*Author:*
**Batelis, Stamatis**

*Title:*
**The impact of groundwater representation in land surface models under current and future climate scenarios in Great Britain**

# *THE IMPACT OF GROUNDWATER REPRESENTATION IN LAND SURFACE MODELS UNDER CURRENT AND FUTURE CLIMATE SCENARIOS IN GREAT BRITAIN*

By

STAMATIOS CHRISTOS BATELIS


Department of Civil Engineering
UNIVERSITY OF BRISTOL

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of DOCTOR OF PHILOSOPHY in the Faculty of Engineering

November 2020

Word count: 38561

# ABSTRACT

Land Surface Models simulate the exchange of water and energy cycle between the surface and the lower atmosphere and are widely used for hydrological applications. However, they commonly ignore the impact of groundwater flow on land surface processes, usually relying on the free drainage approximation. This affects the realism of the model and consequently the water partitioning. The impact of explicitly representing groundwater in an LSM is studied in this thesis to explain and quantify its effects on the water balance components at the land surface. The Joint UK Land Environment Simulator (JULES) model is used in this analysis. The limitations of JULES to simulate runoff are first investigated in 47 British catchments. The missing groundwater representation and the low performance of JULES in groundwater dominated catchments motivated us to develop a 2-D groundwater model below JULES soil domain. The new model, called JULES Groundwater Flow Boundary (JULES-GFB), has been extensively tested against two synthetic experiments (i.e., a column-scale infiltration and a tilted-V catchment experiment) as well as applied to a real-case experiment. JULES-GFB improves soil moisture dynamics, while successfully representing lateral water flow in the saturated zone, and consequently better estimating the contribution of groundwater to river discharge, when compared to its default model version. Model evaluation of JULES-GFB in a region characterized by groundwater-dominated catchments in Great Britain shows an increase of model performance relative to catchment-integrated streamflow and evapotranspiration. The new model is used to assess the climate change impact using the UKCP18 climate projection dataset. After bias correction is applied to the meteorological data, 25 different 5-year periods are picked to represent climate scenarios with different change of precipitation and temperature compared to the climatological values. Results show that runoff could decrease up to 23% and recharge up to 90% under extreme dry and warm conditions.

# DEDICATION AND ACKNOWLEDGEMENTS

This thesis would not have finished without the generous support of many people. First, I would like to thank my two supervisors, Rafael Rosolem and Shams Rahman. They helped me shape my work and motivate me through our discussions. I thank, also, Ross Woods and Thorsten Wagener for their fruitful discussions during my PhD and because they gave me the opportunity to be part of the WISE CDT program. The implementation of JULES-GFB would have not completed without the contribution of Stefan Kollet. Finally, I thank my friends, my colleagues from Woodland Road and my family for their unconditional support and encouragement.

# AUTHOR'S DECLARATION

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... DATE: .....

# TABLE OF CONTENTS

**Page**

# LIST OF TABLES

**TABLE** **PAGE**

# LIST OF FIGURES

**FIGURE** **PAGE**

# 1 INTRODUCTION

## 1.1 Background

Water cycle is the perpetual movement of water throughout the various component of the global Earth's system (Labat, 2006) due to the solar radiation that acts as the driving factor (Quante and Matthias, 2006). Understanding the water cycle is essential to mitigate disasters caused by water and figure out the mechanisms that contribute to climate change that are related to water (Dooge, 2001). The main mechanisms of water movement through the water cycle are precipitation, infiltration, evaporation, runoff, and subsurface runoff (Chow et al., 1988).

Regardless of the importance of groundwater for water management, as it represents 97% of available freshwater resources worldwide (Guppy et al., 2018) and the second highest percentage of freshwater storage corresponding to about 30% (Gleick, 1993), it has not received as much attention in the scientific literature about the impact of climate change as the surface hydrology (Stoll et al., 2011). According to Alley et al. (2002), more than two billion people depend on groundwater supply, as their main water source, while most of the water for irrigation comes from groundwater sources (Siebert et al., 2010). Groundwater is also the last critical national resource during droughts (Famiglietti et al., 2011), thus it will be key in future water management, knowing that climate change will likely increase the frequency of drought (e.g., Lehner et al., 2006). In England, about 30% of public water is supplied by groundwater (Environment Agency, 2006); whereas there are areas, like the Southern region of England that more than 70% of public water supply comes from groundwater (Environment Agency, 2014). Hence, it is crucial to know the impact of climate change on groundwater dominated catchments.

