



Diagnosing hydrological limitations of a land surface model: application of JULES to a deep-groundwater chalk basin

N. Le Vine¹, A. Butler¹, N. McIntyre^{1,2}, and C. Jackson³

¹Department of Civil and Environmental Engineering, Imperial College London, London, UK

²Centre for Water in the Minerals Industry, the University of Queensland, St. Lucia, Australia

³British Geological Survey, Keyworth, UK

Correspondence to: N. Le Vine (nlevine@imperial.ac.uk)

Received: 29 June 2015 – Published in Hydrol. Earth Syst. Sci. Discuss.: 7 August 2015

Revised: 9 November 2015 – Accepted: 27 November 2015 – Published: 18 January 2016

Abstract. Land surface models (LSMs) are prospective starting points to develop a global hyper-resolution model of the terrestrial water, energy, and biogeochemical cycles. However, there are some fundamental limitations of LSMs related to how meaningfully hydrological fluxes and stores are represented. A diagnostic approach to model evaluation and improvement is taken here that exploits hydrological expert knowledge to detect LSM inadequacies through consideration of the major behavioural functions of a hydrological system: overall water balance, vertical water redistribution in the unsaturated zone, temporal water redistribution, and spatial water redistribution over the catchment's groundwater and surface-water systems. Three types of information are utilized to improve the model's hydrology: (a) observations, (b) information about expected response from regionalized data, and (c) information from an independent physics-based model. The study considers the JULES (Joint UK Land Environmental Simulator) LSM applied to a deep-groundwater chalk catchment in the UK. The diagnosed hydrological limitations and the proposed ways to address them are indicative of the challenges faced while transitioning to a global high resolution model of the water cycle.

damental weaknesses in the simulation of hydrology. Hydrology (as well as other soil–vegetation–atmosphere interactions) in climate models is represented via land surface models (LSMs) that partition water between evapotranspiration, surface runoff, drainage, and soil moisture storage. The deficiencies in hydrological processes representation lead to incorrect energy and water partitioning at the land surface (Oleson et al., 2008) that propagates into precipitation and near-surface air temperature biases in climate model predictions (Lawrence and Chase, 2008). Furthermore, improving the representation of hydrology is a step towards the development of a global hyper-resolution model for monitoring the terrestrial water, energy, and biogeochemical cycles that is considered as one of the *grand challenges* to the community (Wood et al., 2011).

The most recent third generation LSMs operate in a continuous time and distributed space mode, and simulate exchanges of energy, water, and carbon between the land surface and the atmosphere using physics-based process descriptions (Pitman, 2003). The physics-based nature of third generation LSMs allows widely available global data sets, such as soil properties, land use, weather states, etc., to be used as model parameters and inputs, thus making predictive modelling with LSMs very appealing.

A significant body of literature exists on LSM hydrology assessment and inter-comparison, including comparison with observed point scale evapotranspiration fluxes, soil moisture, observed river flow rates and depths to groundwater (Balsamo et al., 2009; Blyth et al., 2011; Boone et al., 2004; Lohmann et al., 2004; Maxwell and Miller, 2005). Blyth et al. (2011) used point-scale evapotranspiration fluxes from

1 Introduction

Guidance to support adaptation to the changing water cycle is urgently required, yet the ability of water cycle models to represent the hydrological impacts of climate change is limited in several important respects. Climate models are an essential tool in scenario development, but suffer from fun-

10 FLUXNET observation sites covering the major global biomes as well as river flows from seven large rivers to assess the performance of the JULES (Joint UK Land Environmental Simulator) model. The evaluation used monthly average fluxes, over a period of 10 years, and demonstrated a number of model weaknesses in energy partitioning as well as in water partitioning and routing, thus providing a direction for further model improvements. Balsamo et al. (2009) revised the soil representation in the TESSEL LSM (used by the European Centre for Medium-Range Weather Forecasts) and showed better agreement of the new H-TESSEL model with soil moisture point observations from the Global Soil Moisture Bank. Lohmann et al. (2004) evaluated four LSMs coupled to a surface runoff routing model over 1145 small and medium size basins in the USA, and found that “the modeled mean values of the water balance terms are of the same magnitude as the spread of the models around them”. The authors name both parameter selection and model structure improvements as the key factors to achieve better model performance for hydrological predictions.

LSMs focus on modelling processes in the near-surface layer (typically, the top three metres). Typically, a unit gradient (free drainage) or other simple lower boundary condition is generally assumed in place of explicitly representing the groundwater boundary (e.g. Best et al., 2011; Kriner et al., 2005; Yang and Niu, 2003). However, in permeable basins the depth to the water table is often much deeper, for example in the Kennet case study, introduced below, this can be as much as 100 m (Jackson et al., 2008), calling into question the adequacy of a relatively shallow lower boundary condition. This can result in unrealistically dry lower soil layers (e.g. Li et al., 2008). To address this problem, the NASA Seasonal-to-Interannual Prediction Project Catchment Model (NSIPP) model relies on an approximate relationship (derived from detailed simulations) to estimate the soil moisture transfer rate between the root zone and water table at a catchment scale (Koster et al., 2000). In contrast, Community Land Model (CLM) uses the hydraulic gradient between the bottom of the soil column and the water table to approximate the drainage rate from the soil column (Oleson et al., 2008). Another approach is to use the location of the water table as a lower boundary condition. The soil, water, atmosphere and plant (SWAP) model uses a variable depth soil column, whose base is located at the water table (Gusev and Nasonova, 2003). Maxwell and Miller (2005) developed this further by coupling CLM to a physics-based 3-D groundwater model ParFlow (Kollet and Maxwell, 2008) at the land surface, replacing the soil column/root-zone soil moisture formulation in CLM with the ParFlow formulation. They concluded that the resulting model provided reasonable predictions for runoff rates and shallow groundwater levels on monthly time steps. However, the explicit inclusion of the deep unsaturated zone requires the estimation of hydraulic properties that are generally not included in existing soil databases.

The tendency for LSMs to use relatively shallow soil column depths and a simplistic or non-existent representation of groundwater also questions their applicability to catchments with deep-groundwater systems (where an average water table is tens of metres deep). Such systems represent a major storage of water and their interaction with the unsaturated zone can influence river flows, soil moisture, and evapotranspiration rates (Maxwell and Miller, 2005). Consequently, the addition of a groundwater modelling capability into LSMs will not only address these issues, but also be a step forward for multi-purpose modelling (e.g. representing groundwater levels for water resources) (Wood et al., 2011).

Most LSMs assume a 1-D vertical flow in a soil column neglecting lateral flow (e.g. Kriner et al., 2005; Gusev and Nasonova, 2003). Although this assumption is sufficiently accurate only for soils that are relatively homogeneous in horizontal and vertical directions (Protopapas and Bras, 1991), it is a fairly common feature for LSMs that employ a gridded surface representation. A further complicating factor is that 1-D flow is usually described in physics-based LSMs using Richards' equation, which was derived at the point scale and used to represent single permeability, single porosity soils. The validity of this is questionable for a wide variety of soils, particularly at larger scales (Beven and Germann, 2013). Chalk is an example of a soil–rock system that consists of both matrix and fractures, whose properties are significantly distinct from each other, forming a dual porosity, dual-permeability system (Price et al., 1993). Therefore, a traditional single domain soil water representation is unsuitable to adequately characterize its properties (Ireson et al., 2009). To the best of our knowledge, there is no currently operational LSM that is capable of realistically representing such dual porosity, dual porosity-dual permeability.

Another important challenge in improving hydrological fluxes in LSMs is the representation of surface and near-surface heterogeneity, in particular how it affects partitioning between surface runoff, evaporation and infiltration. For example, 15 LSMs and a two-layer conceptual hydrological model were used to represent river discharge in the Rhone, one of Europe's major basins, in the Rhone-Aggregation Land Surface Scheme Inter-comparison Project (Boone et al., 2004); it was concluded that an LSM's ability to provide a good performance for daily discharge simulation is linked to their ability to generate sub-grid runoff, that is, to the representation of top-soil heterogeneity.

In light of these concerns, the scope of the study is to assess the hydrological behaviour of a typical third-generation LSM, the JULES, in a comprehensive and consistent way and adapt the model accordingly. For this, an evaluation strategy focusses on the primary functions of a hydrological system in a hierarchical way. While other alternatives exist (Black, 1997; Wagener et al., 2007), the following four hydrological functions are considered (Yilmaz et al., 2008): (1) to maintain an overall water balance (i.e. water partitioning between different water cycle components), (2) to redistribute

water vertically through the soil, (3) to redistribute water in time, and (4) to redistribute water spatially over the catchment's groundwater and surface-water systems. The hierarchical evaluation strategy (or diagnosis) allows for inferences to be made about the specific aspects of the model structure that are causing the problems via targeted evaluations of the model response. The diagnostic evaluation makes use of multiple measures of model performance that are relevant for each of the four functions evaluated. When model performance is poor in a particular hydrological aspect, model modifications are based on hydrological expert knowledge that, whilst subjective, is the only currently available way to adjust the model. The Kennet catchment in southern England is chosen as a complex case study that represents a number of the modelling challenges; however, the methodology and the results are of interest beyond this study due to the similarities across the hydrological modules of different third-generation LSMs, and also the broad importance of chalk aquifers and deep-groundwater systems (Brouyere et al., 2004; Downing et al., 1993; Kloppmann et al., 1998; Pinault et al., 2005; Dahan et al., 1998, 1999; Nativ and Nissim, 1992; Nativ et al., 1995).

2 Case study

2.1 The Joint UK Land Environmental Simulator

JULES is a community land surface model, based upon the established UK Met Office Surface Exchange Scheme (MOSES) (Cox et al., 1999). In addition to representing the exchange of fluxes of heat and moisture between the land surface and the atmosphere, the model also represents fluxes of carbon and some other gases, such as ozone and methane (Clark et al., 2011). It includes linked processes of photosynthesis and evaporation, soil and snow physics as well as plant growth and soil microbial activity. These processes are all linked through a series of equations that quantify how soil moisture and temperature govern evapotranspiration, energy balance, respiration, photosynthesis, and carbon assimilation (Best et al., 2011; Clark et al., 2011). JULES includes multi-layer, finite-difference models of subsurface heat and water fluxes, as described in Cox et al. (1999). There are options for the specification of the hydraulic and thermal characteristics, the representation of soil moisture and the subsurface heterogeneity of soil properties (for more details see Best et al., 2011). JULES can be used as a stand alone land surface model driven by observed forcing data, or can be coupled to an atmospheric model (for example, the UK Met Office Unified Model). The model runs at a sub-daily time step, using meteorological drivers of rainfall, incoming radiation, temperature, humidity, and wind speed. When meteorological data have coarser temporal resolution than required by the model, the standard model (version 2.2) disaggregates the data as constant values.

