Hydrol. Earth Syst. Sci., 18, 855–873, 2014
www.hydrol-earth-syst-sci.net/18/855/2014/
doi:10.5194/hess-18-855-2014

Hydrology and Earth System Sciences
Open Access

PERSiST: a flexible rainfall-runoff modelling toolkit for use with the INCA family of models

M. N. Futter[1], **M. A. Erlandsson**[2], **D. Butterfield**[3], **P. G. Whitehead**[4], **S. K. Oni**[1], **and A. J. Wade**[2]

[1]Department of Aquatic Science and Assessment, Swedish University of Agricultural Sciences, Uppsala, Sweden
[2]School of Human and Environmental Science, University of Reading, Reading, UK
[3]ENMOSYS, P.O. Box 2800, Glen Ellyn, IL 60138, USA
[4]School of Geography and the Environment, University of Oxford, Oxford, UK

Correspondence to: M. N. Futter (martyn.futter@slu.se)

Received: 24 June 2013 – Published in Hydrol. Earth Syst. Sci. Discuss.: 3 July 2013
Revised: 31 December 2013 – Accepted: 6 January 2014 – Published: 28 February 2014

Abstract. Runoff generation processes and pathways vary widely between catchments. Credible simulations of solute and pollutant transport in surface waters are dependent on models which facilitate appropriate, catchment-specific representations of perceptual models of the runoff generation process. Here, we present a flexible, semi-distributed landscape-scale rainfall-runoff modelling toolkit suitable for simulating a broad range of user-specified perceptual models of runoff generation and stream flow occurring in different climatic regions and landscape types. PERSiST (the Precipitation, Evapotranspiration and Runoff Simulator for Solute Transport) is designed for simulating present-day hydrology; projecting possible future effects of climate or land use change on runoff and catchment water storage; and generating hydrologic inputs for the Integrated Catchments (INCA) family of models. PERSiST has limited data requirements and is calibrated using observed time series of precipitation, air temperature and runoff at one or more points in a river network. Here, we apply PERSiST to the river Thames in the UK and describe a Monte Carlo tool for model calibration, sensitivity and uncertainty analysis.

1 Introduction

Understanding the fate and transport of pollutants in surface waters is dependent on credible simulations of water movement through the landscape. There are several approaches to simulating hillslope and catchment-scale water fluxes, depending on the purpose of the modelling exercise (Kampf and Burges, 2007). Models have been developed for flood flow forecasting (Bergström and Singh, 1995; Vehviläuinen, 2007, and references therein), as tools to better understand the hydrology of well-instrumented catchments (Fenicia et al., 2011; Kavetski and Fenicia, 2011; Hrachowitz et al., 2013b) and for simulating pollutant transport and transformations in catchments and surface waters (Whitehead et al., 1998; Wade et al., 2002; Andersson et al., 2005; Andersen et al., 2006; Lindström et al., 2010). The Integrated Catchments (INCA) family of models are widely used for simulating the behaviour of nitrogen (Whitehead et al., 1998; Wade et al., 2002), phosphorus (Crossman et al., 2013a, b), sediment (Lazar et al., 2010), dissolved organic carbon (Futter et al., 2007, 2009a) and a number of other solutes and pollutants in streams and rivers (Jin et al., 2011; Futter et al., 2012, and references therein). Here we present a new rainfall-runoff model, the Precipitation, Evapotranspiration and Runoff Simulator for Solute Transport (PERSiST) designed for use with the INCA family of models. PERSiST is a semi-distributed bucket-type modelling framework which allows model users to specify a perceptual model of the runoff generation process.

1.1 PERSiST and INCA

The intended audiences for PERSiST are scientists and catchment managers who use the INCA family of models to assess potential effects of climate and land-management change on surface water quality. PERSiST was developed primarily (but not exclusively) to address two challenges associated with the use of INCA. INCA uses external hydrologic time series inputs from a rainfall-runoff model, and the perceptual model of catchment hydrology in INCA is not

always appropriate outside the temperate and boreal ecoregions (Bernal et al., 2004; Medici et al., 2008, 2010).

INCA relies on external time series of hydrologically effective rainfall (HER; the fraction of precipitation which contributes to runoff) and soil moisture deficits (SMD; the difference between the current depth of water and the water-holding capacity). In the past, these time series have been obtained from rainfall-runoff models including the Meteorological Office Rainfall and Evapotranspiration Calculation System (MORECS; Hough and Jones 1997), the Hydrologiska Byråns Vattenbalansavdelning model (HBV, Bergström and SIngh, 1995; Sælthun 1996), the system for Identification of unit Hydrographs And Component flows from Rainfall, Evaporation and Streamflow data (IHACRES, Jakeman et al., 1990) and the Watershed Simulation and Forecasting System (WSFS, Vehviläuinen, 2007). There are conceptual and practical problems when using any of these models to generate time series inputs for INCA. The conceptual representation of water stores may differ between INCA and the rainfall-runoff models used to estimate SMD and HER. While it is possible to obtain credible time series of hydrologic inputs to INCA using any of the current generation of lumped rainfall-runoff models (including PERSiST), it is not entirely satisfactory to use one perceptual model of the runoff generation process for hydrological estimation and another for water chemistry simulations.

1.2 PERSiST and other models

There is a broad range of models available for catchment-scale runoff and solute simulations. The HYPE (Hydrological Predictions for the Environment) family of models (Arheimer et al., 2010; Lindström et al., 2010) is excellent for multi-catchment and regional predictions. PERSiST has been designed primarily for single catchment simulations where the routing of water through the landscape can be specified by the model user. While there is a great deal of research interest in Prediction in Ungauged Basins (PUB; Hrachowitz et al., 2013a, and references therein), PERSiST has been designed for use in flow simulations at gauged catchments. Thus, regionalization studies such as Hellebrand et al. (2011) have very different modelling goals than those motivating the development of PERSiST, which may be more useful for understanding the differences in hydrological response between apparently similar catchments such as those reported by Oni et al. (2013).

There may be markedly different perceptual representations of hydrologic response between mountainous and flat catchments. Because of its internal "steepest descent" flow routing algorithm, TAC-D (Uhlenbrook et al., 2004; Uhlenbrook and Sieber, 2005) is probably more suited to mountainous catchments or other regions of high relief. Using such an approach, it might be hard to obtain adequate hydrological gradients in the very flat agricultural catchments modelled by van der Velde et al. (2012). It would be informative to test PERSiST in catchments with very high and very low relief.

Like HBV-Light (Seibert and Vis, 2012), PERSiST has been designed to be easy to use with a graphical user interface that facilitates immediate feedback about the effect of parameter changes on simulated streamflow. However, the HBV-Light parameter estimation toolbox is more sophisticated than that presented here, and HBV-Light incorporates a fixed perceptual model of the runoff generation process.

A perceptual model of water fluxes in which the landscape is represented as one or more buckets that receive, store or transmit water has a long history, starting in the late 1960s with the Sacramento Soil Moisture Accounting model (Burnash et al., 1973). While this approach has been criticized recently (Gupta et al., 2012) it is still widely used for operational hydrology. Notably, HBV (Bergström and Singh, 1995) and WSFS (Vehviläuinen, 2007) are used for operational flood forecasting in Fennoscandia and IHACRES (Jakeman et al., 1990) and its variants for operational purposes in Australia. Simple bucket-type models such as HYMOD (Wagener et al., 2001) are also a key component of hydrologic research programmes investigating Bayesian techniques for model calibration (Vrugt et al., 2009)

The HBV-NP (Andersson et al., 2005), NAM-Mike 11 (Andersen et al., 2006), HYPE (Lindström et al., 2010) and INCA families of models (Whitehead et al., 1998; Wade et al., 2002) are all semi-distributed bucket-type solute transport models in which a single perceptual model is used to represent the movement of water through a catchment. While these models often work well, greater flexibility in representing the perceptual model of the runoff generation process is desirable. For example, Bernal et al. (2004) note that the perceptual model of runoff generation used in INCA, which is derived from observations of temperate catchments, is not ideal for simulating hydrochemistry in intermittent Mediterranean streams.

Because they have been developed for specific regions, spatial scales and climatic conditions, most currently used rainfall-runoff models can be difficult to apply successfully outside the conditions for which they were designed. There is an implicit "one-size-fits-all" approach in most rainfall-runoff models where a single perceptual model of the runoff generation process is assumed to be applicable under all conditions. Clearly, this is not an ideal situation. One possible response to the one-size-fits-all problem is the development of modular frameworks in which individual model components can be assembled so as to represent the modeller's perceptual model of the runoff generation process. PERSiST shares some of the design goals of FLEX (Fenicia et al., 2006), SUPERFLEX (Kavetski and Fenicia, 2011; Fenicia et al., 2011) and DYNAMIT (Hrachowitz et al., 2013b). These three modelling frameworks give the modeller an ability to specify different model structures by linking an arbitrary number of reservoirs to represent water flow through the landscape. This flexibility can be very useful for representing different

perceptual models of the runoff generation process (Fenicia et al., 2013; Hrachowitz et al., 2013b). One clear advantage of flexible modelling frameworks over traditional rainfall-runoff models is that they provide a great deal of flexibility in model structure, thereby facilitating a more credible representation of underlying perceptual models and improved understanding of catchment hydrology.

Here, we describe the structure and assumptions of the PERSiST framework, present a model application simulating flows at 8 sites along the main branch of the river Thames in the UK, and describe a simple Markov chain Monte Carlo calibration strategy useable with PERSiST and other models in the INCA family.

2 Model description

PERSiST is a watershed-scale hydrological model suitable for simulating terrestrial runoff and streamflow across a range of spatial scales from headwaters to large river basins. It is a conceptual, daily time step, semi-distributed model designed primarily for use with the INCA family of models. PERSiST simulates water fluxes from precipitation through the terrestrial part of a catchment and into rivers and streams. Key model features include (i) a user specified model structure suitable for simulating multiple perceptual models of catchment water stores and flow pathways; (ii) semi-distributed flow routing incorporating runoff production from multiple hydrologic response types; (iii) an ability to simulate flows at multiple points in a river network; (iv) capacity to simulate inundation and infiltration of riparian areas; (v) simple temperature index snowmelt and evapotranspiration routines; (vi) abstraction and discharge from industrial sources including drinking water supply; (vii) simulation of biogeochemically important low flow runoff events, (viii) a full water balance; (ix) an ability to simulate mean transit times of different water stores throughout the landscape; and (x) generation of input time series files for use with INCA. The model has been implemented as a set of first-difference equations.

At its core, PERSiST is a conceptual, bucket-type model. A watershed is represented as a series of subcatchments made up of one or more landscape units (Fig. 1). Landscape units are analogous to hydrological response units (sensu Wade et al., 2001). Landscape units consist of one or more user-specified water stores which control hydrologic response. These stores can be conceptualized as buckets that receive inputs of water from the atmosphere, from other buckets and potentially from river water. An arbitrary number of buckets can be specified and connected. By giving the model user, as opposed to the model developer, an ability to represent different patterns of water storage and landscape connectivity, PERSiST can be used to evaluate different perceptual models and explore effects of model structural uncertainty on runoff prediction. It should be stressed that PERSiST is limited to representing perceptual models of the runoff process based on a conceptualization of the catchment as a series of linked buckets.

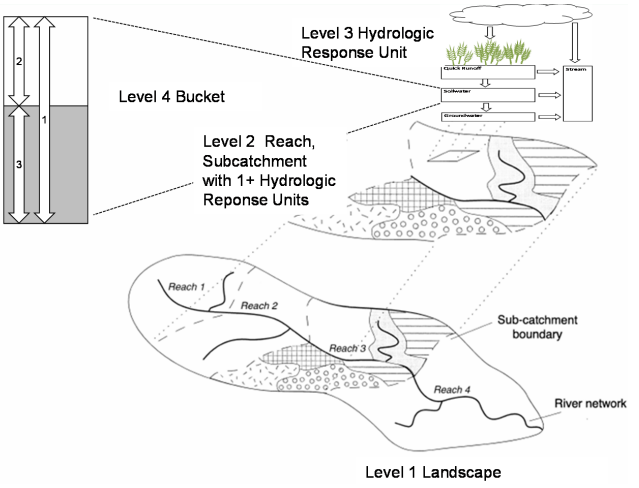

Fig. 1. Conceptual representation of the landscape in PERSiST adapted from Wade et al. (2002). A watershed (level 1) is represented as one or more reach/subcatchments (level 2). Within each subcatchment, there are one or more hydrologic response units (level 3). Each hydrologic response unit is made up of one or more buckets through which water is routed (level 4).