It is known that climate change will affect recharge (e.g., Scibek and Allen, 2006; Konikow, 2011; Cuthbert et al., 2019). In the UK, there are previous studies tried to understand the impact of climate change on recharge (e.g., Herrera-Pantoja and Hiscock, 2008; Holman et al., 2009; Jackson et al., 2011). All the studies mentioned the uncertainty from the projections and the high sensitivity of their models to predict the impact of climate change on recharge. However, there is limited knowledge on how climate change will affect all the water components (e.g., recharge, runoff, soil moisture) on groundwater dominated catchments.

It is inevitable to rely on modelling of the components of the water cycle for water resources assessment and planning under changing climate conditions or land cover changes. The reason is that there are water components that are very difficult to be measured (infiltration, subsurface, evapotranspiration, runoff) or very expensive at large scale (Hari et al., 2009; Hemlin Söderberg, 2015; Potočki, 2018). However, even for variables that are easier to get measurements (precipitation, runoff), these are site specific, sparse, with many sources of errors, and many times they do not transmit the measurements in real time (e.g., Villarini et al., 2008). Remote sensing could be an alternative, but these products suffer from many sources of uncertainty, as well (Zhang et al., 2016), Thus, someone cannot avoid the use of modelling to capture the components of the water cycle.

There are many classifications for hydrological models, in terms of spatial representation (lumped, distributed or semi-distributed), temporal representation (event-based, continuous, steady state and seasonal steady state), and model structure (conceptual, physical-based, black box) (Dingman, 2015). Classically, the hydrological models have focused on the water balance ignoring energy balance, which is the driving factor of the water cycle. In that case, hydrological models impose a mass conservation constraint in their simulations but not an energy conservation constraint (Mitchell et al., 2008). Furthermore, in hydrological models there are processes, like snowmelt, heat transport in the soil, evaporation, permafrost melting that are either ignored or poorly represented due to the absence of the energy balance equation (Wild and Liepert, 2010; Dandar et al., 2017).

The models that can simulate both the exchange of water and energy cycle between the surface and the lower atmosphere are called Land Surface Models (LSMs hereafter). Due to their additional capacity to simulate the energy balance, LSMs are considered superior to models that rely only on water balance for simulating evapotranspiration (Overgaard et al., 2006). Typically, LSMs are used to simulate the interactions between climate and land surface (i.e., heat, mass, and momentum exchanges) (e.g., Dickinson, 1986). In theory, LSMs rely on

the explicit representation of processes and on data (e.g., soil, land use, forcing data) that can define the model parameters, thus they can provide continuous simulation of runoff without calibration (Beven, 2001).

Recently, the improved quality and resolution of land-surface and near-surface data from remote sensing in combination with the increase in computing power have significantly boosted the use of LSMs for hydrological modelling (Overgaard et al., 2006) and for operational applications in global hydrology (e.g., Beck et al., 2017; Sutanudjaja et al., 2018; Givati et al., 2016). For instance, in USA, the National Oceanic and Atmospheric Administration (NOAA) launched a new operational flood forecasting tool, called National Water Model (Gochis et al., 2015), in which the core of the system is an LSM (NOAH-MP). The European Centre for Medium-Range Weather Forecasts (ECMWF) applies the Land Surface Model H-TESSEL within the Integrated Forecast System (Balsamo et al., 2009).

However, even though LSMs are now more frequently used for hydrological purposes, they generally do not represent the impact of groundwater flow on land surface processes (Clark et al., 2015), which is a limitation and requires improvement. The missing groundwater representation was the reason of the inferior performance of LSMs against hydrological models in multimodel inter-comparison (e.g., Beck et al., 2017). Typically, the LSMs consider a shallow column extending vertically from the surface to a depth of 2 to 5 meters. Regarding the conditions constraining water dynamics at the bottom of the soil domain, most LSMs still apply the free drainage condition (i.e., water flow at the bottom boundary of the model domain is controlled solely by gravity). The implication of this parameterization is the exacerbation of drying because there is no physical constraint at the bottom of the soil column to prevent the water to be retained within the soil domain (Rahman et al., 2019). This will ultimately lead to reduced values of evapotranspiration flux (Niu et al., 2007). Furthermore, there is no interaction between the unsaturated zone and the water table (Maxwell and Miller, 2005).