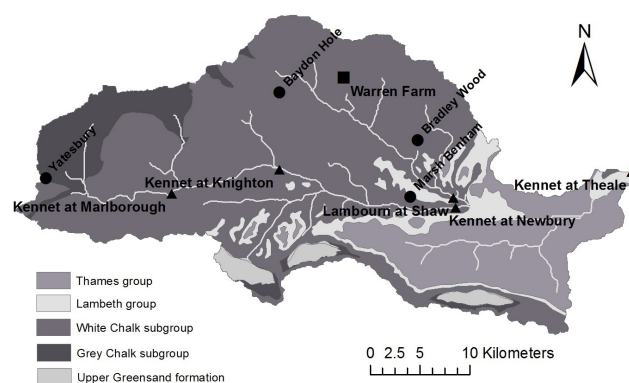


Figure 1. Hydrogeological map of the Kennet catchment. The square indicates the Warren Farm site, the triangles are flow gauging stations, and the circles are observational boreholes.

JULES is typically employed with a 3 m fixed depth of soil, a unit hydraulic head gradient lower boundary condition, and no groundwater component. Shallow groundwater can be optionally represented via the (topography-based model) TOPMODEL approach (Beven and Kirkby, 1979; Clark and Gedney, 2008). Further, top-soil heterogeneity can be optionally represented via a probability distributed model (PDM) (Moore, 2007; Clark and Gedney, 2008). Both options require specification of parameters that are conceptual in nature and are not directly related to the existing data on soil/vegetation properties. JULES is able to generate an infiltration-excess (when PDM is used, or when rainfall intensity exceeds the near-surface infiltration rate) as well as saturation-excess (when the TOPMODEL is used, or through the upward movement of water from saturated soil layers) surface runoff.

The study uses and implements modifications to JULES version 2.2, termed the standard JULES. The standard set-up is used with a 3 m depth of soil, four soil layers: 0.1, 0.25, 0.65, and 2 m deep, starting from the surface. The model is spun-up over 3 years, repeating the weather inputs for the first year of available data 3 times (one of the model warming-up options provided), and initialising soils with saturated conditions.

2.2 Case study catchment

The Kennet is a groundwater-dominated catchment in southern England (Fig. 1). The topographic catchment has an area of 1030 km² with an annual average rainfall of 759 mm (1961–1990). It is predominantly a permeable catchment (Upper Cretaceous Chalk). The western and northern parts of the catchment have exposed bedrock with only a thin, permeable soil. However, in the southern and eastern parts of the catchment there is significant drift cover, and, in its lowest quarter, it is largely impermeable due to overlying Palaeogene deposits (Fig. 1). It is a primarily rural catchment with scattered settlements. The flow regime is domi-

Table 1. Data used for JULES set-up and performance evaluation.

JULES input type	Data description	Source
Catchment grid	(1) 50 m resolution raster file (2) catchment outlet coordinates	(1) http://digimap.edina.ac.uk/ (2) http://www.environment-agency.gov.uk/hiflows/station.aspx?39016
Land use	50 m IGBP 2007 reclassified from 17 IGBP classes to 9 JULES classes (Smith et al., 2006)	MODIS land cover product: http://webmap.ornl.gov/wcsdown/dataset.jsp?ds_id=10004
Soil properties	1 km NSRI soil maps (Brooks and Corey parameterization)	The Cranfield Soil and AgriFood Institute: http://www.landis.org.uk/data/
Meteorological inputs	Daily, 1 km CHES data, 1971–2007	E. Blyth, personal communication, 2012 with CEH, UK
Observations	(1) Soil moisture and soil matric potential measurements at Warren Farm, 2003–2006 (2) automatic weather station data at Warren Farm, 2002–2009 (3) Daily river flow data (4) Groundwater levels at observation boreholes	(1) N. Hewitt, personal communication, 2011 with CEH, UK, and LO-CAR project data (2) N. Hewitt, personal communication, 2011 with CEH, UK, and LO-CAR project data (3) http://www.ceh.ac.uk/data/nrfa/data/search.html (4) National Groundwater Level Archive: http://www.bgs.ac.uk/research/groundwater/datainfo/levels/ngla.html

nated by the slow response of the groundwater held within the chalk aquifer (the base flow index, that is, the proportion of total flow as base flow, is 0.87). Where the chalk outcrops, there is generally little surface runoff. At the Chalk–Palaeogene boundary, surface runoff from low permeability deposits gives rise to focussed recharge into the chalk. As a consequence, there are a number of swallow holes in the area (West and Dumbleton, 1972) that serve as surface-water sinks. The flow at the catchment outlet at Theale is monitored using a crump profile weir, where bypassing of the weir occurs above $29 \text{ m}^3 \text{ s}^{-1}$. The unsaturated zone of the chalk has two characteristic behaviours: slow drainage over summer, and bypass flow during rainfall events (Ireson and Butler, 2013). Both behaviours are important under extreme conditions (i.e. droughts or extreme rainfall) for sustaining river flows and rapid water table response.

2.3 Case study data sets

A number of gridded data types are required for JULES parameterization and forcing (Table 1), including land cover and soil profile data, and meteorological drivers. Using a 50 m resolution topographic map, the Kennet catchment is discretized into 1 km^2 grids, which matches the resolution of the soil and meteorological data. Soil property data are provided by the National Soil Resources Institute (NSRI). Most soil profiles from the NSRI database extend as deep as 1.5 m for the basin (about 70 % of the profiles) and are provided with vertically variable Brooks and Corey soil moisture retention parameters. At the surface, the NSRI database differentiates between soil hydraulic parameters depending on land use (arable, permanent grassland, ley grassland, and other). Land use cover is provided from data collected by the

International Geosphere-Biosphere Programme (IGBP). The IGBP 2007 data are utilized to determine land cover types from 17 IGBP classes. These are re-classified to the 9 JULES land use types (Smith et al., 2006). The outcome is that cropland and mosaic/natural land use are the dominant land use types in the area (97 %).

Meteorological inputs to JULES were provided by the data from the Climate, Hydrology and Ecology research Support System (CHES) project. The data set, produced by the Centre of Ecology and Hydrology (CEH), UK (E. Blyth, personal communication, 2012), includes 1 km gridded daily rainfall amounts derived from the UK rain gauge network measurements for the period 1971–2008 (Keller et al., 2006). In addition, air temperature, vapour pressure, long and short wave downward radiation, and wind speed, derived by downscaling the observed meteorology from the Meteorological Office Rainfall and Evaporation Calculation System (MORECS) $40 \text{ km} \times 40 \text{ km}$ data set (Hough and Jones, 1997) accounting for the effects of topography, are also included. Daily observations of river flow at a number of gauging stations, along with groundwater levels at various observation frequencies (daily to monthly) from boreholes in the catchment, are used to evaluate model performance (Fig. 1). Groundwater levels at the same observational boreholes were previously examined by Jackson (2012), who used a conceptual model to estimate recharge rates to groundwater.

Chalk hydraulic properties are not available from standard national/global soil data sets (in the NSRI data set it is classified as a rock). Instead, these properties are estimated using soil moisture and matric potential observations at Warren Farm in the Kennet along with data from an on-site automatic weather station (Ireson et al., 2006) (Fig. 1). Soil moisture was measured between May 2003 and February 2006 us-

ing neutron probe measurements at different depths between 0.1 and 4.1 m taken fortnightly, on average. Either pressure transducer tensiometer (wet conditions) or equitensiometer (dry conditions) readings were taken for the same period of time to measure soil matric potential at 1 m depth every 15 min (Ireson et al., 2006; Ireson, 2008). Weather data include hourly observations of rainfall, downward short-wave solar and downward long-wave radiation, air temperature, specific humidity, and wind speed for the period between October 2002 and January 2009. The sub-daily weather data are used to account for any soil moisture sensitivity to the rainfall timing and intensity.

3 Method

The hydrological process representation in JULES is assessed with respect to the four primary functions of a hydrological system (Yilmaz et al., 2008): (1) overall water balance, (2) vertical redistribution, (3) temporal redistribution, and (4) horizontal spatial redistribution. Table 2 lists the assessment metrics for each of the four functions: the examined model assumptions/simplifications, the implemented model modifications, and the information sources used to inform the model modifications. Each of these information sources is described in the following sub-sections. The implemented model modifications considered below consist of a sub-daily weather generator, representation of sub-grid scale heterogeneity, dual Brooks and Corey curve representation of chalk hydraulic properties, change of the lower boundary condition, and coupling to a groundwater model.

3.1 Sub-daily weather generator

The daily CHES weather data are downscaled in time (15 min) by a weather generator (WG) (D. Clark, personal communication, 2013). The code provided by CEH uses a cosine variation for sub-daily temperature defined by the average daily temperature and temperature variation range (defined as 7 °C based on the local automatic weather station – AWS). Sub-daily incoming long-wave radiation is calculated using the same phase of the cosine function as that used for the temperature disaggregation. Sub-daily downward short-wave radiation is calculated as a product of the daily average downward shortwave radiation and a normalized fraction of a daily total solar radiation defined by a geographical location, time of year and day. Sub-daily specific humidity is assumed to be equal to the minimum of the saturated specific humidity (for a given sub-daily temperature) and the average daily specific humidity. Wind speed and air pressure are assumed to be constant throughout the day. Sub-daily precipitation is divided into large-scale rainfall, convective rainfall and large-scale snow. This differentiation is based on the mean daily temperature. Precipitation is defined as snow if the temperature is below 0 °C; convective if the temperature is above

20 °C; and large-scale rainfall, otherwise. It is set to start at a random time during a day and to continue for a specified number of hours over the entire corresponding model grid: 2 h for a convective storm, and 5 h for large-scale precipitation. The model configuration that includes the weather generator is referred to as JULES + WG (see Table 2).