The model requires daily time series of air temperature and precipitation from one or more sites as driving data. PERSiST can also be used for projecting possible future patterns of runoff by using downscaled temperature and precipitation time series from regional or global climate models. PERSiST is calibrated against stream flow measured at one or more points in a river. The necessary spatial data to run the model include descriptions of all relevant hydrological response unit types, subcatchment areas, the proportional coverage of different hydrologic response types within each subcatchment, and reach (river or stream) information including length and average width. When available, additional data on abstraction and discharge volumes or stream flow velocity can aid in model calibration. Data on soil moisture or depth to groundwater can be used for soft calibration sensu Seibert and McDonnell (2002). Calibration can be based on Pearson's correlation (R^2), Nash–Sutcliffe statistics for untransformed (NS) and log-transformed (logNS) series, mean absolute error (MAE) and root mean square error (RMSE).

Elements of PERSiST draw on both HBV and INCA. PERSiST uses temperature index representations of snow dynamics similar to those in HBV and the semi-distributed landscape representation from INCA. Like INCA, PERSiST represents a watershed as one or more subcatchments connected by a stream or stream network (Fig. 1). The stream is divided into reaches, with one reach per subcatchment. One or more hydrologic response types with different hydrologic properties based on land cover, land use or underlying geology are distributed across the watershed. A key feature

Table 1. Model parameters applicable at a hydrologic response type (l_1–l_6), subcatchment (s_1–s_3), reach (r_1–r_7) and bucket (b_1–b_{10}) level. "SS" indicates the subscript used for parameter identification. The min and max values represent suggested parameter ranges.

Source	SS	Name	Description	Units	Min	Max
Hydrologic response	1	Snow threshold	Temperature threshold for liquid or solid water	°C	−5	5
Hydrologic response	2	Snow multiplier	Adjustment factor relating measured precipitation to estimated snowfall	null	0.5	1.5
Hydrologic response	3	Rain multiplier	Adjustment factor relating measured precipitation to estimated rainfall	null	0.5	1.5
Hydrologic response	4	Degree day melt factor	Temperature-dependent rate at which snow melts	$mm\,°C^{-1}$	1	4
Hydrologic response	5	Degree day ET	Maximum possible temperature-dependent rate at which evapotranspiration occurs	$mm\,°C^{-1}$	0.01	0.2
Hydrologic response	6	Growing degree threshold	Temperature threshold above which evapotranspiration can occur	°C	−5	5
Hydrologic response	7	Snow interception	Depth of precipitation intercepted by canopy when air temperature is less than or equal to the snow threshold	mm	0	5
Hydrologic response	8	Rain interception	Depth of precipitation intercepted by canopy when air temperature is greater than the snow threshold	mm	0	5
Subcatchment	1	Snow multiplier	Adjustment factor relating measured precipitation to estimated snowfall	null	0.5	1.5
Subcatchment	2	Rain multiplier	Adjustment factor relating measured precipitation to estimated rainfall	null	0.5	1.5
Subcatchment	3	Area	Subcatchment area	km^2	0.01	
Reach	1	Length	Length of the main stem of the reach	m	1	
Reach	2	Width	Width of the main stem of the reach	m	0.01	
Reach	3	a	Flow velocity multiplier	null	1×10^{-6}	1
Reach	4	b	Flow velocity exponent	null	0.3	1
Reach	5	Abstraction	Rate of water removal from reach	$m^3\,s^{-1}$	0	
Reach	6	Effluent	Rate of water addition to the reach	$m^3\,s^{-1}$	0	
Reach	7	Infiltration offset	Offset for different water level baselines between reach and buckets receiving infiltration	mm	-1×10^5	1e5
Bucket	1	Max capacity	Maximum depth of water that can be held in the bucket	mm	0	
Bucket	2	Retained water depth	Depth below which water no longer freely drains	mm	0	
Bucket	3	Runoff time constant	Characteristic time constant for water drainage	d	1	
Bucket	4	Relative ET	The fraction of total evapotranspiration in a landscape unit occurring in a given bucket	null	0	1
Bucket	5	ET adjustment	Exponent for limiting evapotranspiration	null	0	20
Bucket	6	Infiltration	The maximum depth of water that may infiltrate into a bucket from any source	mm	0	
Bucket	7	Drought runoff fraction	The fraction of incoming precipitation contributing to runoff when the soil water will not freely drain	null	0	1
Bucket	8	Relative area index	Fraction of surface area covered by bucket	null	0	1
Bucket	9	Inundation threshold	The depth at which water from the reach can inundate a hydrologic response unit type	mm	0	
Bucket	10	Porosity	The void fraction of a bucket (used for calculating height of the water column)	null	0	1

differentiating PERSiST from both HBV and INCA is that it has a more flexible representation of terrestrial hydrology which gives greater flexibility in model structure and an ability to simulate a wider range of hydrologic conditions.

2.1 Precipitation and evapotranspiration

Some parameters related to precipitation and evapotranspiration (ET) in PERSiST are specified for individual hydrologic response types and are applicable across the entire watershed (Table 1). Parameters include landscape-scale snow threshold temperature (l_1; °C), snowfall (l_2) and rainfall (l_3) multipliers. When air temperatures are below snow threshold temperatures, precipitation is assumed to fall as snow and accumulate in the snowpack. The depth of snowfall is calculated by multiplying observed precipitation by the snowfall multiplier. When air temperature is above the snow threshold temperature, precipitation is assumed to fall as rain. Depth of rainfall is estimated by multiplying observed precipitation by the rainfall multiplier.

Actual ET is calculated in a manner similar to that presented by Durand (2004). However, instead of using Penman

potential ET as the baseline, a degree day evapotranspiration parameter is used (l_5; mm °C^{-1} d^{-1}) which defines the maximum (i.e. potential) ET when air temperatures are above the growing degree day threshold (l_6; °C). When air temperatures are below the growing degree day threshold, it is assumed that no ET occurs. The hydrologic-response-type-specific potential evapotranspiration ($E(l)$; mm d^{-1}) is calculated as the difference between observed air temperature (T; °C) and the growing degree day threshold multiplied by the degree day ET parameter.

$$E(l) = (T - l_6) \cdot l_5 \quad (1)$$

The actual rate of ET can be less than the maximum potential rate, depending on the amount of moisture available. This is further described in Sect. 2.3.

The model simulates canopy interception of snow (l_7; mm) and rain (l_8; mm) depending on whether the air temperature is below or above the snow threshold temperature (l_1). The interception is subtracted from precipitation before it enters the soil or snowpack.

2.2 Subcatchment and reach

A watershed is represented as one or more subcatchments. Within a watershed, the stream is divided into reaches. There is a one-to-one correspondence between subcatchments and reaches (Fig. 1). A subcatchment contains one or more hydrological response types. A different temperature and precipitation time series can be associated with each subcatchment.

The model calculates runoff from rainfall in an ad hoc manner by moving water between compartments in an arbitrary order dependent on both the perceptual model of the runoff generation process and catchment topology. Water movement within a catchment is simulated in the following manner. The reach network is represented as a directed tree graph, with the root at the catchment outlet. Stream flow is estimated in terminal (furthest upstream) reaches first. The graph is then traversed and stream flow estimated in all other reaches based on stream flow from upstream reaches and inputs from the local subcatchment. Within each subcatchment, runoff is estimated sequentially for each hydrologic response type. For each day of simulation (i) interception, rainfall, snow accumulation and snowmelt are estimated based on measured temperature and precipitation; (ii) rainfall and snowmelt are routed through the uppermost buckets in a hydrologic response type; (iii) ET from the uppermost bucket is estimated; and (iv) outflow from the bucket is estimated. Steps (i)–(iv) are repeated for each bucket in the hydrologic response type.

Additional parameters may be specified (Table 1) including snowfall (s_1) and rainfall (s_2) multipliers for each subcatchment. Effective snowfall and rainfall multipliers are determined by multiplying the landscape-scale and subcatchments-scale parameter values. The subcatchment area (s_3; km^2) and the proportional cover of each hydrologic response type must be specified. Reach parameters including length (r_1; m), width (r_2; m) and the parameters necessary to determine flow velocity (v) as a function of flow (Q) must be specified.

$$v = r_3 Q^{r_4} \quad (2)$$

Rates of water abstraction from and effluent input to individual reaches may be specified either as constant values (r_5, r_6; m^3 s^{-1}) or as time series of daily average values. The effects of land use change on hydrology can be simulated in PERSiST by allowing the relative area of hydrologic response types within a subcatchment to vary over time.

2.3 Bucket

A hydrologic response type consists of one or more buckets. Buckets can store water, return it to the atmosphere through ET, transfer it to other buckets or to surface waters. Buckets can be conceptualized as dual-porosity reservoirs in which water is divided into stagnant and freely draining fractions (Šimůnek et al., 2003).

Each bucket has the following properties (Fig. 3, Table 1): depth of water in the bucket at time t (z_t: mm); the maximum depth of water that can be held in a bucket (b_1; mm); and the retained water depth (b_2; mm), which is the depth below which water no longer freely drains. When water is below the retained water depth, ET can continue and drought-related runoff (described below) can occur. A characteristic time constant (b_3; d) specifies the rate at which water drains from a bucket. The depth of water draining on day t is calculated as follows:

$$\Delta z_t = \frac{z_t - b_2}{b_3}. \quad (3)$$

Water can be returned to the atmosphere through ET and canopy interception. While water returned to the atmosphere through evaporation is released from surfaces, water returned as transpiration may be derived from different depths, depending on the root structure of the vegetation community in the simulated landscape unit. Thus, there is a need to simulate different rates of ET from different buckets. Typically, the rate will be highest in a surface bucket as it will include evaporation and transpiration. The rate of evapotranspiration from a bucket ($E(b)$; mm d^{-1}) is determined by multiplying the landscape-scale maximum possible rate ($E(l)$; mm d^{-1}) by the relative ET index (b_4). Note that relative ET indices must sum to unity within a hydrologic response unit so the base degree day ET is consistent with the total ET produced. When the depth of water in a bucket is above the retained water depth, ET occurs at the maximum rate. When the depth of water is below the retained water depth, the rate of ET can be limited as follows:

$$E(b) = \left(\frac{z}{b_2}\right)^{b_5} \cdot b_4 \cdot E(l). \quad (4)$$

Changing values of b_5 adjust the rate at which ET slows when the depth of water in a bucket is below the retained water depth. A value of 0 indicates that ET will be unchanged by soil moisture status, while high values (10+) effectively stop ET when there is no longer any freely draining water.

The amount of water that can be added to a bucket in any given time step is limited. It cannot exceed the infiltration rate (b_6) or the difference between the maximum and current depths of water (b_1-z). Water which is prevented from percolating due to the infiltration capacity being exceeded is referred to as "infiltration excess" and is routed to the adjacent stream. Water that is prevented from percolating due to the maximum storage capacity of the receiving box being exceeded is referred to as "saturation excess", and will be routed back to the quick box as described in the section on hydrologic response units.

Low-flow events can have a disproportionate importance for surface water solute chemistry. Small increases in flow, which have a negligible effect on the overall hydrograph, can transport high concentrations of solute to streams. This has been observed for the flushing of nitrate to the Thames (Jin et al., 2012) or organic carbon from boreal catchments (Ledesma et al., 2012, and references therein). In many hydrological models, rainfall during dry conditions is assumed to contribute only to recharging soil and groundwater. However, a small fraction of the rain may contribute to stream flow and solute transport, even when soils are very dry. This phenomenon is simulated in PERSiST using the drought runoff fraction (b_7). When the depth of water in a bucket is below the retained water depth, the amount of water entering the bucket which contributes to runoff is estimated by multiplying the total input by the drought runoff fraction. Note that the default behaviour is for all water inputs to contribute only to recharge when depth of water in the bucket is below the freely draining depth.