Recently improvements have been made towards the coupling of groundwater models below LSMs (e.g., Gutowski et al., 2002; York et al., 2002; Gedney and Cox, 2003; Maxwell and Miller, 2005; Yeh and Eltahir, 2005a, b; Gulden et al., 2007; Niu et al., 2007; Fan and Miguez-Macho, 2011; Maxwell et al., 2011; Vergnes and Decharme, 2012; Tian et al., 2012; Fan et al., 2013; de Graaf et al., 2015, 2017; Guzman et al., 2015; Zeng et al., 2018; Ganji and Sushama, 2018; Koirala et al., 2018; Reinecke et al, 2019; Tian et al., 2020). These studies generally showed that adding the groundwater component led to a more realistic partitioning of water components at the land surface. In particular, LSMs that incorporate a groundwater

parameterization had better performance to simulate water storage variations against the Gravity Recovery and Climate Experiment (GRACE) water storage dataset (Niu et al., 2007; Nie et al., 2018; Sutanudjaja et al., 2018; Huang et al., 2019) and better agreement with the water table observations (Liang et al., 2003; Niu et al., 2007). The higher values of soil moisture and evapotranspiration from LSMs with groundwater component led to more accurate results against remote sensing data of evaporation (Maxwell and Miller, 2005; Danierhan et al. 2012; Nie et al., 2018). Finally, a few studies showed improved simulation of monthly runoff (Maxwell and Miller, 2005; Sutanudjaja et al., 2018; Huang et al., 2019) and daily runoff (Gedney and Cox, 2003).

Despite the recent model developments of LSMs with a groundwater scheme, the implementation of LSMs with the traditional free drainage scheme for assessing the impact of climate change is a common approach. Trenberth (2011) mentioned that the key to understand changes in the hydrological cycle due to climate change is the energy cycle, and this can only be simulated by LSMs. Furthermore, as LSMs do not heavily rely on model calibration (Beven, 2001), they can examine the impact of climate change in areas with sparse data, like Africa. For instance, Betts et al. (2018) applied JULES to simulate the change of mean runoff and mean length of flood events globally at 2 °C global warming. Koutroulis (2019) and Koutroulis et al. (2019) used JULES to estimate the impact of climate change on drylands and the exposure of freshwater stress globally under the Representative Concentration Pathway 8.5 (RCP8.5) scenario, respectively. Leng et al. (2015) analysed the impact of climate change on drought in China using the VIC model. Finally, many other studies (e.g., Hagemann et al., 2013; Dankers et al., 2014; Van Huijgevoort et al., 2014; Giuntoli et al., 2015a; Vetter et al., 2017) used a multimodel approach with both LSMs and Global Hydrological Models to assess the impact of climate change on water resources. Some of the models used in these applications (e.g., JULES, VIC, H08, ORCHIDEE) rely on the free drainage case, thus possibly their conclusions are not valid. This means that it is crucial to use LSMs that use a proper groundwater representation to accurately simulate water partition and the impact of climate change on surface-subsurface interactions. The implementation of an LSM coupled with a groundwater parameterization can help us to identify the impact of climate change on surface-subsurface interactions.

## 1.2 Research objective and scope

The Joint UK Land Environment Simulator (JULES) is a widely used land surface model, developed by the UK Met Office, and used for operational services and research to simulate the

energy, carbon and water balance between the land surface and the lower atmosphere (Best et al., 2011; Clark et al., 2011). Until now, JULES has been used mostly to assess the impact of climate change scenarios (e.g., Davie et al., 2013, Hagemann et al., 2013, Dankers et al., 2014; Prudhomme et al., 2014; Schewe et al., 2014; Papadimitriou et al., 2016). The hydrological components of JULES have been tested for runoff predictions at monthly and inter-annual scales (Gudmundsson et al., 2012a, b; MacKellar et al., 2013) and at daily resolution (Dadson and Bell, 2010; Dadson et al., 2011; Zulkafli et al., 2013; Weedon et al., 2015; Martinez-de la Torre et al., 2019). However, the current JULES model relies on the more typical free drainage assumption, like other LSMs.