3.2 Representation of sub-grid scale heterogeneity of near-surface soil hydraulic properties

A statistical approach is chosen to represent sub-grid scale heterogeneity of soil hydraulic properties; therefore, the upper soil layer storage capacity is assumed to be heterogeneous and to have a Pareto probability distribution with shape parameter b and upper soil layer depth dz (Bell et al., 2009). This representation is available in the standard version of JULES, but there is no guidance on selection of the two parameters. This approach limits the amount of water available for infiltration according to the soil moisture state, with the rest of the rainwater becoming surface runoff. The infiltrated water is then routed vertically through the soil using Richards' equation.

Since there is limited information to constrain both parameters, the effective upper layer soil depth dz is fixed to the JULES default value of 1 m. A regionalized base flow index (BFI) from the HOST soil classification (BFIHOST) (Boorman et al., 1995; Bulygina et al., 2009) is used to specify the Pareto distribution shape parameter b for each soil type in the catchment. The parameter is calibrated using water partitioning between surface runoff and drainage by JULES. The parameter value that results in the drainage-to-total-runoff ratio closest to the expected BFIHOST for that soil classification is chosen to be representative of the soil heterogeneity. Due to the high computational requirement of JULES, only 21 regularly spaced values between 0 and 2 are considered. The considered parameter b range is found to provide suitable drainage-to-total-runoff ratios for the catchment soils and meteorological conditions. The model configuration that includes both the weather generator and the PDM model is referred to as JULES + WG + PDM (Table 2).

3.3 Chalk hydraulic properties estimation

Modelling vertical soil water flow in JULES using Richards' equation requires the following descriptors: air entry pressure head, Brooks and Corey exponent (Brooks and Corey, 1964), saturated hydraulic conductivity, soil moisture at saturation, residual soil moisture, soil moisture at the critical point when transpiration starts to decrease, and soil moisture at wilting point. Due to the two distinct flow domains in chalk–matrix and fractures, two intersecting Brooks and Corey curves are employed when fitting a chalk soil moisture retention curve. The effective soil moisture at the curves' intersection is estimated using available observations. This leads to a double curve representation of hydraulic conduc-

Table 2. Metrics and information used for model diagnosis and modification.

Metrics	Examined model assumptions/simplifications	Implemented model modifications	Resulting model configuration	Information sources
Relative bias in total runoff: $RBias_Q = \frac{\sum_i Q_i^{mod} - \sum_i Q_i^{obs}}{\sum_i Q_i^{obs}}$ Relative bias in subsurface runoff ² : $RBias_{SR} = \frac{\sum_i SR_i^{mod} - \sum_i SR_i^{obs}}{\sum_i SR_i^{obs}}$	Weather inputs are constant when disaggregated to finer timescales; soil properties are constant over relatively large spatial scales (1 km)	Sub-daily weather generator (Sect. 3.1); PDM model and its parameterization (Sect. 3.2)	JULES + WG JULES + WG + PDM	Observed flows at the catchment outlet, precipitation from CHES data set, regionalized base flow index HOST (BFL-HOST)
II. Vertical redistribution				
Mean square relative error for soil moisture ³ : $MSRE = \frac{1}{N} \sum_{n=1}^N \frac{1}{T_n} \sum_{t=1}^{T_n} \frac{(\theta_{n,t}^{mod} - \theta_{n,t}^{obs})^2}{var(\theta_{n,t}^{obs})}$	A single permeability Richards' equation is capable of representing chalk soil hydraulic behaviour	Chalk representation via a double Brooks–Corey soil moisture retention curve and calibrated K_{sat} (Sect. 3.3)	JULES + WG + PDM + CHALK	Observations of soil matric potential and soil moisture
III. Temporal redistribution				
Nash–Sutcliffe efficiency for flow: $NS = 1 - \frac{\sum_i (Q_i^{mod} - Q_i^{obs})^2}{\sum_i (Q_i^{obs} - E Q_i^{obs})^2}$ and log-transformed flow: $NS = 1 - \frac{\sum_i (\log(Q_i^{mod}) - \log(Q_i^{obs}))^2}{\sum_i (\log(Q_i^{obs}) - E \log(Q_i^{obs}))^2}$	Horizontal unsaturated zone disconnection; no root uptake from deep saturated zone; unit gradient lower boundary condition; no surface/subsurface routing	Change of the lower boundary condition; approximation of water travel through the deep unsaturated zone (Sect. 3.4); and coupling to the groundwater model ZOOMQ3D (Sect. 3.5)	JULES + WG + PDM + CHALK + GW	States/fluxes from a detailed physics-based model of a chalk hillslope, observed flows at the catchment outlet
V. Spatial redistribution				
Nash–Sutcliffe efficiency for raw and log-transformed flow, relative bias in total runoff, visual inspection of groundwater levels at selected boreholes	Groundwater model parameterization	Change of specific yield parameters in the groundwater model	JULES + WG + PDM + CHALK + GWadj	Observed flows and groundwater level hydrographs at internal catchment points

¹ mod refers to modelled values; obs refers to the observed values; Q_i denotes runoff value at time i ; SR_i denotes subsurface runoff value at time i . ² Subsurface runoff is calculated via a hydrograph separation for observations. ³ $\theta_{n,t}$ denotes effective soil moisture at time t at the n th soil depth.

tivity dependence on soil moisture. The JULES soil module is modified accordingly to allow for a dual curve soil moisture retention representation. Although preferential *bypass* flow can occur in chalk (Ireson et al., 2012), it is considered to be relatively rare in the Kennet (Ireson and Butler, 2011). Consequently, it is not a major component in groundwater recharge and JULES has not been modified to include this effect.

The residual soil moisture content cannot be readily observed in the field, as chalk never dries out sufficiently to reach this state (Ireson, 2008). Therefore, the residual soil moisture content is estimated as a difference between maximum observed soil moisture and the effective porosity. The effective porosity (which includes matrix and fractures) is fixed at 0.36 (i.e. matrix porosity of 0.35 and fracture porosity of 0.01) (Bloomfield, 1997; Price et al., 1993). While fracture porosity tends to be higher at the soil surface due to the chalk weathering process (Ireson, 2008), this is not represented here due to the lack of comprehensive observations of soil moisture dynamics at multiple vertical levels; and the effects of the assumption are discussed in Sect. 4. Two sets of Brooks and Corey parameters are estimated by fitting the dual curve soil moisture retention representation to measurements of soil moisture and matric potential at 1 m depth obtained from field data collected at Warren Farm, Berkshire (Lowland Catchment Research – LOCAR – experiment data described in Ireson, 2008). Mean Square Error (MSE) is used to measure goodness of fit. Then, using the derived soil moisture retention curve, soil moisture at the critical point is calculated using the *wet* end curve at -40 kPa matric potential, while soil moisture at wilting point is calculated using the *dry* end curve at -1500 kPa.

Chalk saturated hydraulic conductivity is estimated by fitting the simulated soil moisture profiles to the available soil moisture neutron probes at multiple depths down to 4 m. For calibration purposes, 100 random values of saturated hydraulic conductivity are sampled logarithmically between 0.001 and 10 m day^{-1} . Mean square relative error (MSRelE; see definition in Table 2) for soil moisture between modelled and observed soil moisture for all observation depths is used as an objective function. The objective function increases error weights for the deeper layers that have less variable soil moisture, which is deemed to be important for drainage evaluation purposes. The model configuration that includes the weather generator, the PDM model and the chalk representation is referred to as JULES + WG + PDM + CHALK (Table 2).

3.4 A detailed physics-based model of a chalk hillslope

A physics-based model for 2-D flow in chalk (Ireson and Butler, 2013) represents a hillslope transect through unconfined chalk in the Pang catchment, located in close proximity to the river Kennet. While this model is an approximation itself and can only represent one set of hillslope properties,

it is built upon the best current knowledge of the hydrology of chalk hillslopes and is the best available test bed for simpler approximations. Flows in the 2-D model are governed by Richards' equation in both the saturated and unsaturated zones; and the properties of the chalk matrix and fractures are represented using an equivalent continuum approach (Peters and Klavetter, 1988; Doughty, 1999; Ireson et al., 2009). The Richards' equation is solved using a finite volume method.

Fluxes and states of the chalk hillslope model for the period 1970–2000 are examined to assess the following two assumptions underlying the JULES hydrology: (a) there is no hydrological interaction between neighbouring vertical soil columns, and (b) a unit gradient flow is a satisfactory approximation of the lower boundary condition at the 3 m base of the soil column on a hillslope location with a typically deep unsaturated zone. Further, the hillslope model is used to evaluate the nature of coupling between the unsaturated zone and groundwater, as well as the nature of water transport in the deep unsaturated zone located between the base of the JULES soil column and the water table. For these purposes, lateral fluxes in the unsaturated and saturated zones, hydraulic gradients and drainage rates at the soil column base, transpiration volumes extracted from the saturated and unsaturated zones by plants, and recharge rates at the groundwater table are extracted from the model. To reduce boundary condition effects at the upper and lower ends of the hillslope, the above variables are considered in the middle of the hillslope.

3.5 ZOOMQ3D distributed groundwater model

Groundwater flow in the Kennet is simulated using the ZOOMQ3D finite difference code (Jackson and Spink, 2004). The groundwater model is set up to simulate fluctuations in groundwater level, river baseflow, and spring discharge on a daily time step. The model uses gridded catchment representations at two scales; a 2 km base grid is locally refined to 500 m over the central part of the catchment. Rivers are simulated using an interconnected set of river reaches that exchange water with the aquifer according to a Darcian type flux equation. The vertical variations in rock hydraulic properties are represented using a three-layer model based on geological models of the hydrostratigraphy within the London Basin. The model is assessed to be a relatively good representation of the processes in the region in comparison with other chalk modelling examples (Jackson et al., 2011; Power and Soley, 2004).

ZOOMQ3D requires a significant number of parameters including horizontally and vertically distributed hydraulic conductivity and storage coefficient values. The parameters were zonally regularized and calibrated to approximate regional water table elevations (Jackson et al., 2011). For parameter estimation purposes, recharge has been modelled using a distributed recharge model ZOODRM (Mansour and Hughes, 2004) based on a conceptual Penman–Grindley soil

moisture deficit model (Penman, 1948; Gridley, 1967). As a result, it needs to be understood that the calibrated groundwater parameters are only representative for the ZOODRM recharge field and are not, therefore, adjusted for recharge fluxes obtained using JULES.