There are two special bucket types. Quick-flow buckets simulate surface processes, while bidirectional buckets can receive inputs of water from the river through infiltration or inundation. Quick-flow buckets receive inputs of rainfall and snowmelt. Saturation excess flow generated by other buckets in a hydrologic response type is routed through a quick bucket to the adjacent reach. Each hydrologic response type must include one or more quick-flow buckets.

The vast majority of water flows within a watershed are from land to surface waters. However, flows from rivers to the land can also occur. During flood conditions, a river can overflow its banks and inundate the surrounding land. It is also possible for water to infiltrate from a river to the surrounding land. Both of these phenomena can be biogeochemically important. Inundation can deposit large amounts of sediments and nutrients (Bayley, 1995), while infiltration can alter rates of nitrogen processing (Grischek et al., 1998). To simulate either inundation or infiltration in PERSiST, a bidirectional flow bucket must be identified. It is only possible to have one inundation and one infiltration bidirectional flow bucket in a hydrologic response type.

Inundation can only be simulated for bidirectional quick-flow buckets. Inundation is simulated when the depth of water in the reach (z_R ; mm) exceeds the bucket-specific inundation threshold (b_9; mm). The volume of water inundating per unit time (V_{Inundate}; $\text{m}^3\,\text{d}^{-1}$) is estimated as the fraction of the total flow through the reach outflow which occurs at a depth exceeding the inundation threshold. A rectangular channel cross section and uniform flow velocity throughout the reach is assumed when calculating inundation.

$$V_{\text{Inundate}} = \left(\frac{z_R - b_9}{z_R}\right) \cdot Q \cdot 86\,400 \tag{5}$$

The depth of water inundating the bucket per unit time (z_{Inundate}; $\text{mm}\,\text{d}^{-1}$) which corresponds to V_{Inundate} is calculated by dividing the inundation volume by the relative area of the bucket in the relevant hydrologic response unit all multiplied by the subcatchment area.

$$z_{\text{Inundate}} = \frac{V_{\text{Inundate}}}{b_8 \cdot s_3 \cdot 1000} \tag{6}$$

Inundating water is treated the same way as other inputs to a quick-flow bucket.

Infiltration can be simulated in PERSiST by moving water from the reach to either a regular or quick-flow bucket. The volume of water infiltrating is dependent on stream flow and the difference in height between water in the stream and the bucket. The height of water in a bucket is determined by dividing water depth (z; mm) by porosity (b_{10}). Infiltration occurs when water depth in the reach (z_R) exceeds the height of the water column of the receiving bucket ($z_R - r_7 > z/b_{10}$), where r_7 is an offset to account for cases where water depths in the stream and riparian soil are measured against different reference levels. The volume of water infiltrating from the stream to a bucket per unit time ($V_{\text{Infiltrate}}$; $\text{m}^3\,\text{d}^{-1}$) is calculated in a similar manner as for inundating water. After some algebraic rearrangement, $V_{\text{Infiltrate}}$ can be expressed as follows:

$$V_{\text{Infiltrate}} = \left(1 - \frac{1}{z_R}\left(\frac{z}{b_{10}} + r_7\right)\right) \cdot Q \cdot 86400. \tag{7}$$

The increase in water depth in the bucket per unit time due to infiltration ($z_{\text{infiltrate}}$; $\text{mm}\,\text{d}^{-1}$) can be calculated as follows:

$$z_{\text{Infiltrate}} = \frac{b_{10} \cdot V_{\text{Infiltrate}}}{b_8 \cdot s_3 \cdot 1000}. \tag{8}$$

Both inundation and infiltration are estimated for each hydrologic response unit in a subcatchment. Inundating and infiltrating volumes are estimated first for the hydrologic response unit with the lowest water depth in a subcatchment. Infiltration and infiltration are then estimated sequentially until the depth of water in the terrestrial part of the subcatchment is equal to the depth of water in the reach.

Table 2. Example square matrix for the Thames PERSiST application. The upper square matrix corresponds to the hydrological response unit shown in Fig. 2 (upper). The lower square matrix contains the values used in the simulations presented here.

	Direct runoff	Soil water	Groundwater
Direct runoff	a	b	
Soil water	c	d	e
Groundwater			f
	Direct runoff	**Soil water**	**Groundwater**
Direct runoff	0.1	0.9	0
Soil water	1	0.4	0.6
Groundwater	0	0	1

2.4 Hydrological response (landscape) unit types

A landscape (or hydrological response) unit type consists of one or more buckets linked together in a user-specified manner. Hydrologic response units in PERSiST are analogous to land cover types in INCA. Water is routed from a hydrologic response unit directly to the reach. Thus, there is no movement of water between landscape units, but PERSiST is able to represent perceptual models, for example, of runoff generated from recharge and discharge areas by appropriate combinations of buckets. Each hydrologic response unit type must contain one or more quick buckets to receive inputs of precipitation. Flows of water between buckets are described using a square matrix (Table 2). Element (i,j) of the square matrix represents the fraction of water leaving bucket i which is added to bucket j. Diagonal elements (i,i) on the square matrix define the fraction of water leaving the bucket that is routed directly to the stream. Off-diagonal elements $(i,\ j;\ j > i)$ in the upper quadrant define the fraction of water leaving bucket i and entering bucket j. Values of cells in each row from the diagonal to the rightmost entry must sum to 1. Below diagonal elements of the square matrix are used to identify the quick bucket to which saturation excess flow can be routed.

$$\sum_{i \geq r}^{i=c} m_{r,i} = 1 \tag{9}$$

The configuration of buckets shown in Fig. 2 is only one possible representation of the runoff generation process. It is similar to the INCA representation of terrestrial hydrology which has proved successful in temperate (Whitehead et al., 1998; Wade et al., 2002), montane (Ranzini et al., 2007; Futter et al., 2009b) and boreal (Rankinen et al., 2004; Futter et al., 2009a) conditions. It has been less successful in Mediterranean conditions (Bernal et al., 2004). However, PERSiST can be set up to include alternative model structures proposed for the simulation of Mediterranean hydrology (e.g. Medici et al., 2008) or diverse montane catchments (Kavetski and

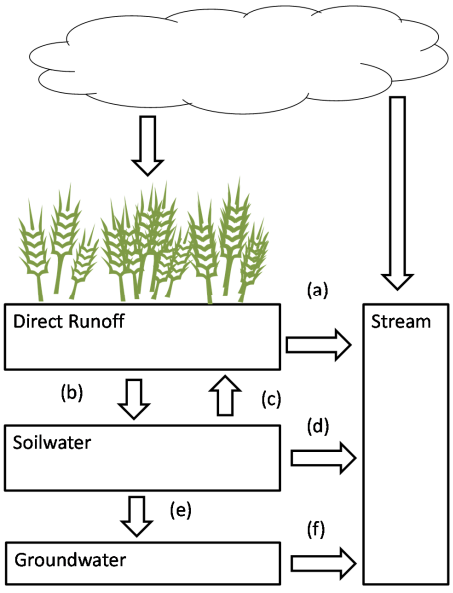

Fig. 2. Simple hydrologic response unit comprised of three buckets representing direct runoff, soil water and groundwater. The arrow labels **(a)**–**(f)** identify different fluxes and can be linked to the flow partitioning matrix shown in Table 2. **(a)** represents direct runoff to the river; **(b)** is infiltration from the direct runoff to soil water bucket; **(c)** represents saturation excess return flow from the quick bucket to the soil surface, while **(d)** is flow from the soil water bucket to the river. Flux **(e)** is from the soil water to groundwater bucket, while flux **(f)** represents water flow from the groundwater bucket to the river.

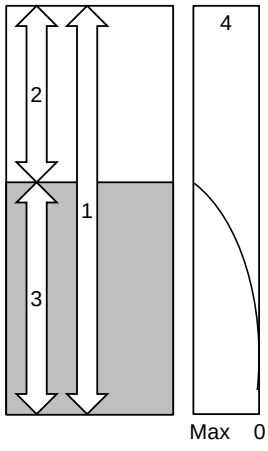

Fig. 3. Generic bucket structure (left) and relative evapotranspiration rate as a function of water depth (right). The maximum depth of water in a bucket (1; Table 1 b_1) can be partitioned into freely draining water (2) and water that may contribute to evapotranspiration but not drainage (3; b_2). The rate of evapotranspiration (4) is not constrained when there is freely draining water in the bucket. When the water depth drops below the freely draining depth, evapotranspiration is limited as a power function of water depth.

Fenicia, 2011). It is possible with PERSiST to link an arbitrary number of buckets in whatever manner provides the most appropriate representation of the modeller's perceptual model of the runoff generation process. For example, riparian and upland areas can be simulated. Because of the lateral connectivity between upland and riparian areas, representing watershed hydrology using a series of vertically stacked buckets may not be appropriate for simulating the hydrochemistry of headwater forest (Löfgren et al., 2011) or agricultural (Stutter et al., 2009) watersheds. When simulating laterally connected bucket systems used to represent riparian and upland areas, some additional parameters must be specified for the different buckets so as to ensure that hydrology is represented correctly. The relative area index (b_8) is used to describe the areal contribution of each bucket to a hydrologic response unit type. For example, a hydrologic response unit type consisting of upland and riparian areas might have relative areas of 0.9 and 0.1 respectively for upland and riparian soil water buckets.

2.5 Calculating fluxes of water in a hydrologic response type

Unlike many other models which are based on the simultaneous solution of a set of first-order differential equations, fluxes of water through the landscape are calculated sequentially in PERSiST. Fluxes are calculated in the same order as that in which buckets are identified in the square matrix. The first row in the square matrix should represent a quick-flow bucket and the last row a bucket where water drains only to the stream. All fluxes between buckets are calculated in units of $m^3\ d^{-1}$ for a representative 1 km^2 landscape unit type (level 3 in Fig. 1). Actual fluxes from the subcatchments to the reach are then estimated by multiplying by the appropriate subcatchment area.

Fluxes are calculated as follows. First, the depth of incoming precipitation is reduced by the appropriate canopy interception factor (l_7 or l_8). Second, ET is calculated according to Eq. (4). Next, outputs to other buckets are calculated. The volume of water transferred from bucket i to bucket j ($V_{i,j}$) is calculated as follows. First, the default output volume is estimated by multiplying the appropriate value in the square matrix ($m_{i,j}$) by the area of the depth of water (Δz_i; Eq. 4) able to leave the bucket all multiplied by relative bucket area (b_8). For each transfer, a test is performed to see if the volume of water being transferred is greater than the available volume in the receiving bucket. If the available volume in the receiving bucket (j) is smaller than the potential volume of water leaving the source bucket (i), then the volume leaving the source bucket is reduced accordingly.

$$V_{i,j} = \min \begin{bmatrix} b_{8,i} \cdot \frac{z_i - b_{2,i}}{b_{3,i}} \\ b_{8,j} \cdot \left(b_{1,j} - z_j\right) \end{bmatrix} \quad (10)$$

After each transfer, the depths of water in buckets i and j are adjusted accordingly. Third, if a bucket is a quick-flow type, then snowmelt, rainfall and inundation are added. When the temperature is warm enough for snowmelt, the depth of snowmelt (z_{snowmelt}) is calculated as follows:

$$z_{\text{snowmelt}} = \min \begin{bmatrix} l_4 \cdot (T_{\text{Air}} - l_1) \\ z_{\text{Snow}} \end{bmatrix}, \quad (11)$$

where l_4 represents the degree day snowmelt factor and (T_{Air}-l_1) is the number of degrees above the snowmelt offset. Rainfall (z_{rain}), which only occurs when the air temperature is above $l_1 + s_1$, is equal to the observed depth of precipitation (P) multiplied by the watershed (l_3) and subcatchment (s_2) precipitation multipliers.

$$z_{\text{rain}} = l_3 \cdot s_2 \cdot P \quad (12)$$

Saturation excess flow is simulated when the depth of water in a bucket exceeds the maximum possible water depth. All saturation excess flow must be routed to a quick bucket where it is treated the same way as other inputs from precipitation or inundation.