The broad objective of this research is to explicitly represent groundwater processes into the JULES model and assess the climate change impact using the new parameterization that can possibly improve the model realism. Before that, we investigate the limitations of JULES model to simulate runoff under different conditions in order to assess the impact of the free drainage assumption on the runoff predictions. The focus of the study is within the UK domain which is more specific to humid catchments. This thesis focuses on addressing six main research questions:

(1) How does JULES performance to simulate runoff contrast in overland and groundwater dominated catchments?
(2) What are the dominant factors (land use, soil type, soil texture) that affect the performance of the model to simulate streamflow? Are there any missing hydrological processes in JULES?
(3) Under which conditions (e.g., water table depth, soil type) can we identify improvements in the JULES model when groundwater dynamics are explicitly represented?
(4) How do soil water dynamics, due to explicit representation of groundwater, potentially impact other hydrological fluxes at the land surface (e.g., streamflow and evapotranspiration) at the regional scale?
(5) What is the impact on water components (e.g., recharge, runoff, soil moisture) under climate change scenarios on groundwater dominated catchments?
(6) What new we can learn from the JULES model with a groundwater representation in terms of climate change?

In Figure 1.1, the flowchart of this thesis is presented. Chapter 2 presents the most used LSMs and makes a brief comparison of their characteristics and the main hydrological processes they simulate. Subsequently, there is an extensive literature review about recent

developments in groundwater models with LSMs and their main findings. Finally, studies that assess the climate change impact on UK hydrology are presented. Chapter 3 introduces the current JULES LSM version with the needed forcing data, the main hydrological processes, and the limitations it has. Focus will be given on the soil parameterization and how JULES calculates the important fluxes related to water partition, namely infiltration, drainage, runoff generation and moisture movement in the soil column. Furthermore, the new groundwater parameterization (JULES-GFB) that replaces the one-way downward free drainage flux with a two-way flux is presented, with the fundamental equations and the main assumptions. In the last part, the forcing data, the spatial data, and the data used for model evaluation are introduced. Chapter 4 evaluates the default version of JULES (v4.4) in 47 catchments in Great Britain. The scope is to investigate the strengths and the limitations of JULES model to simulate daily runoff in catchments with different dominant hydrological processes (i.e., overland and groundwater dominated). In Chapter 5, a groundwater parameterization has been coupled to JULES (JULES-GFB) and is assessed the potential impact on key hydrological variables in the model. A more complex 3-D hydrological model is used as a benchmark in a set of synthetic experiments. Then, the new model is applied in a regional domain characterized by groundwater-dominated catchments in the UK. The model's ability to represent soil moisture patterns, streamflow, and evapotranspiration within the domain is compared against physical characteristics of the regional domain, as well as observations from UK streamflow database and evapotranspiration products from remote sensing. In Chapter 6, the new model parameterization in JULES (JULES-GFB), that showed improvement in the realism of water partition, is applied to assess the impact of climate change in Central United Kingdom using climate projection data obtained from the UK Climate Projections (UKCP18). Our main interest is to identify how the climate change will affect the water cycle components and the water partition in groundwater dominated catchments. The last chapter reminds the main objective of this study. Then, it cites the most important conclusion extracted from the whole dissertation. In the last subsection, there are recommendations for further investigation related to hydrological processes and to further improvements of JULES' parameterization.