The model configuration JULES + WG + PDM + CHALK is coupled to ZOOMQ3D based on the findings from the detailed 2-D model (Sect. 3.4) and is referred to as JULES + WG + PDM + CHALK + GW (Table 2). When groundwater model parameters are adjusted to examine sensitivity of the model response, the configuration is referred to as JULES + WG + PDM + CHALK + GWadj.

3.6 Surface runoff routing

The standard JULES configuration (version 2.2) does not have a surface-water routing option. Therefore, given the catchment size, flows are averaged over 10-day intervals to reduce the impact of routing effects. For the chosen flow averaging interval, any inaccuracy in the estimated river discharge due to the lack of surface routing is believed to be minor when compared to the total flow magnitude (groundwater contributes 87 % of the flow, on average), and inaccuracies in both actual baseflow index estimation (when BFIHOST is used) and in groundwater routing representation. Further, swallow holes in the catchment (West and Dumbleton, 1972) and river-soil water exchange for surface runoff (i.e. infiltration of surface runoff into the river bed) are not represented in the model, and possible consequences of this are discussed later in the Results and discussion, Sect. 4.

3.7 Other JULES parameters

The remaining JULES parameters are assigned as follows. Land use fractions are taken from the IGBP 2007 data set, and re-classified into the nine land use types commonly used for JULES applications (Smith et al., 2006). Soil hydraulic parameters are taken from the NSRI soil database with the exception of soil layers that are classified as chalk. Soil hydraulic properties below the deepest NSRI horizon, typically at 1.5 m, are assigned the deepest horizon properties. Chalk hydraulic properties derived in this study are assigned to soil horizons that are classified as chalk in the NSRI database. The dominant agricultural crop for the area is spring barley (Limbrick et al., 2000). The root depth for the crop was chosen as 1 m (average value based on Breuer et al., 2003) and canopy height was chosen as 0.8 m (Hough and Jones, 1997; Mauser and Schädlich, 1998). Leaf area index (LAI) changes seasonally, with maximum of $\text{LAI} = 3$ (Mauser and Schädlich, 1998; Petr et al., 2002). The maximum interception capacity per unit leaf area is fixed at 0.2 mm, so that the upper limit to interception is $0.2 \times \text{LAI}$ (Hough and Jones, 1997). Other vegetation parameters are set at their recommended default value for JULES (Cox et al., 1999).

4 Results and discussion

Observations of water fluxes, soil moisture and groundwater levels in the Kennet catchment are compared with the simulated values derived using the sequentially modified JULES model structure to represent the four hydrological functions of a catchment.

4.1 Water balance

The long-term water balance is calculated for the period 1972–2007 from observations and various model configurations, and three metrics are calculated – relative bias for total runoff (RBias_Q) and surface runoff (RBias_{SR}), and MSRelE (Table 3). The unmodified JULES (version 2.2) is found to overestimate the total runoff by 24 % and, correspondingly, underestimate the evapotranspiration (ET) by 15 %. This is attributed to the constant temporal disaggregation of weather variables that is hard-coded into the model. When weather variables are temporally disaggregated using the WG, described in Sect. 3.1, the total runoff is only 2 % lower than the observed value. However, neither configuration is capable of producing any surface runoff. This is because the hydraulic conductivities of the catchment soils (derived from the NSRI parameter database), even for relatively clayey soils, are sufficiently high to enable virtually all the instantaneous rainfall rates obtained using temporal disaggregation to infiltrate into the soil.

The parameter b of the PDM model is selected based on regionalized information from BFIHOST (Sect. 3.2) and ranges, mainly, between 0 and 0.4, except for two grids where b is set as 0.7 and 1. Further, the parameter b is assigned 0 value over approximately 60 % of the catchment for the locations with permeable (chalky) top-soils. Addition of the PDM model (JULES + WG + PDM configuration) with parameters selected based on regionalized information from BFIHOST (Sect. 3.2) generated, on average, 70 mm yr^{-1} of surface runoff (compared to 39 mm yr^{-1} derived by baseflow separation at the catchment outlet). This is likely to originate from the regionalization error – the catchment average regionalized BFIHOST value (0.78) is lower than the BFI value calculated from observed flow at the catchment outlet (0.87). This difference may arise from a number of locally relevant soil properties and processes that are not represented in the regionalized BFIHOST, for example there is a focussed recharge into sink or swallow holes of runoff from the Palaeogene deposits in the lower reaches of the Kennet catchment (West and Dumbleton, 1972). Such localized processes could, in principle, be explicitly represented in the land surface model, but this would be difficult in practice due to the scales involved; for example, representing the sink holes would require fine-scale data (at 0.1 to 1 m resolution) describing the land surface features.

It is to be noted that the proposed model modification with PDM and its parameterization is not the only possible model

Table 3. Observed and simulated water balance and metrics of model performance.

Source/model configuration	Rainfall, mm yr ⁻¹	ET, mm yr ⁻¹	Total runoff, mm yr ⁻¹		RBias _Q	RBias _{SR}	MSRelE
			Surface runoff, mm yr ⁻¹	Subsurface runoff, mm yr ⁻¹			
Obs ¹	784	485	299		–	–	–
			39	260			
Standard JULES (version 2.2) ²	784	410	370		0.24	0.42	– ³
			0	370			
JULES + WG	784	489	292		–0.02	0.12	–
			1	291			
JULES + WG + PDM	784	489	299		0.00	–0.12	3.64
			70	229			
JULES + WG + PDM + CHALK	784	495	293		–0.02	–0.13	1.12
			67	226			
JULES + WG + PDM + CHALK + GW	784	496	293		–0.02	–0.13	1.07
			67	226			
JULES + WG + PDM + CHALK + GWadj	784	496	293		–0.02	–0.13	1.07
			67	226			

¹ For observations, ET is calculated as a residual between the long-term precipitation and runoff; surface and subsurface runoff are calculated based on the hydrograph separation. ² For model configurations, surface and subsurface runoff are taken as surface runoff and drainage fluxes, respectively, produced by a model. ³ MSRelE is calculated starting from the JULES + WG + PDM configuration.

modification. An alternative, which potentially leads to increased surface runoff production, includes spatial and, perhaps, further temporal downscaling of rainfall to produce more intense events over parts of the 1 km discretization grids. Table 3 shows that the model modifications used to improve the representation of the additional processes observed in the catchment (and outlined in Table 2) do not compromise the simulated water balance.

4.2 Vertical redistribution through the soil

Both JULES + WG + PDM and JULES + WG + PDM + CHALK configurations use 4.5 m long soil columns with 0.1 m thick soil layers to facilitate the comparison with the observed soil moisture. The JULES + WG + PDM configuration results in overly dry soils between 1 and 4.1 m depth when compared to the observations (Fig. 2; a representative subset of the soil moisture time series is shown); and the corresponding MSRelE metric equals 3.64. This soil dryness is attributed to incorrect representation of chalk soil hydraulic properties. Figure 3 shows two Brooks and Corey soil moisture retention curves fitted to the pairs of soil moisture and matric potential observations at 1 m depth in chalk; the curves intersect at an effective soil moisture of 0.31 (effective soil moisture equals soil moisture with subtracted residual soil moisture). The figure illustrates a threshold change

in the chalk soil moisture retention curve and consequently, through the Brooks–Corey–Mualem model, the unsaturated hydraulic conductivity relationship. This change in properties is related to the dual porosity–dual permeability nature of the chalk soil (Ireson et al., 2009). Estimated chalk hydraulic properties are given in Table 4. Further, the time varying vertical distribution of soil moisture estimated by the JULES + WG + PDM + CHALK configuration is shown in Fig. 2; this corresponds to an MSRelE metric of 1.12. This value stays approximately the same throughout further model modifications to include additional functions of the hydrological system. The inclusion of chalk properties into the model produces a better simulation of soil moisture content at the Warren Farm site than that from the JULES + WG + PDM configuration. This corresponds well with the observed soil moisture below ~1 m depth. However, the upper soil tends to be wetter than the observed moisture levels. This is attributed to the chalk’s vertical heterogeneity; fractures appear more frequently and are larger in the upper chalk. Depth-variable soil hydraulic properties are required to capture the phenomenon. This is not attempted here due to the lack of soil moisture–matric potential observational pairs at multiple vertical levels to define entire soil moisture retention curves.

Table 4. Hydraulic double Brooks and Corey curve parameters for chalk.

Parameter	Description	Wet end	Dry end	Source
b	Exponent	30.2	1.3	Calibration to soil moisture and matric potential at 1 m
α , m	Soil matric potential at saturation	0.15	12.2	Calibration to soil moisture and matric potential at 1 m
K_{sat}^* , m day ⁻¹	Saturated hydraulic conductivity	0.016 (0.02)		JULES calibration to soil moisture at multiple depths down to 4.1 m
θ_s^{eff}	Effective saturated soil moisture		0.36	Price et al. (1993), Bloomfield (1997)
θ_r	Residual soil moisture		0.11	Soil moisture observations and θ_s^{eff} value
$\theta_{\text{cr}}^{\text{eff}}$	Effective saturated soil moisture at critical point		0.32	Brooks and Corey equation at -40 kPa
$\theta_{\text{wilt}}^{\text{eff}}$	Effective saturated soil moisture at wilting point		0.05	Brooks and Corey equation at -1500 kPa
$\theta_{\text{inter}}^{\text{eff}}$	Effective soil moisture at the two curves intersection		0.31	Calibration to soil moisture and matric potential at 1 m

* K_{sat} is fitted using JULES + WG + PDM + CHALK as well as JULES + WG + PDM + CHALK + GW configurations. The value for the latter is shown in the parenthesis.