2.6 PERSiST time series

PERSiST generates terrestrial and aquatic time series including daily inputs, outputs and changes in storage for each bucket in each hydrologic response unit type in every subcatchment (Table 3). At each time step, depths of water in each bucket and all transfers between buckets, total fluxes of water from subcatchment to reach, reach volumes, stream flow, infiltration and inundation are all recorded. Atmospheric exchange (precipitation, interception and ET) and snowpack dynamics are also reported.

2.7 INCA compatibility

One of the design criteria for PERSiST is to generate input data files for the INCA model. Currently, INCA requires the use of an external rainfall-runoff model to generate SMD and HER time series. Subcatchment and watershed-scale estimates of SMD are produced in PERSiST based on average differences between depth of water in a bucket (z; mm) and bucket maximum water-holding capacity (b_1; mm). There is a possibility to adjust this depth by a user-specified offset so as to obtain SMD time series with a minimum value of 0.

HER is an estimate of the precipitation entering a watershed which eventually contributes to runoff. In PERSiST, this can be estimated at the subcatchment scale as precipitation minus interception and ET. These calculations take into account the effect of soil moisture status on ET rates. HER is estimated by working backwards through time series of simulated ET and precipitation inputs. Starting from the last day in the simulation, ET is accumulated and then the precipitation for that day is subtracted. If the value is less than 0 (i.e. precipitation is greater than the cumulative ET deficit) then the difference is recorded as HER for that day and the cumulative ET set to 0. This is repeated until the start of the

Table 3. PERSiST Internal Time Series.

Code	Description	Bucket	Hydrologic response Unit	Subcatchment	Reach	Watershed	Units
$R(l)$	Rainfall					Yes	$\mathrm{mm\,d^{-1}}$
$S(l)$	Snowfall					Yes	$\mathrm{mm\,d^{-1}}$
$R(s)$	Rainfall			Yes			$\mathrm{mm\,d^{-1}}$
$S(s)$	Snowfall			Yes			$\mathrm{mm\,d^{-1}}$
$Z(s)$	Depth of snow in a subcatchment			Yes			mm
Q	Stream flow				Yes		$\mathrm{m^3\,s^{-1}}$
v	Stream velocity				Yes		$\mathrm{m\,s^{-1}}$
z_R	Depth of water in a reach				Yes		m
z	Depth of water in a bucket	Yes	Yes	Yes			mm
$E(l)$	Maximum possible rate of evapotranspiration					Yes	$\mathrm{mm\,d^{-1}}$
$V(i,j)$	Volume of water transferred from bucket i to bucket j	Yes	Yes	Yes			$\mathrm{m^3\,d^{-1}}$

simulation is reached. The adequacy of this assumption is tested by comparing total estimated HER to total runoff.

2.8 Markov chain Monte Carlo tool

A simple Markov chain Monte Carlo (MCMC) tool was developed to identify credible parameter sets, assess parameter sensitivity and generate ensembles of model predictions. The MCMC code was based on the Metropolis–Hastings algorithm (Chib and Greenberg, 1995) and did not use Gibbs sampling (Smith and Roberts, 1993). Each instance of the MCMC sampler operates as follows:

1. An initial manual calibration provides a basis for identifying credible parameter ranges to be sampled during MCMC analysis. The system is initialized using the parameter set from the manual calibration to specify the best model performance and parameter set. Model performance is estimated using Nash–Sutcliffe statistics (Nash and Sutcliffe, 1970) from one or more untransformed (NS) and log-transformed (logNS) flow time series Σ ((NS-1) + (logNS-1)) for modelled and observed flows. The statistic has a maximum value of 0 and a minimum of $-\infty$.

2. A random starting point is drawn from the parameter space, and model performance is recorded. If model performance is better than the best model performance, the starting point is accepted and the best model performance and parameter set are updated. If model performance from the random starting point is worse than the best model performance, the ratio between the two values is compared to a random number between 0 and 1. If the ratio exceeds the random threshold, the starting point is accepted; otherwise a new starting point is drawn and model performance is evaluated.

3. A jump is defined which randomly perturbs the parameter values.

4. The jump is applied to the parameter set, the model is run with the new parameter set and model performance is assessed. If performance is better than that from the previous parameter set, the jump is repeated. If model performance is better than the best model performance, the best model performance and parameter set are updated. This process is repeated until no further improvement in model performance is obtained or until a jump would cross the maximum or minimum value identified in (i). If the jump would cross the maximum or minimum threshold, a new value is randomly selected within the range of that parameter.

5. If a jump does not lead to an improvement in model performance, it may still be accepted if the ratio of new and old model performance exceeds a random number between 0 and 1. If the jump is rejected, a counter is incremented. If the counter exceeds a user-specified threshold, control returns to (ii) and a new random starting point is defined. If the threshold is not exceeded, control returns to (iii).

6. This process is repeated an arbitrary number of times and the best-performing parameter set retained for further analysis.

Steps (i)–(vi) are repeated a user-specified number of times so as to generate an ensemble of credible parameter sets. The credible parameter sets are not globally optimal, but are the result of a non-exhaustive MCMC exploration of the parameter space.

Parameter sensitivity was assessed by comparing the cumulative distribution of parameters to a rectangular distribution which would be indicative of parameter randomness. Parameters with a non-rectangular posterior distribution as identified by a Kolmogorov–Smirnov d statistic $>=0.2$ were assumed to be sensitive. Using a KS d statistic of 0.2 as a criterion for significant deviation from rectangular is a conservative assumption designed to control for the possibility of spurious estimates of statistical significance that could be obtained from repeated testing (Futter et al., 2009b).

3 Model application

3.1 Site description

An application of PERSiST is presented for the river Thames in the UK. The Thames watershed supplies drinking water for ∼14 million people in the greater London region (Bloomfield et al., 2009; Jin et al., 2012; Whitehead et al., 2013; Crossman et al., 2013a). The watershed has an area of ∼10 000 km^2 above the tidal limit at Teddington and the main river is ∼235 km long. Land use is predominantly agricultural in the upper subcatchments and becomes more urban near the outflow. There are approximately 45 locks and other water control structures along the main stem of the Thames and numerous effluent discharges and water abstraction points. Catchment elevation ranges from 0 to 325 metres above sea level (Bloomfield et al., 2009). Precipitation is between 600 and 900 mm yr^{-1} (Marsh and Hannaford, 2008). Base flows throughout the watershed are variable, ranging between 0.3 and 0.95 (Marsh and Hannaford, 2008).

Underlying geology in the Thames watershed can be divided into three regions: the Midlands Shelf in the northwest, the central London Basin and the Wealdean Anticline in the southeast (Bloomfield et al., 2009). The Midlands Shelf is primarily Jurassic and Cretaceous oolitic limestone, clay and chalk. The London Basin has significant quaternary deposits and the Wealdean Anticline includes Lower Cretaceous clay and sandstone (Bloomfield et al., 2009).

INCA-N and *P* applications to the Thames catchment have conceptualized hydrologic response in the catchment primarily on the basis of land use (Jin et al., 2012; Crossman et al., 2013a; Whitehead et al., 2013). Preliminary PERSiST model applications to the Thames suggested that there were few meaningful differences between hydrologic responses of different land use types. Much of the difference in hydrologic response across the Thames Basin seemed to be related to underlying geology. We hypothesized that there would be significant differences between subcatchment hydrologic response depending on whether or not chalk bedrock was present.

3.2 Model application

For the application presented here, three hydrologic response types were used representing chalk bedrock, non-chalk bedrock and Quaternary sand, slit and clay (Fig. 4, Table 4). Areas of the different hydrologic response types were obtained by generalizing data from British Geological Survey 1 : 62 500 maps of surficial and bedrock geology. The Thames was divided into 8 reaches based on Jin et al. (2012). Reach boundaries were established at the location of flow measuring stations (Fig. 4). Subcatchment areas, reach length and proportional cover of different hydrologic response types are shown in Table 4. A single time series of temperature and precipitation obtained from the UK Met Office (Crossman et al., 2013a) was used to drive the model. The meteorological time series was based on a synthesis of data from observing stations in the Thames watershed.

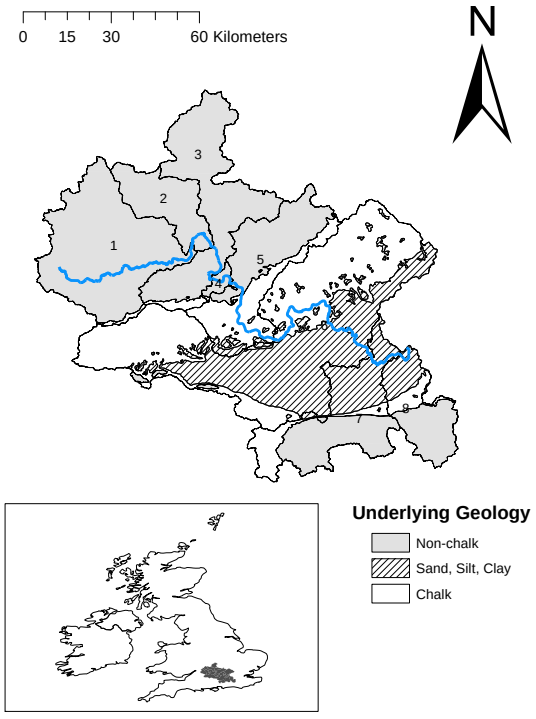

Fig. 4. Map of the river Thames watershed showing the main branch of the Thames, underlying geology and boundaries for the 8 subcatchments used in the present study. The flow gauging stations are located where the main branch of the river crosses the subcatchment boundary.

Each hydrologic response type was simulated as three vertically stacked buckets representing direct runoff, soil water and groundwater (Fig. 2). This is similar to the terrestrial hydrological representation used in INCA (Whitehead et al., 1998; Wade et al., 2002). It was assumed that all precipitation entering the direct runoff bucket percolated to the soil water bucket. Water could leave the soil water bucket as runoff, percolation to groundwater or as saturation excess flow which was routed immediately to the stream. All water entering the groundwater bucket was assumed to flow to the stream. It was assumed that there were no losses to deep groundwater.

The model application was performed for the period 1 January 1999–31 December 2008. PERSiST was calibrated against observed stream flows at 8 locations along the main stem of the Thames (Fig. 4). The model was calibrated by first manually adjusting parameters so as to improve the fit between modelled and observed stream flow. In all cases, model performance was assessed using untransformed and log-transformed data for each flow measurement location. Thus, a total of 16 goodness-of-fit statistics were evaluated for each candidate parameter set. Using both untransformed and log transformed data achieves some balance between fitting high and low flows.

Table 4. Subcatchment and reach dimensions along with proportional cover of different landscape element types.

No.	Name	Sand, silt & clay	Chalk	Non-chalk	Area (km^2)	Reach length (m)
1	Crickdale Castle to Pinkhill	0	1.3	98.7	1609	54 100
2	Pinkhill to Osney	0	0	100	526	12 420
3	Osney to Culham	0	4	96	1288	18 960
4	Days Weir	0	4	96	58	9320
5	Days Weir to Caversham	5	40	55	1154	35 150
6	Caversham to Shepperton	42	57	1	3632	70 410
7	Mosley	35	17	48	1102	9540
8	Teddington	29	17	64	589	7740

The manual calibration first attempted to obtain the best possible fit to the uppermost reach. Once this was obtained, attempts were made to fit subsequent reaches. Manual calibration continued until it was no longer possible to obtain obvious improvements in model performance. The subsequent automated calibration was based on maximizing the sum of NS and logNS statistics. Thus, it was possible to accept poorer model performance in one reach if the parameter set under evaluation led to better fits in other reaches.