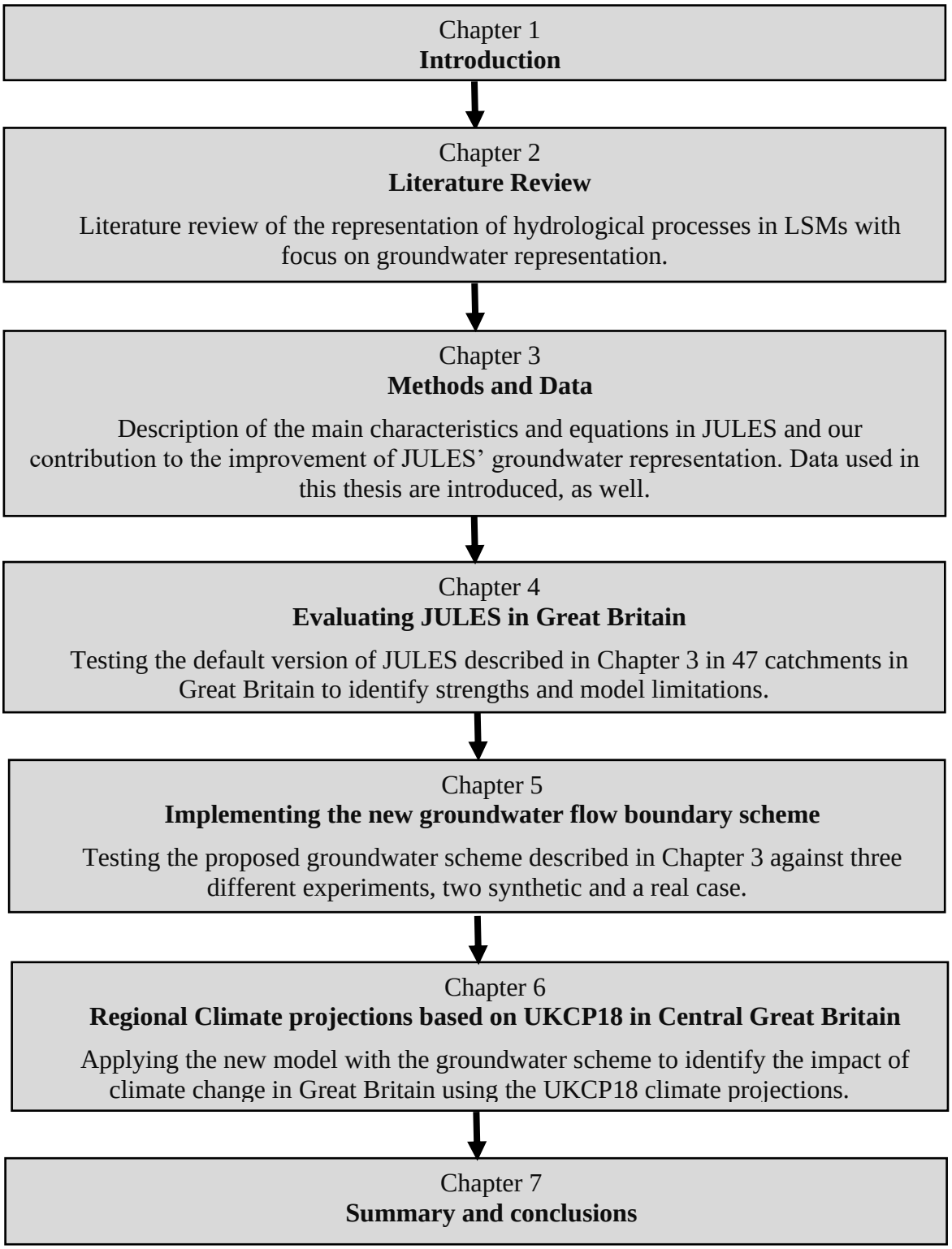


Figure 1.1: Schematic overview of the thesis organisation.

# 2 LITERATURE REVIEW

## 2.1 Hydrological mechanisms associated with runoff predictions in land surface models

The introduction of an LSM (Manabe, 1969) followed by Deardorff (1978) lead to the first generation LSMs (Dickinson, 1984; Mahrt and Pan, 1984; Sellers et al., 1986; Pan and Mahrt; 1987). Since then, many models have been developed by different research groups, while recently a number of LSMs are being used for operational purposes (e.g., National Water Model in USA (Gochis et al., 2015), JULES by Met Office in UK (Best et al., 2011), HTESSEL by ECMWF (Balsamo et al., 2009)). The most common Land Surface Models are HTESSEL (e.g., Balsamo et al., 2009; Pappenberger et al., 2010), JULES (Best et al., 2011; Clark et al., 2011) developed by UK researchers, working with UK Met Office (UKMO) and Centre for Ecology and Hydrology (CEH), Variable Infiltration Capacity (VIC) (e.g., Liang et al., 1994) developed by the University of Washington, ISBA (e.g., Habets et al., 1999) and SURFEX (e.g., Le Moigne et al., 2009) developed by the Centre National de Recherches Météorologiques, Community Land Model (CLM) (e.g. Oleson et al., 2010, 2013) developed by the National Center for Atmospheric Research (NCAR) and the CESM Land Model and Biogeochemistry Working Groups, Noah-MP (e.g., Niu et al., 2011) developed by the University of Texas at Austin, ORCHIDEE (e.g., Ducharne et al., 2016; Krinner et al., 2005) developed by the Institut Pierre Simon Laplace, MATSIRO (e.g., Takata et al., 2003) developed by the Atmosphere and Ocean Research Institute at the University of Tokyo, the National Institute of Environmental Studies, and the Frontier Research Center for Global Change in Japan, Canadian Land Surface Scheme (CLASS) (e.g., Verseghy, 1991) developed by the Canadian Centre for Climate