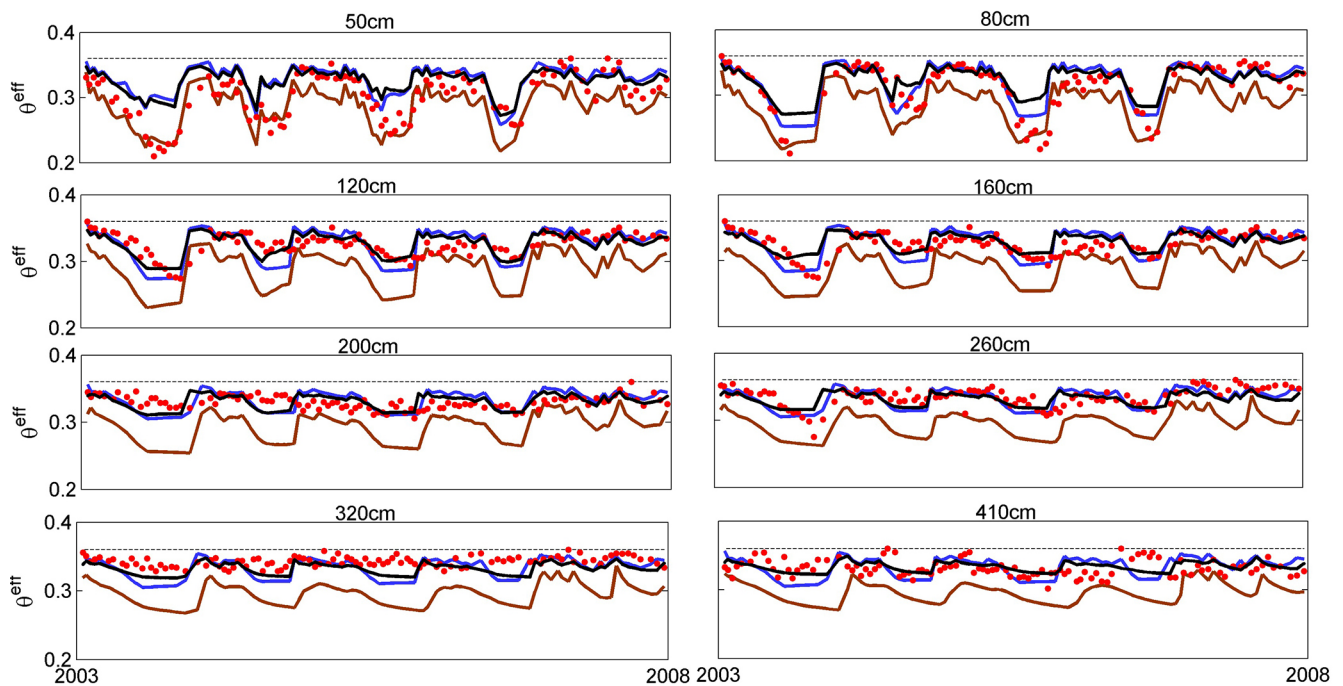


Figure 2. Comparison of the optimized effective soil moisture (θ^{eff}) time series with the observed soil moisture (red dots) at various depths at Warren Farm, Berkshire, UK. Brown lines show soil moisture estimated by JULES + WG + PDM; blue lines show soil moisture estimated by JULES + WG + PDM + CHALK; and black lines show soil moisture estimated by JULES + WG + PDM + CHALK + GW. Grey horizontal line shows the effective soil moisture at saturation 0.36. Note, only a representative subset of soil moisture time series utilized in the analysis is shown in the figure.

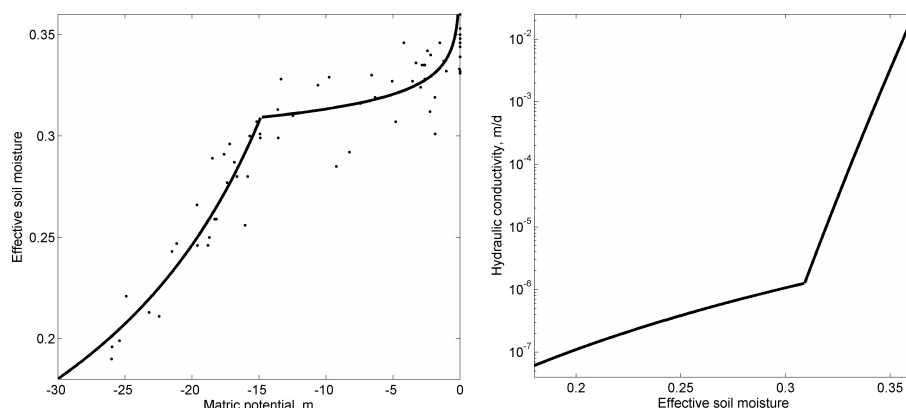


Figure 3. Chalk soil moisture retention fit using a dual Brooks and Corey curve and the corresponding hydraulic conductivity dependence on effective soil moisture. Black dots are observed data.

4.3 Temporal redistribution

The original as well as the modified model configurations are only capable of partitioning water fluxes at the point/grid scale, and do not have a mechanism for further routing to provide temporal water redistribution at the catchment outlet. Here, some assumptions about the nature of such water redistribution (Sect. III in Table 2) are assessed, and lateral routing through the saturated zone is achieved through coupling the model to a distributed groundwater model.

Fluxes extracted from the physics-based 2-D model of a chalk hillslope imply that there are two simplifications that can be made with regards to the 2-D nature of hillslope hydrological processes. First, lateral fluxes in the chalk unsaturated zone are found to be insignificant when compared to the net (vertical and lateral) fluxes in the unsaturated zone. Hence, the simplifying assumption about inter-soil-column-independent hydrological behaviour is a reasonable and sufficiently accurate approximation for the area. Second, evapotranspiration losses from the chalk saturated zone are found to be negligible compared to those from the unsaturated zone. It is therefore assumed that evapotranspiration processes can be restricted to the unsaturated zone when coupling the unsaturated zone to groundwater for the study area investigated herein.

Extracted vertical hydraulic gradients from the 2-D hillslope model are compared to the unit gradient lower boundary condition along with a number of alternative lower boundary conditions (using mean absolute difference as an objective function). Of these, it is found that a *persistent gradient* condition is the most consistent and accurate approximation of the lower boundary condition for the area. The persistent gradient condition assumes that hydraulic gradient is time varying but almost constant with depth at the soil column base. The condition can be approximated using the hydraulic gradient between soil column nodes just above the column base, requiring a relatively fine node mesh at the column base. The persistent gradient condition can be seen as a general conse-

quence of the following lower boundary condition $\frac{\partial^2 h}{\partial z^2} = 0$, where the unit gradient lower boundary condition $\frac{\partial h}{\partial z} = 1$ is a special case. Note that only the hydraulic gradient at the soil column base is approximated herein. This gradient is used to substitute the unit gradient in the formula for the drainage flux in JULES. This implies that hydraulic conductivity at the base of the soil column is based on the nearest to the bottom node state.

Further, the persistent gradient approximation is evaluated for multiple soil column depths to optimize its applicability. The mean absolute difference between the persistent hydraulic gradient at the lower boundary and hydraulic gradients extracted from the hillslope model at a number of depths is used as an objective function. It is found that the objective function improves with increasing depth of soil column but less significantly after 6 m. As a trade-off between the soil column depth and the lower boundary approximation accuracy, an optimal depth to apply the persistent gradient lower boundary condition is chosen to be 6 m. Figure 4 compares hydraulic gradients at a 6 m column base extracted from the 2-D model to the unit gradient as well as to the gradient just above 6 m (approximately at 5.5 m depth, based on the model mesh), representing the persistent boundary condition approximation.

Lastly, to draw a connection between the modelled potential recharge at 6 m depth and the modelled actual recharge at the water table, temporally averaged vertical fluxes extracted from the 2-D hillslope model are considered for 6 hourly (the model step), daily, weekly and 30-day periods. The correlations between the time series of actual and potential recharges for the averaging periods are 0.75, 0.8, 0.89, and 0.94, respectively. Total actual and total potential recharges for the 1970–2000 period are found to be less than 1 % different. Average daily (the regional groundwater time step) potential simulated recharge at 6 m and actual simulated recharge at the water table extracted from the 2-D model are shown in Fig. 5. It can be seen that the potential recharge

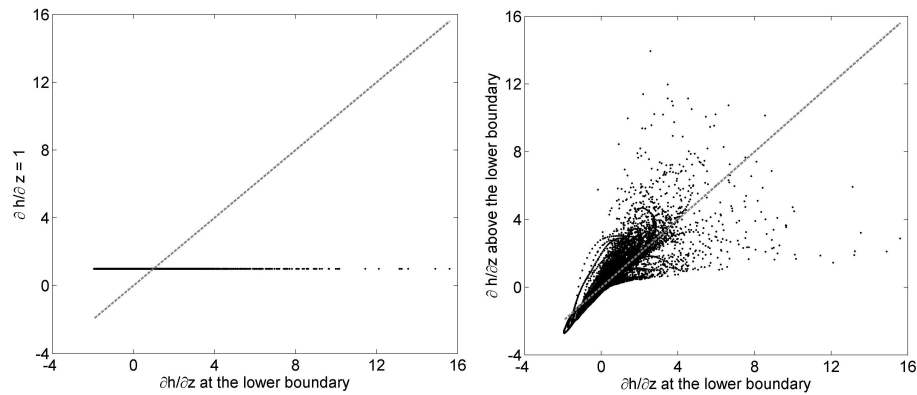


Figure 4. Comparison of hydraulic head gradient $\partial h/\partial z$ at a 6 m depth extracted from the 2-D model with a unit gradient condition (left panel), and with a persistent gradient condition extracted at a 5.5 m depth (right panel).

is widely spread at low actual recharge rates (below about 2 mm day^{-1}). However, the potential recharge becomes quite a consistent predictor for the actual recharge at mid- to high actual recharge rates.

Based on the above findings, JULES + WG + PDM + CHALK is used with a 6 m deep soil column and 0.1 m thick soil layers, and is coupled via a weak two-way coupling to the groundwater model ZOOMQ3D implemented through the lower boundary condition (persistent gradient). The *weak* coupling assumes that the drainage flux from JULES is used as an upper boundary condition by ZOOMQ3D, and any upward water fluxes from the saturated zone to the upper unsaturated zone are calculated based on the (persistent gradient) lower boundary condition. Note, the saturated hydraulic conductivity for chalk soil is re-calibrated following the procedure given in Sect. 3.3 (Table 4) as the new persistent gradient lower boundary condition impacts the soil moisture dynamics.