The final manual calibration was used as the starting point for a simple MCMC exploration of parameter space. A total of 39 parameters were allowed to vary during the MCMC analysis (Table 5). Five parameters representing the growing degree threshold temperature (l_6), ET rate limitation (b_5) and time constants for the quick, upper and lower soil buckets were allowed to vary for each hydrologic response type. The subcatchment-specific rain multiplier (s_2) and reach flow : velocity a and b parameters (r_3, r_4) were allowed to vary individually for each of the 8 subcatchments. The MCMC tool was run 500 times, with each chain consisting of 5000 model runs so as to generate an ensemble of credible parameter sets.

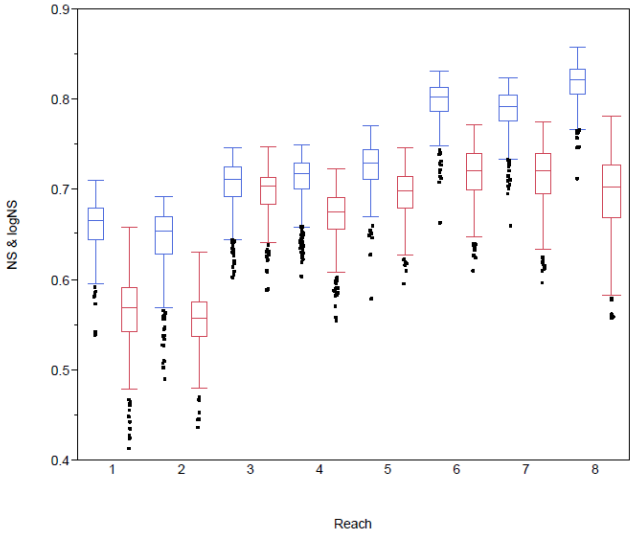

Fig. 5. Values for Nash–Sutcliffe (blue) and log Nash–Sutcliffe (red) goodness-of-fit statistics obtained for each reach from the ensemble of 500 credible model calibrations.

4 Results

Two general trends are apparent in the model goodness-of-fit statistics from the ensemble of parameter sets obtained during the MCMC analysis (Fig. 5). Model goodness of fit improved lower in the river (higher reach numbers) and the NS statistics were typically higher than logNS statistics.

The ensemble of best-performing parameter sets were re-run through PERSiST to generate ensembles of predicted values. Results are shown for the uppermost (Fig. 6, Pinkhill) and lowermost (Fig. 7, Teddington) reaches. The ensemble of PERSiST parameter sets was able to capture the timing of peak flows at Pinkhill (Fig. 6a); however it missed the absolute magnitude with over-predictions in 2001 and 2003 and under-predictions in 2002, 2004 and 2006. Examination of the log-transformed values for Pinkhill (Fig. 6b) shows that in almost all years PERSiST consistently was too late in the predicted timing of lowest flows. An examination of measured versus observed flows at Pinkhill showed an apparent measured maximum of 70 m^3 s^{-1}, while modelled flows in some cases exceed 100 m^3 s^{-1}. As corroborated by the higher NS and logNS statistics at Teddington versus Pinkhill (median 0.82, 0.70 versus 0.66, 0.56), there is a better fit between modelled and observed flows at the lowest measurement point in the Thames (Fig. 7a, b). At Teddington, the model did a better job of simulating the timing and height of high-flow events (Fig. 7a). The model was also more successful at reproducing both timing and amounts during low-flow periods (Fig. 7b).

With the exception of the ET adjustment, model performance was only sensitive to land-phase parameters for non-chalk underlying geology (Fig. 8, Table 6). There was a slight tendency towards lower growing degree day thresholds for the non-chalk hydrologic response type (Fig. 8a). The distribution of ET adjustment parameter values was strongly skewed towards lower values for the non-chalk hydrologic response type. Lower ET adjustment parameter values facilitate greater simulated rates of ET when the depth of water in a bucket is less than the freely draining depth. The posterior

Table 5. Parameter ranges used during MCMC analysis. Note that the range of values used for individual subcatchments and reach were smaller than the overall range shown in the table.

Location	Parameter	Min	Max	n	Units
Hydrologic response type	Growing degree threshold temperature	−2	0	3	°C
Hydrologic response type	ET rate limitation	1	10	3	
Hydrologic response type	Quick bucket time constant	1	3	3	days
Hydrologic response type	Soil water (upper) bucket time constant	1	15	3	days
Hydrologic response type	Groundwater (lower) bucket time constant	50	80	3	days
Reach	Flow a (r_3)	0.001	0.1	8	
Reach	Flow b (r_4)	0.1	0.95	8	
Subcatchment	Rain multiplier	0.5	1.25	8	

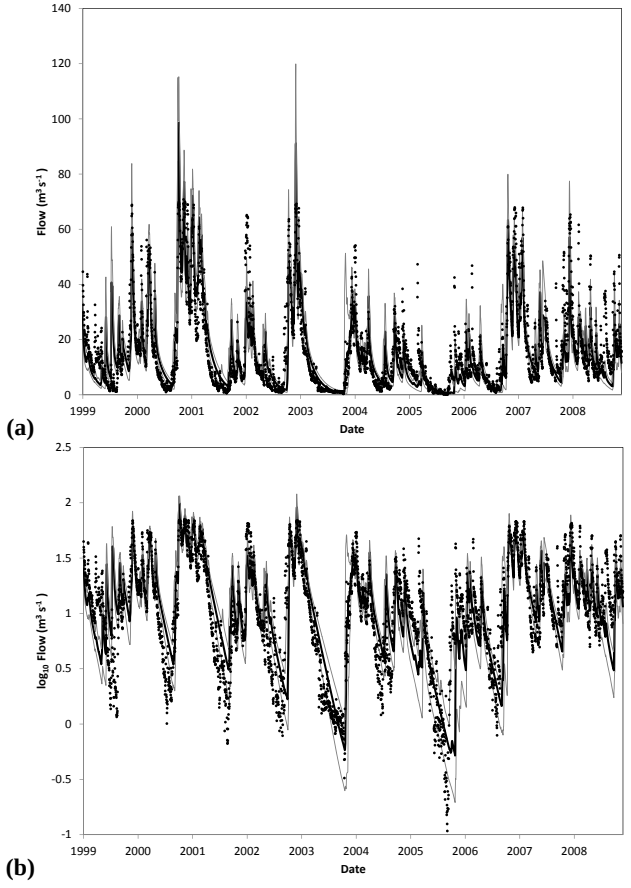

Fig. 6. (a) Observed (dots), maximum average and minimum simulated flows at Pinkhill, the uppermost flow measurement site in the Thames (reach 1). **(b)** Plot of $\log_{10}$ transformed observed (dots), minimum, average and maximum modelled flows at Pinkhill (reach 1).

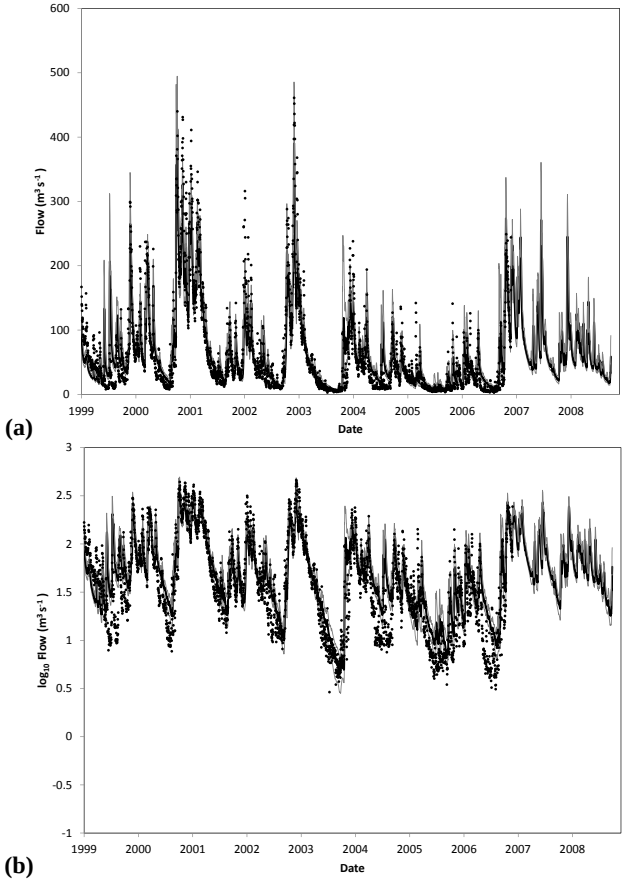

Fig. 7. (a) Observed (dots), maximum, average and minimum simulated flows at Teddington, the lowermost site simulated in the present study (Reach 8). **(b)** Plot of $\log_{10}$ transformed observed, maximum, average and minimum modelled flows at Teddington (reach 8).

distributions of the two ET-related parameters (Fig. 8a, b) suggest that parameter sets with higher rates of simulated ET for the non-chalk hydrologic response types are more likely to be included in the ensemble of credible parameter sets. The posterior distribution of the ET adjustment for chalk and Quaternary sand, silts and clay are not as extreme as for the non-chalk hydrologic response type, but they do support some simulated ET when water is below the freely draining depth (Fig. 8b). Both the upper (Fig. 8c) and lower (Fig. 8d) time constants are skewed towards the lower end of the sampled parameter range for the non-chalk hydrologic response type.

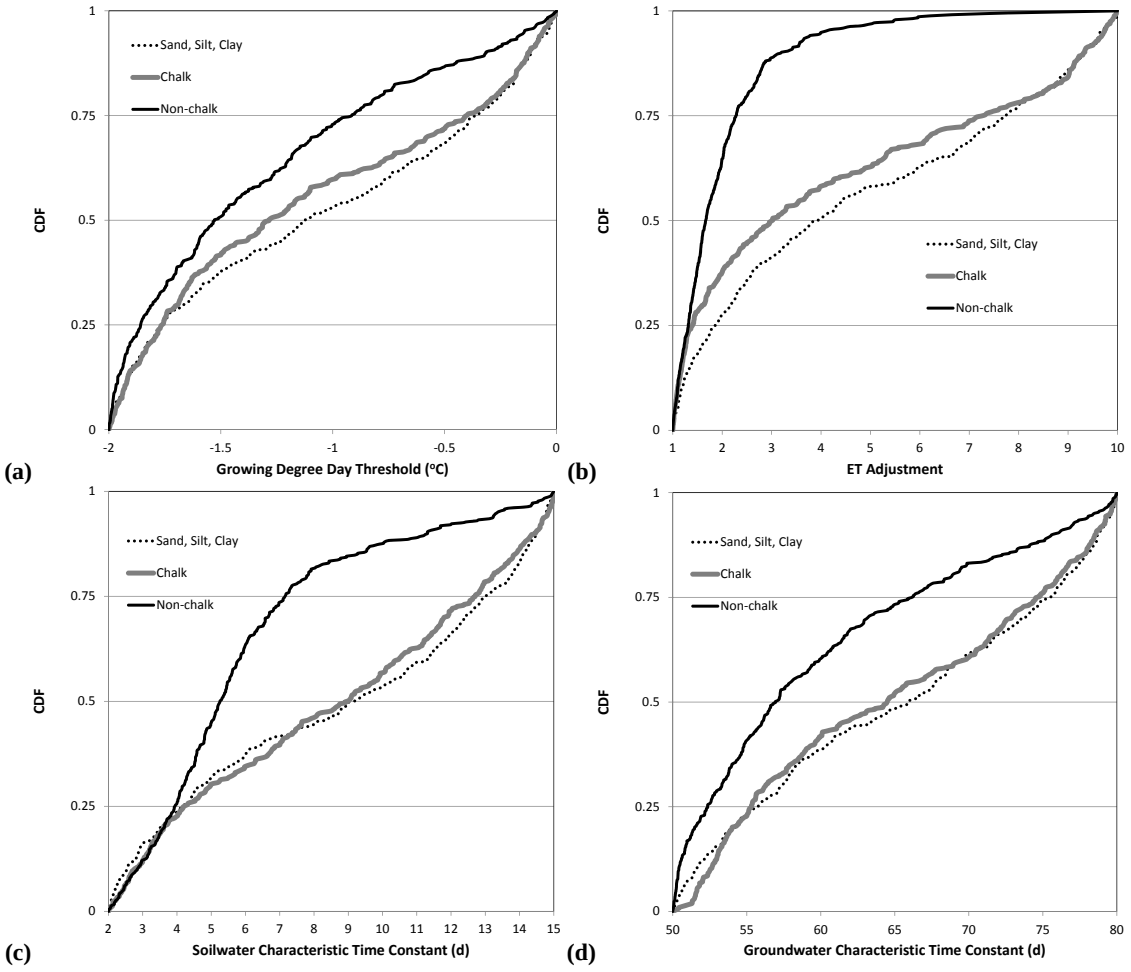

Fig. 8. (a) Cumulative distribution plot of growing degree thresholds for the three hydrologic response types. The KS statistics for the chalk and Quaternary (sand, silt, clay) hydrologic response types are not statistically significant, suggesting model performance is only sensitive to values in the non-chalk hydrologic response type. **(b)** Cumulative distribution plot of evapotranspiration adjustment factors for the three hydrologic response types. **(c)** Cumulative distribution plot of soil water characteristic time constants for the three hydrologic response types. **(d)** Cumulative distribution plot for groundwater time constants for the three hydrologic response types.

