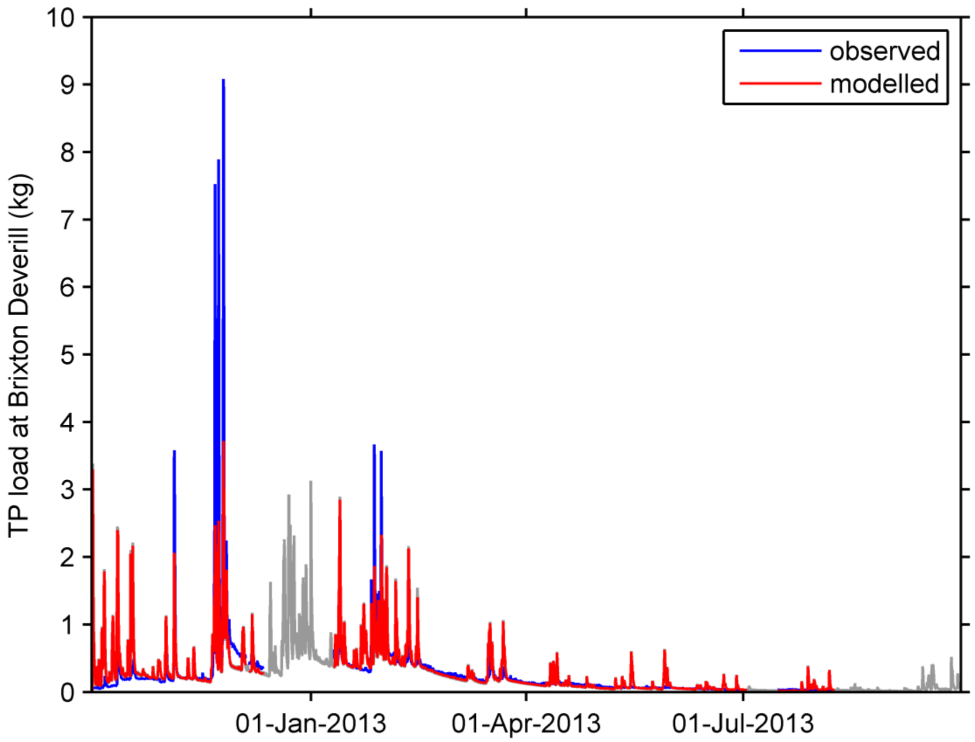

Figure 5 Second-order model between effective rainfall and total phosphorus (TP) load at Wylye for the identification period 1 October 2012 – 30 September 2013. Continuous-time model with structure [2, 2, 6] (see Table 3); R_t^2 = 0.67. The grey band indicates the 95% confidence limits on the model prediction due to parameter uncertainty (at this scale, only visible during periods where TP data are missing).

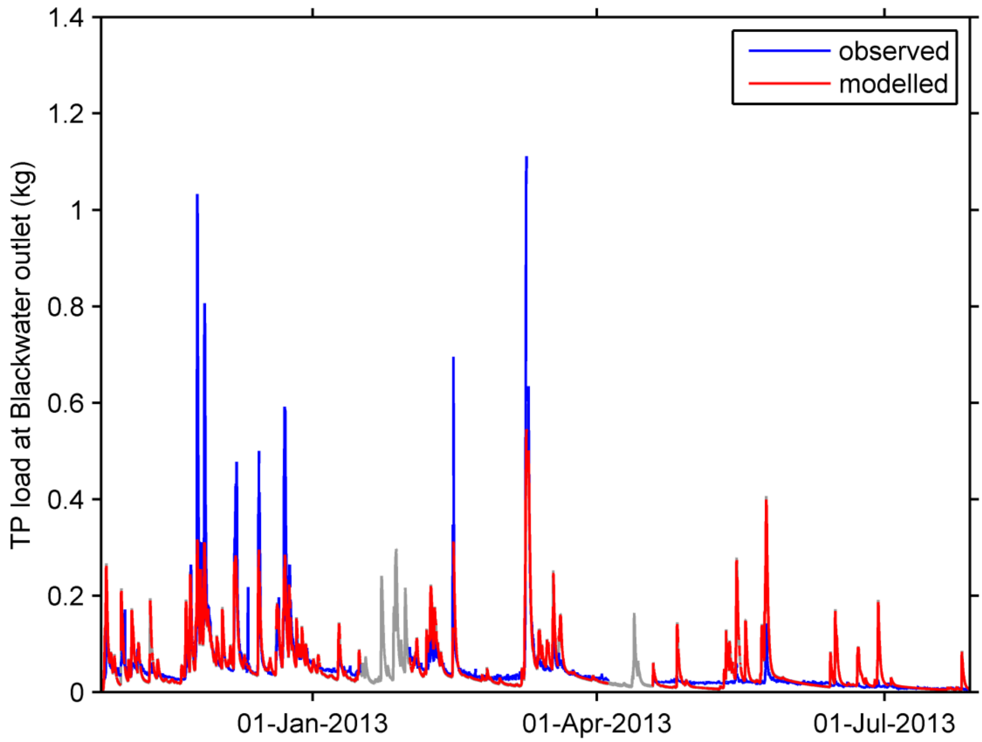

Figure 6 Second-order model between rainfall and total phosphorus (TP) load at Blackwater for the identification period 26 October 2012 – 28 July 2013. Continuous-time model with structure [2, 2, 4] (see Table 3); R_t^2 = 0.67. The grey band indicates the 95% confidence limits on the model prediction due to parameter uncertainty (at this scale, only visible during periods where TP data are missing).