The resulting 10-day-averaged river flow at the catchment outlet (Theale) for the period 1994–2006 is shown on Fig. 6. The period includes two droughts in the region (1995–1998 and 2003–2006) as well as substantially wet 1999–2001 period that led to groundwater flooding. Figure 6 also shows model performance measures for the total simulation period of 1972–2007, with a Nash–Sutcliffe efficiency for simulated flow ($NS = 0.82$) and log-transformed simulated flow ($NS_{\log} = 0.81$), as well as a relative bias for the total flow ($RBias_Q = 0.01$). Note, the relative bias is calculated using the flow at the catchment outlet, not the sum of surface and drainage fluxes produced by the land surface model component of the configuration. This explains the slight difference between the $RBias_Q$ values in Table 3 and Fig. 6 for the model configuration.

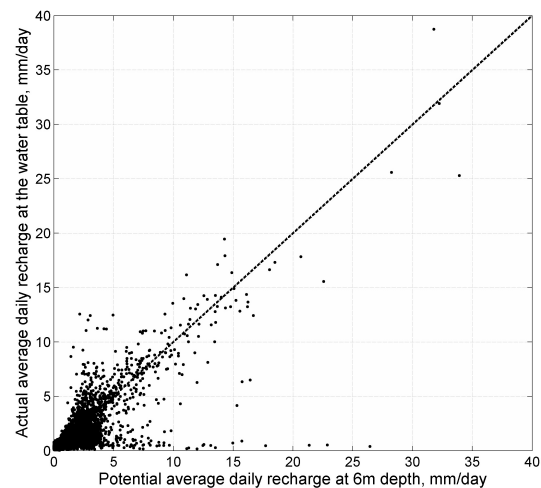


Figure 5. Correspondence between potential and actual daily recharge rates extracted from the 2-D model.

4.4 Spatial redistribution over the groundwater and surface-water systems

Due to the distributed nature of the coupled model configuration, flows (Fig. 6) and groundwater levels (Fig. 7) can be examined at the internal catchment points shown in Fig. 1. It can be seen that total flow tends to be underestimated in the smaller sub-catchments such as the Kennet at Marlborough and the Lambourn at Shaw. Inspection of water movement patterns inside the groundwater model ZOOMQ3D offers a possible explanation. The Lambourn groundwater catchment area is found to be underestimated by ZOOMQ3D when compared to the groundwater catchment area extracted from local observational boreholes and spring head data (Parker, 2011; Parker et al., 2015). Further, the model tends to direct some water from the Lambourn to the middle part of the Kennet (Parker, 2011, S. Parker, personal communication, 2013), which helps to explain the total flow overestimation

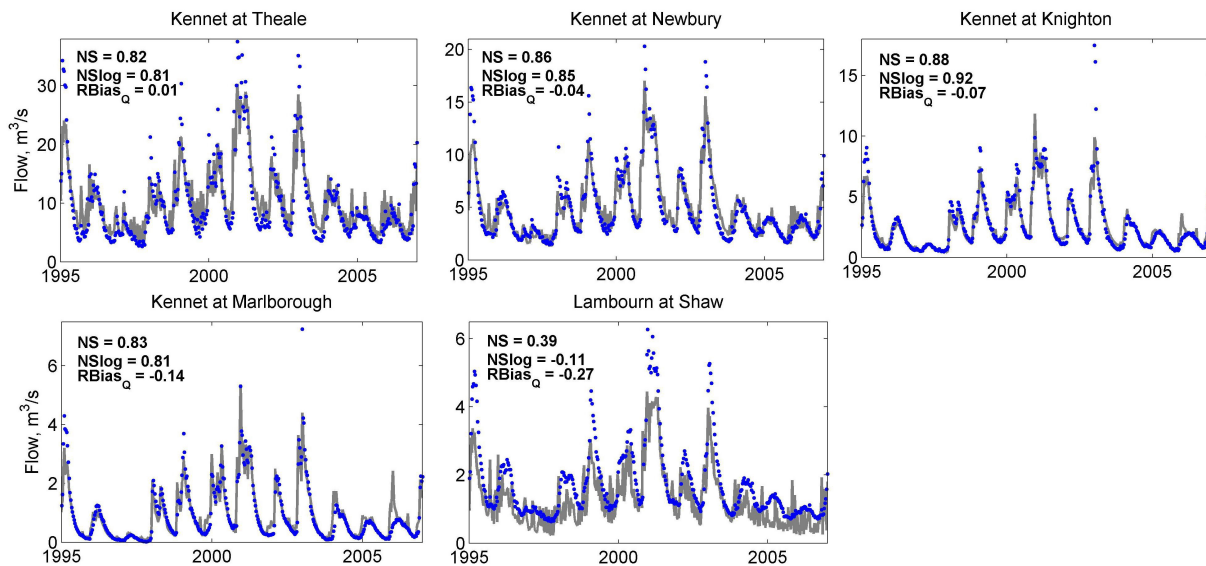


Figure 6. 10-day average flows at five gauging stations in the Kennet generated by the JULES + WG + PDM + CHALK + GW model configuration. Grey lines denote simulated flows, and blue dots are observations.

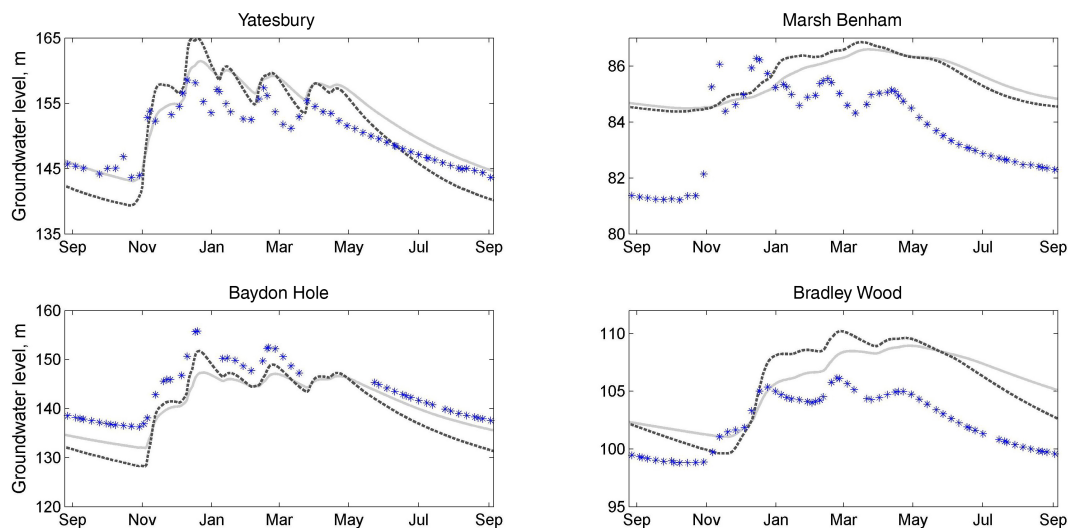


Figure 7. Water levels at four observational boreholes in the Kennet. Blue stars are observed levels, grey lines represent groundwater levels generated by JULES + WG + PDM + CHALK + GW configuration; and black dotted lines represent groundwater levels generated by JULES + WG + PDM + CHALK + GWadj configuration.

at the Marlborough, Newbury, and Knighton gauges. Further, during wet years, peak flows appear to be underestimated at all gauging stations; meanwhile, low flows are slightly overestimated for the Kennet at Theale and the Kennet at Newbury, and underestimated for the Lambourn at Shaw. Treating potential recharge from the land surface model as actual recharge to ZOOMQ3D might partly explain the low flow overestimation.

Because of the mismatch of scales between an observation borehole (order of 1 m) and JULES and ZOOMQ3D grid scales (1 km and 500 m, respectively), only a visual assess-

ment of the predicted water levels is attempted. Figure 7 illustrates simulated water levels at four selected boreholes for September 2000–August 2001 representing an unusually wet year leading to a groundwater flooding in the area. Similar to the results from Jackson (2012), who considered the same period and boreholes, water levels are mainly overestimated at the Marsh Benham and Bradley Wood boreholes. Moreover, the modelled response at Marsh Benham and Bradley Wood is more attenuated than the observed response. At the model scale (1 km), the estimated groundwater levels are indicative of the boreholes partly due to soil heterogeneity. For exam-

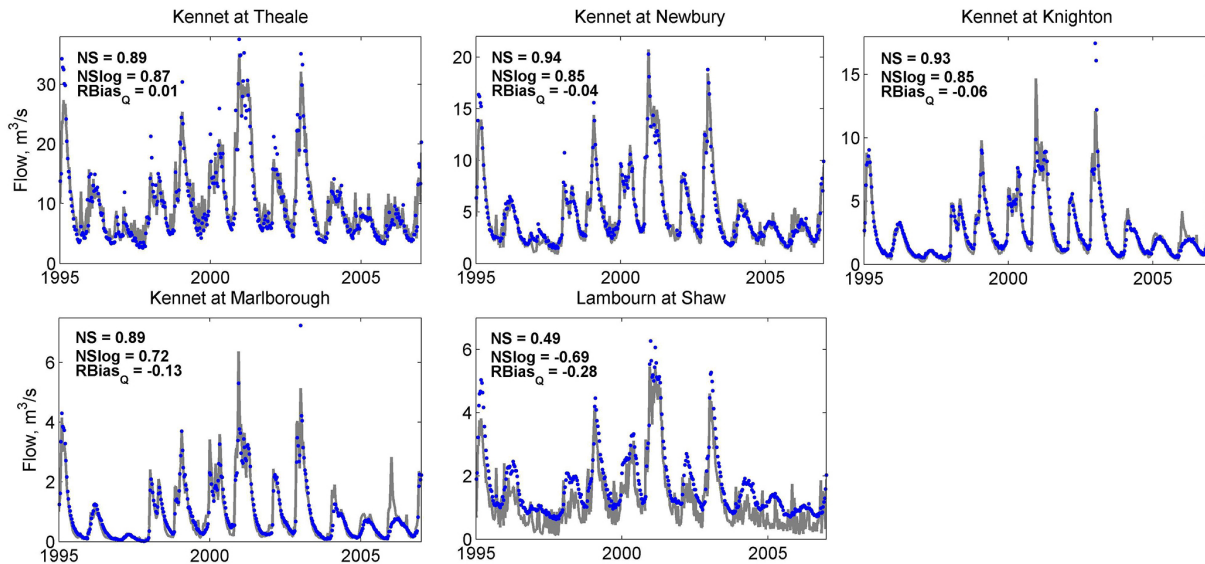


Figure 8. 10-day average flows at five gauging stations in the Kennet generated by the JULES + WG + PDM + CHALK + GWadj model configuration. Grey lines denote simulated flows, and blue dots are observations.

ple, the PDM model parameter b (as well as soil hydraulic properties) is chosen based on a dominant soil type; therefore, the recharge to total runoff ratios are 0.52 and 0.17 for Bradley Wood and Marsh Benham, correspondingly. However, other soils present in the model grids have very different recharge to total runoff ratios, for example, 0.98 and 0.88 for Bradley Wood and Marsh Benham, respectively. Incorporating the hydrology of these soils can potentially lead to more responsive water level behaviour at the boreholes as more water will infiltrate.