Model performance was insensitive to values for the quick time constant and the drought runoff fraction in all hydrologic response classes (Table 6). There was relatively little difference in the posterior distributions of growing degree thresholds and upper and lower bucket time constants between the Chalk and Quaternary sand, silt, and clay hydrologic response types.

Preliminary manual calibrations suggested that model performance was very sensitive to values of the subcatchment rainfall multiplier. Subcatchment-specific parameter ranges for the rainfall multiplier were used in the MCMC analysis since model performance was so strongly influenced by the values for this parameter. Values of the flow velocity relationship parameters did not appear to influence model performance in any reach (Table 6).

Pearson's correlations calculated for all pairs of parameters from the ensemble of credible parameter sets showed a number of statistically significant relationships between parameter values (Table 7). There were negative correlations between time constants in the non-chalk hydrologic response type, suggesting that, if one time constant were too high, it could be compensated for with a lower value of the other time constant. Reach rain multipliers and growing degree thresholds were negatively correlated at several points throughout the catchment. This suggests that bias in the precipitation input could be corrected by a countervailing bias in the amount of ET. There were more significant correlations for parameters related to the uppermost reach than for reaches further downstream.

5 Discussion

Here, we present a new model for simulating fluxes of water through heterogeneous landscapes and a simple Monte Carlo tool for model calibration and sensitivity analysis.

Table 6. Kolmogorov–Smirnov (KS) *d* statistics showing parameter sensitivity assessed as deviation of posterior parameter distribution from rectangular. Parameters with a significant effect on model performance are shown in *italic*. Statistical significance has been adjusted to account for multiple comparisons.

Parameter	Hydrologic response type		
	Sand, silt, clay	Chalk	Non-chalk
Growing degree day threshold	0.14	0.18	*0.27*
Quick time constant	0.12	0.11	0.19
Drought runoff fraction	0.12	0.09	0.11
Upper time constant	0.12	0.08	*0.36*
ET adjustment	*0.20*	*0.28*	*0.67*
Lower time constant	0.10	0.09	*0.29*
Subcatchment/reach	**Parameter**		
	a	b	Rainfall multiplier
Crickdale Castle to Pinkhill	0.12	0.17	*0.53*
Pinkhill to Osney	0.14	0.11	*0.27*
Osney to Culham	0.11	0.09	*0.45*
Days Weir	0.12	0.13	*0.26*
Days Weir to Caversham	0.11	0.10	*0.25*
Caversham to Shepperton	0.13	0.11	*0.60*
Mosley	0.13	0.12	0.13
Teddington	0.13	0.08	0.10

Table 7. Significant Pearson's correlations between parameter values from ensemble of credible parameter sets.

Hydrologic response or Reach	Parameter	Hydrologic response or Reach	Parameter	Correlation	*p*
Non-chalk	Upper time constant	Non-chalk	Quick time constant	−0.43	5.6×10^{-19}
Non-chalk	Growing degree threshold	1	Reach rain multiplier	−0.34	6.2×10^{-12}
Non-chalk	Lower time constant	Non-chalk	Drought runoff fraction	−0.26	1.0×10^{-6}
Non-chalk	ET adjustment	1	Reach rain multiplier	−0.21	2.1×10^{-4}
Non-chalk	Drought runoff fraction	Non-chalk	Growing degree threshold	−0.20	4.6×10^{-4}
2	Reach rain multiplier	Non-chalk	Growing degree threshold	−0.20	5.7×10^{-4}
1	Reach rain multiplier	2	Reach rain multiplier	−0.20	7.5×10^{-4}
3	Reach rain multiplier	Non-chalk	Growing degree threshold	−0.20	8.7×10^{-4}
Chalk	ET adjustment	Chalk	Quick time constant	−0.19	1.1×10^{-3}
Non-chalk	ET adjustment	Non-chalk	Growing degree threshold	−0.17	9.0×10^{-3}
Non-chalk	Upper time constant	Chalk	Upper residence time	−0.15	2.7×10^{-2}
1	Reach rain multiplier	Non-chalk	Drought runoff fraction	−0.15	2.8×10^{-2}
8	Reach rain multiplier	2	Reach rain multiplier	−0.15	3.4×10^{-2}
3	b	5	Reach rain multiplier	−0.15	3.8×10^{-2}
3	Reach rain multiplier	8	Reach rain multiplier	−0.15	4.7×10^{-2}
1	b	Non-chalk	Upper time constant	0.20	7.0×10^{-4}
1	a	Non-chalk	Upper time constant	0.26	1.7×10^{-6}
Non-chalk	ET adjustment	Non-chalk	Upper time constant	0.32	2.9×10^{-10}

PERSiST was able to reproduce the observed stream flow dynamics at 8 sites in the river Thames using time series of air temperature and precipitation with data on areas of different hydrologic response types in the catchment. The INCA perceptual model of runoff and flow generation in which precipitation was routed from a quick runoff box into soil water (upper) and groundwater (lower) buckets was adequate for reproducing patterns of observed streamflow in the Thames.

There may be some uncertainty in both meteorologic and hydrologic time series used here. Beven (2012, and references therein) has noted that it can be difficult to accurately estimate high flows from stage height measurements. Model simulations consistently showed better fits at Teddington, the

most downstream gauging site, than at Pinkhill in the headwaters of the catchment (Fig. 5). There are a number of possible reasons for this. The model structure used here may be inappropriate for simulating flows in the upper reaches of the Thames, or there may be problems with input and calibration data. The single rainfall time series used here may be more representative of precipitation patterns lower in the Thames catchment than in the upper reaches. Marsh and Hannaford (2008) report a wide range (600–900 mm yr^{-1}) in precipitation across the Thames Basin. Results from the Monte Carlo analysis show the importance of getting precipitation inputs right for successful hydrological simulation. Subcatchment-specific precipitation multipliers were amongst the most important parameters controlling model performance. This suggests that additional rainfall time series for the upper and lower reaches of the Thames might have improved model performance. The difference between NS and logNS statistics from the ensemble of credible parameter sets shows that high flows were typically better simulated than low flows. It is possible that high flows were more controlled by rainfall while low flows were influenced by the extensive network of locks and abstraction points in the river. Model performance was not sensitive to the drought runoff fraction, suggesting that summer low-flow events do not need to be simulated for credible estimates of runoff. While the low-flow events are not needed for runoff simulations, they are important for solute simulation in the Thames (Jin et al., 2012). Model performance was not sensitive to the flow velocity a and b parameters, suggesting that flow simulations in the Thames are relatively insensitive to estimated flow velocity. It should be noted that flow velocities are an important control on water residence time within a reach, which in turn influence nutrient cycling, especially rates of denitrification (Jin et al., 2012).

The PERSiST application presented here uses a fairly simple structural model where precipitation is routed vertically through three buckets and to the river. This perceptual model of runoff generation has been widely used for simulating surface water nutrient dynamics in the Thames (Jin et al., 2012; Crossman et al., 2013a; Whitehead et al., 2013) and elsewhere (Whitehead et al., 1998; Wade et al., 2002; Futter et al., 2009a, b). Performing a similar analysis to that of Fenicia et al. (2013) or Hrachowitz et al. (2013b) and evaluating alternate model structures, especially in the upper reaches, might improve model performance and reveal additional information about patterns of runoff generation in the Thames catchment.

The existence of a large number of credible parameter sets clearly demonstrates the existence of equifinality due to parameter uncertainty (Beven, 2006). There were numerous correlations between parameter values from the ensemble of credible parameter set, suggesting that different parameters are able to compensate for each other. It was somewhat surprising that model performance was only sensitive to terrestrial parameters from the non-chalk hydrologic response type. Parameter equifinality, lack of parameter sensitivity and correlation between parameter values all suggest that the simulations presented here are over-determined. Both Beven (1989) and Jakeman and Hornberger (1993) suggest that 3–6 parameters can be uniquely identified from typical precipitation and streamflow time series. Incorporating solute time series in model calibration could assist in better constraining model parameterization (e.g. Tetzlaff and Soulsby, 2008; Birkel et al., 2011; Hrachowitz et al., 2013b). Chloride is widely used as a semi-conservative tracer (Shaw et al., 2008; Jin et al., 2011). We must caution against the uncritical use of chloride as a tracer as it is not conservative (Svensson et al., 2012). This may not be problematic when there are significant chloride inputs, but can lead to inappropriate conclusions when chloride is in short supply.

There are both formal (Vrugt et al., 2009) and informal approaches (Beven, 2006) to Monte Carlo analysis. The approach presented here takes an informal approach that may appear overly simplistic but which should be robust to high dimensionality and potentially non-smooth goodness-of-fit response surfaces. It appeared that a chain length of 5000 was sufficient for identifying local optima in parameter space. Typically, the best-performing parameter set was identified within 200 model runs and it was extremely rare to obtain better model fits after 4000 or more runs. As the goal of each MCMC run was to identify a credible as opposed to optimal parameter set, this chain length was deemed adequate. The approach presented here, with an ensemble of locally optimal parameter sets, recognizes the equifinality inherent in catchment-scale hydrological modelling. There is a potentially infinite number of parameter sets all capable of providing credible fits to the observed data. As noted by Beven (2006), this is a common outcome in hydrological modelling. However, ensembles of credible parameter sets give more plausible simulations of present-day hydrology than can be achieved with single parameter sets (Oni et al., 2014). Unlike Laplace's demon, our understanding of the runoff generation process will always be incomplete and multiple competing hypotheses, expressed as combinations of model structures, parameter sets and environmental data may have equal credibility.

The modelling exercise presented here, and hydrologic modelling in general, can be thought of as an inverse problem in which inferences are made about a phenomenon from partial or incomplete information (O'Sullivan, 1986). It seems reasonable to suggest that most inverse problems in environmental modelling are not well posed and would support the conclusion of O'Sullivan (1986), who stated that "in an ill posed inverse problem, a classical least squares... solution may not be uniquely defined". A quest for optimal parameter sets in environmental modelling may be doomed to fail.

There has been considerable discussion in the hydrological literature about the pros and cons of multi-site calibration (Zhang et al., 2008; Cao et al., 2006; Lerat et al., 2012; Gong et al., 2012). Often, modellers are forced to

use single-site calibrations out of necessity as monitoring agencies rarely collect data at internal points in a catchment. The Thames is a fortunate exception to this trend. While we present a model application based on data from 8 gauging station on the main stem of the river, the UK National River Flow Archive maintains data collected at 122 sites throughout the catchment (http://www.ceh.ac.uk/data/nrfa/index.html, last access: 3 May 2013). Because branched river networks can be simulated in PERSiST, it would be possible to simulate flows at all monitoring sites in the catchment. Such an exercise might play a role in evaluating data quality or in identifying hydrologically meaningful differences in parameter values across the catchment. For example, the model simulations presented here used a single base flow index (BFI) for all hydrologic response units. Future model applications could better constrain characteristic time constants if the hydrologic response units incorporated the range of BFI reported for gauging stations in the Thames by Marsh and Hannaford (2008).