As indicated in Sect. 3.5, the parameters of the ZOOMQ3D groundwater model are derived using recharge from a different near-surface model, and thus are likely to be sub-optimal when recharge produced by JULES is used. A manual sensitivity analysis of model parameters showed that tuning values of specific yield or hydraulic conductivity leads to better agreement with the observed data. For example, Fig. 8 shows flows generated by the coupled model when ZOOMQ3D specific yield parameters are halved over the whole Kennet area. This results in a better representation of high flows, but mixed outcomes for low flows (according to the NS and NS_{log} performance measures). Groundwater levels at the selected boreholes become slightly more responsive, but do not change significantly (Fig. 7). As the primary research objective is to diagnose the hydrological limitations of a land surface model, a formal recalibration of an auxiliary groundwater model is not pursued here.

5 Conclusions

The paper is motivated by the goals of using land surface models as a basis for global hyper-resolution modelling of

the terrestrial water, energy, and biogeochemical cycles, including application to a range of complex hydrological prediction problems. This comes alongside the recognition that there are significant limitations in the accuracy with which hydrological fluxes and storages are represented in general in LSMs due to their focus on supporting large-scale climate modelling problems (Oleson et al., 2008; Wood et al., 2011). The paper uses a case study of the JULES LSM model applied to the Kennet catchment in southern England, which represents the challenging problem of hydrological modelling in a chalk dominated catchment with a predominantly deep unsaturated zone. A diagnostic approach is taken to identify the model inadequacies with respect to the four functions of a hydrological system: overall water balance, vertical redistribution of water through the soil, temporal redistribution of water, and spatial redistribution over the catchment's surface-water and groundwater systems. The approach facilitates a sequential model improvement using hydrological expert knowledge about model assumptions and simplifications relevant for each hydrological aspect considered. The following model modifications are presented and assessed in the paper:

- overall water balance: introduction of a weather generator and statistical description of top-soil heterogeneity via regionalized information;
- vertical redistribution through the soil: approximation of the dual permeability – dual porosity hydraulic soil behaviour;
- temporal redistribution: change of the lower boundary condition and approximation of coupling to a groundwater model;

- spatial redistribution over the catchment: alteration of groundwater model parameters.

It is noted that improving the model physics in sequence preserves model performance quality with respect to the other previously considered functions. For example, improving vertical distribution does not corrupt the water balance achieved at a previous model modifications stage. This might be explained by the physical basis of both the model and reasoning for model modifications. The improvements are illustrative of the potential outcomes of a diagnosis approach, and alternative or additional improvements are possible. These include: the representation of the temporal and spatial distribution of precipitation; inclusions of point/small-scale features such as sink holes; and more physics-based inclusion of the vertical and horizontal distribution of soil hydraulic properties. As a procedural improvement, uncertainty analysis could be used to indicate if output errors can be explained by estimates of particular input uncertainties.

For some applications, the intermediate model configurations might be sufficient. For example, while JULES + WG + PDM configuration cannot provide flow/groundwater level hydrographs as it lacks surface and subsurface water routing, the configuration can still be used to represent the water balance over an area. This is useful for regions where no groundwater model and/or detailed geology data are available. It is to be noted that the findings are catchment specific and result from a weak surface-groundwater coupling, and as such cannot be readily generalized to other environments with shallower water tables.

Diverse sources of information were used to guide the model assessment and include remotely sensed data (topography, land use), spatially extrapolated point data (soils, weather conditions), point measurements (soil moisture and matric potential, flow, groundwater level), regionalized hydrological information (BFIHOST), and states/fluxes extracted from an auxiliary physics based hillslope model (Ireson and Butler, 2013). Fewer data might result in a less detailed representation of the water cycle depending on the specifics the hydrological system being investigated.

Whilst this application of JULES to the Kennet catchment is highly specific, it conveniently illustrates the type of challenges and parameterization of complex and distributed hydrological processes, model coupling using simplified boundary conditions, and assimilation of different sources of information to model identification, which will be encountered in almost any attempt to improve the utility of LSMs for catchment-scale water cycle modelling, arising due to the *uniqueness of place* problem. The paper has demonstrated the considerable accuracy gains that can be achieved using a sequential model error diagnosis strategy and expert-lead model adjustments. These can be taken forward to develop a general comprehensive guidance for transitioning to high resolution land surface modelling.

Acknowledgements. The study was made possible through code and data provided by Andrew Ireson (2-D hillslope model and soil moisture data, University of Saskatchewan, Canada), Doug Clark (weather generator, CEH, UK), Ned Hewitt (AWS and neutron probes, CEH, UK), and Eleanor Blyth (CHESS, CEH, UK). Further, the authors thank Zed Zulkafi (Imperial College London, UK) for help with the initial JULES set-up and Simon Parker (Imperial College London, UK) for specific information on chalk groundwater catchment behaviour. The author Jackson publishes with the permission of the Executive Director of the British Geological Survey. The work was funded under the UK Natural Environment Research Council's Changing Water Cycle program, grant NE/I022558/1.

Edited by: S. Archfield

References

- Balsamo, G., Viterbo, P., Beljaars, A., van den Hurk, B., Hirschi, M., Betts, A., and Scipal, K.: A revised hydrology for the ECMWF model: verification from field site to terrestrial water storage and impact in the integrated forecast System, *J. Hydrometeorol.*, 10, 623–643, 2009.
- Bell, V., Kay, A., Jones, R., Moore, R., and Reynard, N.: Use of soil data in a grid-based hydrological model to estimate spatial variation in changing flood risk across the UK, *J. Hydrol.*, 377, 335–350, 2009.
- Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes, *Geosci. Model Dev.*, 4, 677–699, doi:10.5194/gmd-4-677-2011, 2011.
- Beven, K. and Germann, P.: Macropores and water flow in soils revisited, *Water Resour. Res.*, 49, 3071–3092, doi:10.1002/wrcr.20156, 2013.
- Beven, K. and Kirkby, M.: A physically based, variably contributing model of basin hydrology, *Hydrol. Sci. Bull.*, 24, 2415–2433, 1979.
- Black, P. E.: Watershed functions, *J. Am. Water Resour. As.*, 33, 1–11, 1997.
- Bloomfield, J.: The role of diagenesis in the hydrogeological stratification of carbonate aquifers: an example from the chalk at Fair Cross, Berkshire, UK, *Hydrol. Earth Syst. Sci.*, 1, 19–33, doi:10.5194/hess-1-19-1997, 1997.
- Blyth, E., Clark, D. B., Ellis, R., Huntingford, C., Los, S., Pryor, M., Best, M., and Sitch, S.: A comprehensive set of benchmark tests for a land surface model of simultaneous fluxes of water and carbon at both the global and seasonal scale, *Geosci. Model Dev.*, 4, 255–269, doi:10.5194/gmd-4-255-2011, 2011.
- Boone, A., Habets, F., Noilhan, J., Clark, D., Dirmeyer, P., Fox, S., Gusev, Y., Haddeland, I., Koster, R., Lohmann, D., Mahanama, S., Mitchell, K., Nasonova, O., Niu, G., Pitman, A., Polcher, J., Shmakin, A., Tanaka, K., van den Hurk, B., Verant, S., Verseghy, D., Viterbo, P., and Yang, Z.: The Rhone-aggregation land surface scheme intercomparison project: an overview, *J. Climate*, 17, 187–208, 2004.

- Boorman, D., Hollis, J., and Lilly, A.: Hydrology of Soil Types: A Hydrologically-based Classification of the Soils of the UK, Institute of Hydrology, Wallingford, 1995.
- Breuer, L., Eckhardt, K., and Frede, H.-G.: Plant parameter values for models in temperate climates, *Ecol. Model.*, 169, 237–293, 2003.
- Brooks, R. and Corey, A.: Hydraulic Properties of Porous Media, Hydrology Papers, Colorado State University, Fort Collins, Colorado, 1964.
- Brouyere, S., Dassargues, A., Hallet, V.: Migration of contaminants through the unsaturated zone overlying the Hesbaye chalky aquifer in Belgium: a field investigation, *J. Contam. Hydrol.* 72, 135–164, 2004.
- Bulygina, N., McIntyre, N., and Wheeler, H.: Conditioning rainfall-runoff model parameters for ungauged catchments and land management impacts analysis, *Hydrol. Earth Syst. Sci.*, 13, 893–904, doi:10.5194/hess-13-893-2009, 2009.
- Clark, D. B. and Gedney, N.: Representing the effects of subgrid variability of soil moisture on runoff generation in a land surface model, *J. Geophys. Res.*, 113, D10111, doi:10.1029/2007JD008940, 2008.
- Clark, D. B., Mercado, L. M., Sitch, S., Jones, C. D., Gedney, N., Best, M. J., Pryor, M., Rooney, G. G., Essery, R. L. H., Blyth, E., Boucher, O., Harding, R. J., Huntingford, C., and Cox, P. M.: The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon fluxes and vegetation dynamics, *Geosci. Model Dev.*, 4, 701–722, doi:10.5194/gmd-4-701-2011, 2011.
- Cox, P., Bett, R., Bunton, C., Essery, R., Rowntree, P., and Smith, J.: The impact of new land surface physics on the GCM simulation of climate and climate sensitivity, *Clim. Dynam.*, 15, 183–203, 1999.
- Dahan, O., Nativ, R., Adar, E., Berkowitz, B.: A measurement system to determine water flux and solute transport through fractures in the unsaturated zone, *Groundwater*, 36, 444–449, 1998.
- Dahan, O., Nativ, R., Adar, E., Berkowitz, B., Ronen, Z.: Field observation of flow in a fracture intersecting unsaturated chalk, *Water Resour. Res.*, 35, 3315–3326, 1999.
- Doughty, C.: Investigation of conceptual and numerical approaches for evaluating moisture, gas, chemical and heat transport in fractured unsaturated rock, *J. Contam. Hydrol.* 38, 69–106, 1999.
- Downing, R. A., Price, M., and Jones, G.: The Hydrogeology of the Chalk of Northwest Europe, Oxford University Press, Oxford, 1993.
- Grindley, J.: The estimation of soil moisture deficits, *Meteorol. Mag.*, 96, 97–108, 1967.
- Gusev, Y. and Nasonova, O.: The simulation of heat and water exchange in the boreal spruce forest by the land-surface model SWAP, *J. Hydrol.*, 280, 162–191, 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.
- Ireson, A.: Quantifying the hydrological processes governing flow in the unsaturated Chalk, PhD thesis, Imperial College London, 270 pp., available at: <http://www3.imperial.ac.uk/pls/portallive/docs/1/38553696.PDF> (last access: 1 March 2015), 2008.
- Ireson, A. M. and Butler, A.: Controls on preferential recharge to Chalk aquifers, *J. Hydrol.*, 398, 109–123, 2011.
- Ireson, A. M. and Butler, A. P.: A critical assessment of simple recharge models: application to the UK Chalk, *Hydrol. Earth Syst. Sci.*, 17, 2083–2096, doi:10.5194/hess-17-2083-2013, 2013.
- Ireson, A. M., Wheeler, H., Butler, A., Mathias, S., and Finch, J.: Hydrological processes in the chalk unsaturated zone – insights from an intensive field monitoring programme, *J. Hydrol.*, 330, 29–43, 2006.
- Ireson, A. M., Mathias, S., Wheeler, H., Butler, A., and Finch, J.: A model for flow in the Chalk unsaturated zone incorporating progressive weathering, *J. Hydrol.*, 365, 244–260, 2009.
- Ireson, A. M., Butler, A., and Wheeler, H.: Evidence for the onset and persistence with depth of preferential flow in unsaturated fractured porous media, *Hydrol. Res.*, 43, 707–719, 2012.
- Jackson, B., Browne, C., Butler, A., Peach, D., Wade, A., and Wheeler, H.: Nitrate transport in Chalk catchments: monitoring, modeling and policy implications, *Environ. Sci. Policy*, 11, 125–135, 2008.
- Jackson, C.: Simple automatic time-stepping for improved simulation of groundwater hydrographs, *Ground Water*, 50, 736–745, 2012.
- Jackson, C. and Spink, A.: User's Manual for the Groundwater Flow Model ZOOMQ3D, IR/04/140, British Geological Survey, Nottingham, UK, 107 pp., 2004.
- Jackson, C., Meister, R., and Prudhomme, C.: Modeling the effects of climate change and its uncertainty on UK chalk groundwater resources from an ensemble of global climate model projections, *J. Hydrol.*, 399, 12–28, 2011.
- Keller, V., Young, A., Morris, D., and Davies, H.: Task 1.1: Estimation of Precipitation Inputs, Environment Agency R&D Project w6-101, Centre for Ecology and Hydrology, Wallingford, UK, 1–35, 2006.
- Kloppmann, W., Dever, L., and Edmunds, W.: Residence time of Chalk groundwaters in the Paris Basin and the North German Basin: a geochemical approach, *Appl. Geochem.*, 13, 593–606, 1998.
- Kollet, S. and Maxwell, R.: Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model, *Water Resour. Res.*, 44, W02402, doi:10.1029/2007WR006004, 2008.
- Koster, R., Suarez, M., Ducharme, A., Stieglitz, M., and Kumar, P.: A catchment-based approach to modeling land surface processes in a general circulation model 1. Model structure, *J. Geophys. Res.*, 105, 24809–24822, 2000.
- Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of coupled atmosphere-biosphere system, *Global Biogeochem. Cy.*, 19, GB1015, doi:10.1029/2003GB002199, 2005.
- Lawrence, P. and Chase, T.: Climate impacts of making evapotranspiration in the Community Land Model (CLM3) consistent with the Simple Biosphere Model (SiB), *J. Hydrometeorol.*, 10, 374–394, 2008.
- Li, B., Peters-Lidard, C., Kumar, S., Rheingrover, S., and Anantharaj, V.: Free drainage or not: an evaluation of simulated soil moisture profiles by the Noah land surface model in the Mississippi region, 22nd Conference on Hydrology, 20–24 January 2008, New Orleans, USA, paper 134829, available at: <https://ams.confex.com/ams/88Annual/>

- techprogram/paper_134829.htm (last access: 4 December 2015), 2008.
- Limbrick, K., Whitehead, P., Butterfield, D., and Reynard, N.: Assessing the potential impacts of various climate change scenarios on the hydrological regime of the River Kennet at Theale, Berkshire, south-central England, UK: an application and evaluation of the new semi-distributed model, *INCA, Sci. Total Environ.*, 251/252, 539–555, 2000.
- Lohmann, D., Mitchell, K., Houser, P., Wood, E., Schaake, J., Robock, A., Osgrove, B., Sheffield, J., Duan, Q., Luo, L., Higgins, W., Pinker, R., and Tarpley, J.: Streamflow and water balance intercomparisons of four land surface models in the North American Land Data Assimilation System project, *J. Geophys. Res.*, 109, D07S91, doi:10.1029/2003JD003823, 2004.
- Mansour, M. and Hughes, A.: User's manual for the distributed recharge model ZOODRM, British Geological Survey Internal Report IR/04/150, British Geological Survey, Keyworth, 2004.
- Mausser, W. and Schädlich, S.: Modeling the spatial distribution of evapotranspiration on different scales using remote sensing data, *J. Hydrol.*, 212/213, 250–267, 1998.
- Maxwell, R. and Miller, N.: Development of a coupled land surface and groundwater model, *J. Hydrometeorol.*, 6, 233–247, 2005.
- Moore, R. J.: The PDM rainfall-runoff model, *Hydrol. Earth Syst. Sci.*, 11, 483–499, doi:10.5194/hess-11-483-2007, 2007.
- Nativ, R. and Nissim, I.: Characterization of a desert aquitard – hydrologic and hydrochemical considerations, *Groundwater*, 30, 598–606, 1992.
- Nativ, R., Adar, E., Dahan, O., and Geyh, M.: Water recharge and solute transport through the vadose zone of fractured Chalk under desert conditions, *Water Resour. Res.*, 31, 253–261, 1995.
- Oleson, K., Niu, G.-Y., Yang, Z.-L., Lawrence, D. M., Thornton, P. E., Lawrence, P. J., Stöckli, R., Dickinson, R. E., Bonan, G. B., Levis, S., Dai, A., and Qian, T.: Improvements to the community land model and their impact on the hydrological cycle, *J. Geophys. Res.-Biogeo.*, 113, G01021, doi:10.1029/2007JG000562, 2008.
- Parker, S.: Chalk Regional Groundwater Models and their Applicability to Site Scale Processes, PhD thesis, Imperial College London, 301 pp., available at: https://workspace.imperial.ac.uk/ewre/Public/Parker_PhD.pdf (last access: 1 March 2015), 2011.
- Parker, S., Butler, A., and Jackson, C.: Seasonal and interannual behaviour of groundwater catchment boundaries in a Chalk aquifer, *Hydrol. Process.*, doi:10.1002/hyp.10540, 2015.
- Penman, H. L.: Natural evaporation from open water, bare soil and grass, *P. Roy. Soc. Lond. A*, 193, 120–145, 1948.
- Peters, R. and Klavetter, E.: A continuum model for water movement in an unsaturated fractured rock mass, *Water Resour. Res.*, 24, 416–430, 1988.
- Petr, J., Lipavsky, J., and Hradecka, D.: Production process in old and modern spring barley varieties, *Bodenkultur*, 53, 19–27, 2002.
- Pinault, J., Amraoui, N., and Golaz, C.: Groundwater-induced flooding in macropore-dominated hydrological system in the context of climate changes, *Water Resour. Res.*, 41, W05001, doi:10.1029/2004WR003169, 2005.
- Pitman, A.: The evolution of, and revolution in, land surface schemes designed for climate models, *Int. J. Climatol.*, 23, 479–510, 2003.
- Power, T. and Soley, R.: A comparison of chalk groundwater models in and around the River Test catchment, NC/03/05, Environment Agency of England and Wales Report, Environment Agency, Science Group, Olton, 2004.
- Price, M., Downing, R., and Edmunds, W.: The chalk as an aquifer, in: *The Hydrogeology of the Chalk of North-West Europe*, edited by: Downing, R., Price, M., and Jones, G., Clarendon Press, Oxford, 1993.
- Protopapas, A. and Bras, R.: The one-dimension approximation for infiltration in heterogeneous soils, *Water Resour. Res.*, 27, 1019–1027, 1991.
- Smith, R., Blyth, E., Finch, J., Goodchild, S., Hall, R., and Madry, S.: Soil state and surface hydrology diagnosis based on MOSES in the Met Office Nimrod nowcasting system, *Meteorol. Appl.*, 13, 89–109, 2006.
- Wagener, T., Sivapalan, M., Troch, P., and Woods, R.: Catchment classification and hydrologic similarity, *Geography Compass*, 1, 901–931, 2007.
- West, G. and Dumbleton, M.: Some observations on swallow holes and mines in the chalk, *Q. J. Eng. Geol.*, 5, 171–177, 1972.
- Wood, E., Roundy, J., Troy, T., van Beek, L., Bierkens, M., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffé, P., Kollet, S., Lehner, B., Lettenmaier, D., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., and Whitehead, P.: Hyperresolution global land surface modeling: meeting a grand challenge for monitoring Earth's terrestrial water, *Water Resour. Res.*, 47, W05301, doi:10.1029/2010WR010090, 2011.
- Yang, Z. and Niu, G.: The versatile integrator of surface and atmosphere processes – Part 1. Model description, *Global Planet. Change*, 38, 175–189, 2003.
- Yilmaz, K., Gupta, H., and Wagener, T.: A process-based diagnostic approach to model evaluation: application to the NWS distributed hydrological model, *Water Resour. Res.*, 44, W09417, doi:10.1029/2007WR006716, 2008.