The PERSiST modelling framework makes a number of significant simplifying assumptions. Both snowmelt and ET are dependent on energy balances, which may be poorly represented by temperature index approximations. Despite this shortcoming, simple temperature index snowmelt models are widely used and accepted. Hock (2003) notes that, at a catchment scale, temperature index methods can outperform more-data-hungry energy balance snowmelt models. However, temperature index methods can have limited temporal resolution and spatial accuracy (Hock, 2003). The temperature index approach to estimating ET is similar to the one used in HBV (Sælthun, 1996), which has been the subject of some criticism (Andréasson et al., 2004; Lawrence and Haddeland, 2011). However, the temperature- and soil-moisture-dependent approach to estimating actual ET as the difference between precipitation and runoff appears to be robust and has been successfully defended elsewhere (Crossman et al., 2013b). The PERSiST framework assumes that water is transferred instantaneously between buckets. The developers of the SUPERFLEX framework have shown that adding in lag time for water transfers between buckets can improve model performance (Kavetski and Fenicia, 2011; Fenicia et al., 2013). Unlike the INCA approach, which is semi-distributed but does not route water laterally within the terrestrial environment, PERSiST facilitates the simulation of upslope and riparian areas and fluxes of water between them. As PERSiST has been designed primarily for simulating surface water fluxes, it uses a fairly simplistic representation of groundwater and does not currently include the capacity for loss or input of deep groundwater from outside the catchment.

PERSiST is based on buckets which represent dual-porosity reservoirs in which water is divided into stagnant and freely draining fractions (Šimůnek et al., 2003). By varying the relative depths of the freely draining and retained fraction, it is possible to simulate catchment-scale transit times. Following Soulsby et al. (2009), we define transit time as a measure of the time elapsed between a water molecule entering and leaving a catchment, and mean transit time (MTT) as the total storage divided by the flux of water, Thus, at steady state MTT will be equal to the depth of water in a bucket (L) divided by runoff from the bucket (L/T). The characteristic time constant in PERSiST does not immediately provide insight into transit times, but instead describes the behaviour of the falling limb of the hydrograph. In PERSiST, the transit time distribution is determined from characteristic time constants, storage volumes and the rate at which water enters the system. Storage volume is related to the depth of water in a bucket. It is assumed that the stagnant and freely draining fractions are well mixed; thus water molecules in both the freely draining and retained fractions will eventually leave the bucket. The characteristic time constant only applies to water in the freely draining fraction. With PERSiST, it is possible to simulate arbitrarily long MTT by using large values for the depth of retained water. It should be noted that this approach can only be tested using conservative tracers and assumes instantaneous and complete mixing of water between the stagnant and freely draining fractions. Hrachowitz et al. (2013b) have shown that this assumption may be too simplistic in many cases and future versions of PERSiST should incorporate some form of partial mixing.

Fluxes of water in PERSiST are currently simulated in an ad hoc manner. As noted by Fenicia et al. (2011), this is a shortcoming common to a number of hydrological models. A more appropriate implementation of PERSiST would be dependent on a series of linked first-order differential equations (ODEs). There is a comprehensive literature about the shortcomings of poorly chosen ODE solvers in hydrological modelling (Kavetski and Clark, 2010; Clark and Kavetski, 2010). We must caution that a perception of an appropriate ODE solver as a panacea for numerical problems can lead to difficulties of interpretation and communication. Most rainfall-runoff models, PERSiST included, represent snowmelt and accumulation. Because the model switches between snow accumulation and melt depending on air temperature, the function to be integrated becomes discontinuous. This can lead to challenges in the numerical solution (Hairer et al., 2009, 196–200) which do not appear to be adequately appreciated in the hydrological modelling community. Similarly, the question of how to calibrate models based on simultaneous solution of sets of ODEs must be approached with care. While models are typically calibrated against a time series of mean daily flows, there can be significant sub-daily variation in flows (for a graphic example of this comparing 15 min and daily average flows, see Baggaley et al., 2009). Ideally, calibration would not be based on a single point in time arbitrarily sampled from ODE output but would aggregate model outputs so as to be compatible with the time step of the flow observations used for model calibration.

We do not propose PERSiST as a universal solution to the rainfall-runoff modelling problem. While the model has been

successfully applied to temperate, boreal and Mediterranean rivers, we suspect that PERSiST might have difficulty in credibly simulating snowmelt from mountainous catchments with large elevational gradients where a full energy balance approach would be more appropriate or from extremely flat agricultural areas with large amounts of artificial drainage. In light of the results presented here, we believe that there is an unfilled niche in the rainfall-runoff model ecosystem and that PERSiST is a useful tool for addressing some rainfall-runoff modelling problems, specifically those having to do with the simulation of water quality time series derived from long-term monitoring data. As stated in the name, one goal of PERSiST is improved hydrological simulations for solute transport. It is widely accepted by the hydrological modelling community that tracer data can improve the fit and credibility of hydrological models (Tetzlaff and Soulsby, 2008). In some ways, the modelling of pollutant fate and transport can be conceptualized as the use of non-conservative tracers to improve hydrological understanding. Simulating pollutant transport can help to constrain hydrological model predictions, but, more importantly, use of appropriate hydrological model structures can aid in understanding the mechanisms behind pollutant transport. It is hoped that the PERSiST model makes a contribution – albeit small – to achieving this goal.

Acknowledgements. This paper has been greatly improved by the constructive criticisms of the editor, M. Hrachowitz, and three anonymous reviewers. Parts of this work were funded by the MISTRA Future Forests programme, the FORMAS ForWater strong research environment and the EU project REFRESH (contract number 244121). We thank L. Jin for providing GIS data. Additional GIS data were obtained from the European Environment Agency and British Geological Survey. Meteorological data were obtained from the UK Environment Agency. This work would not have been possible without the unique data set managed and made available through the UK National River Flow Archive.

Edited by: M. Hrachowitz

References

Andersen, H. E., Kronvang, B., Larsen, S. E., Hoffmann, C. C., Jensen, T. S., and Rasmussen, E. K.: Climate-change impacts on hydrology and nutrients in a Danish lowland river basin, Sci. Total Environ., 365, 223–237, 2006.

Andersson, L., Rosberg, J., Pers, B. C., Olsson, J., and Arheimer, B.: Estimating catchment nutrient flow with the HBV-NP model: sensitivity to input data, Ambio, 34, 521–532, 2005.

Andréasson, J., Bergström, S., Carlsson, B., Graham, L. P., and Lindström, G.: Hydrological change-climate change impact simulations for Sweden, Ambio, 33, 228–234, 2004.

Baggaley, N. J., Langan, S. J., Futter, M. N., Potts, J. M., and Dunn, S. M.: Long-term trends in hydro-climatology of a major Scottish mountain river, Sci. Total Environ., 407, 4633–4641, 2009.

Bayley, P. B.: Understanding large river: floodplain ecosystems, BioScience, 45, 153–158, 1995.

Bergström, S. and Singh, V.: The HBV model, in: Computer Models of Watershed Hydrology, edited by: Singh, V. P., Water Resources Publications, Highlands Ranch, CO, 443–476, 1995.

Bernal, S., Butturini, A., Riera, J. L., Vázquez, E., and Sabater, F.: Calibration of the INCA model in a Mediterranean forested catchment: the effect of hydrological inter-annual variability in an intermittent stream, Hydrol. Earth Syst. Sci., 8, 729–741, doi:10.5194/hess-8-729-2004, 2004.

Beven, K.: Changing ideas in hydrology – the case of physically-based models, J. Hydrol., 105, 157–172, 1989.

Beven, K.: A manifesto for the equifinality thesis, J. Hydrol., 320, 18–36, 2006.

Beven, K.: Causal models as multiple working hypotheses about environmental processes, C. R. Geosci., 344, 77–88, 2012.

Beven, K., Lamb, R., Quinn, P., Romanowicz, R., Freer, J., and Singh, V.: Topmodel, in: Computer models of watershed hydrology, edited by: Singh, V. P., Water Resources Publication, Colorado, 627–668, 1995.

Birkel, C., Tetzlaff, D., Dunn, S. M., and Soulsby, C.: Using lumped conceptual rainfall-runoff models to simulate daily isotope variability with fractionation in a nested mesoscale catchment, Adv. Water Resour., 34, 383–394, 2011.

Bloomfield, J. P., Allen, D. J., and Griffiths, K. J.: Examining geological controls on baseflow index (BFI) using regression analysis: An illustration from the Thames Basin, UK, J. Hydrol., 373, 164–176, 2009.

Burnash, R. J., Ferral, R. L., and McGuire, R. A.: A generalized streamflow simulation system, conceptual modeling for digital computers, Report by the Joliet Federal State River Forecasts Center, Sacramento, CA, 204 pp., 1973.

Cao, W., Bowden, W. B., Davie, T., and Fenemor, A.: Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability, Hydrol. Process., 20, 1057–1073, 2006.

Chib, S. and Greenberg, E.: Understanding the Metropolis-Hastings algorithm, Am. Stat., 49, 327–335, 1995.

Clark, M. P. and Kavetski, D.: Ancient numerical daemons of conceptual hydrological modeling: 1. Fidelity and efficiency of time stepping schemes, Water Resour. Res., 46, W10510, doi:10.1029/2009WR008894, 2010.

Crossman, J., Whitehead, P. G., Futter, M. N., Jin, L., Shahgedanova, M., Castellazzi, M., and Wade, A. J.: The interactive responses of water quality and hydrology to changes in multiple stressors, and implications for the long-term effective management of phosphorus, Sci. Total Environ., 454, 230–244, 2013a.

Crossman, J., Futter, M. N., Oni, S., Whitehead, P. G., Jin, L., Butterfield, D., Baulch, H., and Dillon, P.: Impacts of climate change on hydrology and water quality: Future proofing management strategies in the Lake Simcoe watershed, Canada, J. Great. Lakes Res., 39, 19–32, 2013b.

Durand, P.: Simulating nitrogen budgets in complex farming systems using INCA: calibration and scenario analyses for the Kervidy catchment (W. France), Hydrol. Earth Syst. Sci., 8, 793–802, doi:10.5194/hess-8-793-2004, 2004.

Fenicia, F., Savenije, H. H. G., Matgen, P., and Pfister, L.: Is the groundwater reservoir linear? Learning from data in hydrological modelling, Hydrol. Earth Syst. Sci., 10, 139–150, doi:10.5194/hess-10-139-2006, 2006.

Fenicia, F., Kavetski, D., and Savenije, H. H.: Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development, Water Resour. Res., 47, W11510, doi:10.1029/2010WR010174, 2011.

Fenicia, F., Kavetski, D., Savenije, H. H., Clark, M. P., Schoups, G., Pfister, L., and Freer, J.: Catchment properties, function, and conceptual model representation: is there a correspondence?, Hydrol. Process., doi:10.1002/hyp.9726, 2013.

Futter, M. N., Butterfield, D., Cosby, B. J., Dillon, P. J., Wade, A. J., and Whitehead, P. G.: Modeling the mechanisms that control in-stream dissolved organic carbon dynamics in upland and forested catchments, Water Resour. Res, 43, W02424, doi:10.1029/2006WR004960, 2007.

Futter, M. N., Forsius, M., Holmberg, M., and Starr, M.: A long-term simulation of the effects of acidic deposition and climate change on surface water dissolved organic carbon concentrations in a boreal catchment, Hydrol. Res., 40, 291–305, 2009a.

Futter, M. N., Helliwell, R., Hutchins, M., Aherne, J., and Whitehead, P. G.: Modelling the effects of changing climate and nitrogen deposition on nitrate dynamics in a Scottish mountain catchment, Hydrol. Res., 40, 153–166, 2009b.

Futter, M. N., Poste, A. E., Butterfield, D., Dillon, P. J., Whitehead, P. G., Dastoor, A. P., and Lean, D. R. S.: Using the INCA-Hg model of mercury cycling to simulate total and methyl mercury concentrations in forest streams and catchments, Sci. Total Environ., 424, 219–231, 2012.

Gong, Y., Shen, Z., Liu, R., Hong, Q., and Wu, X.: A comparison of single-and multi-gauge based calibrations for hydrological modeling of the Upper Daning River Watershed in China's Three Gorges Reservoir Region, Hydrol. Res., 43, 822–832, 2012.

Grischek, T., Hiscock, K., Metschies, T., Dennis, P., and Nestler, W.: Factors affecting denitrification during infiltration of river water into a sand and gravel aquifer in Saxony, Germany, Water Res., 32, 450–460, 1998.

Gupta, H. V., Clark, M. P., Vrugt, J. A., Abramowitz, G., and Ye, M.: Towards a comprehensive assessment of model structural adequacy, Water Resour. Res., 48, WR08301, do1:10.1029/2011WRR011044, 2012.

Hairer, E., Nørsett, S. P., and Wanner, G.: Solving Ordinary Differential Equations I: Nonstiff Problems, 2nd Edn., Springer, Berlin, 2009.

Hellebrand, H., Müller, C., Matgen, P., Fenicia, F., and Savenije, H. H. G.: A process proof test for model concepts: modelling the meso-scale, Phys. Chem. Earth, 36, 42–53, 2011.

Hock, R.: Temperature index melt modelling in mountain areas, J. Hydrol., 282.1, 104–115, 2003.

Hough, M. N. and Jones, R. J. A.: The United Kingdom Meteorological Office rainfall and evaporation calculation system: MORECS version 2.0-an overview, Hydrol. Earth Syst. Sci., 1, 227–239, doi:10.5194/hess-1-227-1997, 1997.

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P. A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Woods, R. A., Zehe, E., and Cudennec, C.: A decade of Predictions in Ungauged Basins (PUB) A review, Hydrol. Sci. J., 58, 1198–1255, doi:10.1080/02626667.2013.803183, 2013a.

Hrachowitz, M., Savenije, H., Bogaard, T. A., Tetzlaff, D., and Soulsby, C.: What can flux tracking teach us about water age distribution patterns and their temporal dynamics?, Hydrol. Earth Syst. Sci., 17, 533–564, doi:10.5194/hess-17-533-2013, 2013b.

Jakeman, A. J. and Hornberger, G. M.: How much complexity is warranted in a rainfall-runoff model?, Water Resour. Res., 29, 2637–2649, 1993.

Jakeman, A. J., Littlewood, I., and Whitehead, P. G.: Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments, J. Hydrol., 117, 275–300, 1990.

Jin, L., Whitehead, P. G., Siegel, D. I., and Findlay, S.: Salting our landscape: An integrated catchment model using readily accessible data to assess emerging road salt contamination to streams, Environ. Pollut., 159, 1257–1265, 2011.

Jin, L., Whitehead, P. G., Futter, M. N., and Lu, Z.: Modelling the impacts of climate change on flow and nitrate in the River Thames: assessing potential adaptation strategies, Hydrol. Res., 43, 902–916, 2012.

Kampf, S. K. and Burges, S. J.: A framework for classifying and comparing distributed hillslope and catchment hydrologic models, Water Resour. Res., 43, W05423, doi:10.1029/2006WR005370, 2007.

Kavetski, D. and Clark, M. P.: Ancient numerical daemons of conceptual hydrological modeling: 2. Impact of time stepping schemes on model analysis and prediction, Water Resour. Res., 46, W10511, doi:10.1029/2009WR008896, 2010.

Kavetski, D. and Fenicia, F.: Elements of a flexible approach for conceptual hydrological modeling: 2. Application and experimental insights, Water Resour. Res., 47, W11511, doi:10.1029/2011WR010748, 2011.

Lawrence, D. S. L. and Haddeland, I.: Uncertainty in hydrological modelling of climate change impacts in four Norwegian catchments, Hydrol. Res., 42, 457–471, 2011.

Lazar, A. N., Butterfield, D., Futter, M. N., Rankinen, K., Thouvenot-Korppoo, M., Jarritt, N., Lawrence, D. S. L., Wade, A. J., and Whitehead, P. G.: An assessment of the fine sediment dynamics in an upland river system: INCA-Sed modifications and implications for fisheries, Sci. Total Environ., 408, 2555–2566, 2010.

Ledesma, J., Köhler, S. J., and Futter, M. N.: Long-term trends of dissolved organic carbon: implications for drinking water supply, Sci. Total Environ., 432, 1–11, doi:10.1016/j.scitotenv.2012.05.071, 2012.

Lerat, J., Andréassian, V., Perrin, C., Vaze, J., Perraud, J.-M., Ribstein, P., and Loumagne, C.: Do internal flow measurements improve the calibration of rainfall-runoff models?, Water Resour. Res., 48, W02511, doi:10.1029/2010WR010179, 2012.

Lindström, G., Pers, C., Rosberg, J., Strömqvist, J., and Arheimer, B.: Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales, Hydrol. Res., 41, 295–319, 2010.

Löfgren, S., Aastrup, M., Bringmark, L., Hultberg, H., Lewin-Pihlblad, L., Lundin, L., Karlsson, G. P., and Thunholm, B.: Recovery of soil water, groundwater, and streamwater from acidification at the Swedish Integrated Monitoring catchments, Ambio, 40, 836–856, 2011.

Marsh, T. J. and Hannaford, J. (Eds): UK Hydrometric Register. Hydrological data UK series, Centre for Ecology & Hydrology, Wallingford, UK, 210 pp., 2008.

Medici, C., Butturini, A., Bernal, S., Vázquez, E., Sabater, F., Vélez, J., and Francés, F.: Modelling the non-linear hydrological behaviour of a small Mediterranean forested catchment, Hydrol. Process., 22, 3814–3828, 2008.

Medici, C., Bernal, S., Butturini, A., Sabater, F., Martin, M., Wade, A. J., and Frances, F.: Modelling the inorganic nitrogen behaviour in a small Mediterranean forested catchment, Fuirosos (Catalonia), Hydrol. Earth Syst. Sci., 14, 223–237, doi:10.5194/hess-14-223-2010, 2010.

Nash, J. and Sutcliffe, J.: River flow forecasting through conceptual models part I – A discussion of principles, J. Hydrol., 10, 282–290, 1970.

O'Sullivan, F.: A statistical perspective on ill-posed inverse problems, Stat. Sci., 1, 502–518, 1986.

Oni, S. K., Futter, M. N., Molot, L. A., and Dillon, P. J.: Adjacent catchments with similar patterns of land use and climate have markedly different DOC and runoff dynamics, Hydrol. Process., 28, 1436–1449, doi:10.1002/hyp.9681, 2013.

Oni, S. K., Futter, M. N., Molot, L. A., Dillon, P. J., and Crossman, J.: Uncertainty assessments and hydrological implications of climate change in two adjacent agricultural catchments of a rapidly urbanizing watershed, Sci. Tot Environ., 473–474, 326–337, 2014.

Rankinen, K., Lepistö, A., and Granlund, K.: Integrated nitrogen and flow modelling (INCA) in a boreal river basin dominated by forestry: scenarios of environmental change, Water Air Soil Poll. Focus, 4, 161–174, 2004.

Ranzini, M., Forti, M. C., Whitehead, P. G., Arcova, F. C. S., de Cicco, V., and Wade, A. J.: Integrated Nitrogen CAtchment model (INCA) applied to a tropical catchment in the Atlantic Forest, São Paulo, Brazil, Hydrol. Earth Syst. Sci., 11, 614–622, doi:10.5194/hess-11-614-2007, 2007.

Sælthun, N.: The "Nordic" HBV-model, Norwegian Water Resources and Energy Administration, Oslo, Norway, 26 pp., 1996.

Seibert, J. and McDonnell, J. J.: On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, Water Resour. Res., 38, 1241, doi:10.1029/2001WR000978, 2002.

Seibert, J. and Vis, M. J. P.: Teaching hydrological modeling with a user-friendly catchment-runoff-model software package, Hydrol. Earth Syst. Sci., 16, 3315–3325, doi:10.5194/hess-16-3315-2012, 2012.

Shaw, S., Harpold, A. A., Taylor, J. C., and Walter, M. T.; Investigating a high resolution, stream chloride time series from the Biscuit Brook catchment, Catskills, NY, J. Hydrol., 348, 245–256, 2008.

Šimůnek, J., Jarvis, N. J., Van Genuchten, M. T., and Gärdenäs, A.; Review and comparison of models for describing non-equilibrium and preferential flow and transport in the vadose zone, J. Hydrol., 272, 14–35, 2003.

Smith, A. F. and Roberts, G. O.: Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods, J. Roy. Stat. Soc. B. Met., 55, 3–23, 1993.

Soulsby, C., Tetzlaff, D., and Hrachowitz, M.: Tracers and transit times: windows for viewing catchment scale storage?, Hydrol. Process., 23, 3503–3507, 2009.

Stutter, M. I., Langan, S. J., and Lumsdon, D. G.: Vegetated buffer strips can lead to increased release of phosphorus to waters: a biogeochemical assessment of the mechanisms, Environ. Sci. Technol., 43, 1858–1863, 2009.

Svensson, T., Lovett, G. M., and Likens, G. E.: Is chloride a conservative ion in forest ecosystems?, Biogeochemistry, 107, 125–134, 2012.

Tetzlaff, D. and Soulsby, C.: Sources of baseflow in larger catchments – Using tracers to develop a holistic understanding of runoff generation, J. Hydrol., 359, 287–302, 2008.

Uhlenbrook, S. and Sieber, A.: On the value of experimental data to reduce the prediction uncertainty of a process-oriented catchment model, Environ. Modell. Softw., 20, 19–32, 2005.

Uhlenbrook, S., Roser, S., and Tilch, N.: Hydrological process representation at the meso-scale: the potential of a distributed, conceptual catchment model, J. Hydrol., 291, 278–296, 2004.

Van der Velde, Y., Torfs, P. J. J. F., van der Zee, S. E. A. T. M., and Uijlenhoet, R.: Quantifying catchment-scale mixing and its effects on time-varying travel time distributions, Water Resour. Res., 48, W06536, doi:10.1029/2011WR011310, 2012.

Vehviläuinen, B.: Hydrological Forecasting and Real-Time Monitoring: The Watershed Simulation and Forecasting System (WSFS), Water Quality Measurements Series: Hydrological and Limnological Aspects of Lake Monitoring, 13–20, 2007.

Vrugt, J. A., Ter Braak, C. J., Gupta, H. V., and Robinson, B. A.: Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?, Stoch. Env. Res. Risk A, 23, 1011–1026, 2009.

Wade, A. J., Neal, C., Soulsby, C., Langan, S., and Smart, R.: On modelling the effects of afforestation on acidification in heterogeneous catchments at different spatial and temporal scales, J. Hydrol., 250, 149–169, 2001.

Wade, A. J., Durand, P., Beaujouan, V., Wessel, W. W., Raat, K. J., Whitehead, P. G., Butterfield, D., Rankinen, K., and Lepisto, A.: A nitrogen model for European catchments: INCA, new model structure and equations, Hydrol. Earth Syst. Sci., 6, 559–582, doi:10.5194/hess-6-559-2002, 2002.

Wagener, T., Boyle, D. P., Lees, M. J., Wheater, H. S., Gupta, H. V., and Sorooshian, S.: A framework for development and application of hydrological models, Hydrol. Earth Syst. Sci., 5, 13–26, doi:10.5194/hess-5-13-2001, 2001.

Whitehead, P. G., Wilson, E., and Butterfield, D.: A semi-distributed Integrated Nitrogen model for multiple source assessment in Catchments (INCA): Part I – model structure and process equations, Sci. Total Environ., 210, 547–558, 1998.

Whitehead, P. G., Crossman, J., Balana, B. B., Futter, M. N., Comber, S., Jin, L., Skuras, D., Wade, A. J., Bowes, M. J., and Read, D. S.: A Cost Effectiveness Analysis of Water Security and Water Quality: Impacts of Climate and Land Use Change on the River Thames System, Philos. T. Roy. Soc. A, 371, 20120413, doi:10.1098/rsta.2012.0413, 2013.

Zhang, X., Srinivasan, R., and Van Liew, M.: Multi-site calibration of the SWAT model for hydrologic modeling, T. ASABE, 51, 2039–2049, 2008.