Modelling and Analysis and H08 (e.g., Hanasaki et al., 2008) developed by universities and institutes in Japan.

The most common method for calculating the infiltration is the ARNO scheme by Todini (1996) (e.g., SURFEX, ISBA, VIC). ORCHIDEE uses the Green and Ampt method (Green and Ampt, 1911). Regarding the interception, all the models calculate it based on the leaf area index using a single reservoir. There are two common methods to calculate evaporation in LSMs, namely the Penman-Monteith formula (Monteith, 1965) (JULES, HTESSEL, SURFEX, ISBA, VIC and Noah-MP) and the bulk transfer formula (Sene et al., 1991) (CLM, ORCHIDEE, CLASS, H08, MATSIRO).

The root water uptake and the plant water stress are important in land surface model because they affect the soil water balance. Root system produces over 50% of evapotranspiration (Luo et al., 2003). There are two approaches to simulate root water uptake, namely the microscopic model of water flow from the soil to and through the plant roots and the macroscopic empirical functions that simulate uptake using responses to water potential (Kumar et al., 2014). There are three common parameterizations in LSMs to model stomatal conductance, the empirical 'Jarvis' model (Jarvis, 1976), the phenomenological model Ball-Berry by Ball et al., (1987) and the SiB model by Sellers et al. (1986). In the first model stomatal conductance is dependent on the solar radiation, humidity, the humidity deficit, and the mean volumetric soil moisture concentration, while all the factors are independent. The Ball-Berry model uses a linear relationship between the stomatal conductance and the net photosynthesis. Finally, Sellers et al. (1986) use similar approach with the 'Jarvis' model, but they use stomatal resistance as a function of the photosynthetically active radiation (PAR). Noah, HTESSEL and VIC use the model from Jarvis (1976), Mosaic and CLM use the approach by Sellers et al. (1986) and Noah-MP uses the Ball-Berry model.

There are two runoff generation mechanisms, namely the saturation excess and the infiltration excess. In the saturation excess mechanism, runoff is generated when the soil becomes saturated. In the infiltration excess mechanism, runoff is generated when the rainfall intensity is larger than the infiltration rate (Yang et al., 2015). Most LSMs rely on both mechanisms (JULES, SURFEX, ISBA CLM, MATSIRO, Noah-MP). Exceptions are Noah, HTESSEL, ORCHIDEE which use only the infiltration excess mechanism, and H08 and VIC which rely only on the saturation excess mechanism.

However, even though LSMs are used for hydrological purposes, the hydrological processes in these models need improvement, as they generally ignore the impact of groundwater flow on

land surface processes and on surface water bodies (Kollet and Maxwell, 2008; Clark et al., 2015). The soil domain of a typical LSM usually extends vertically from the surface to a depth of 2 to 5 meters. With regards to conditions constraining water dynamics at the bottom of the soil domain, most LSMs still apply the free drainage condition (JULES, CLM 4.0, HTESSEL, Noah, VIC, H08, CLASS) or a shallow aquifer ignoring the lateral flow (MATSIRO, ORCHIDEE). Only a few of them (CLM 4.5, Noah-MP) are fully coupled with a groundwater model below the soil domain. ISBA is coupled with the routing scheme MODCOU (Ledoux et al., 1989) that is used to transfer the groundwater flow, whereas SURFEX has the choice of free drainage, constant transfer time or groundwater diffusive scheme.

The Joint UK Land Environment Simulator (JULES) is a widely used land surface model, developed by the UK Met Office (Best et al., 2011; Clark et al., 2011). The hydrological components of JULES have been tested for runoff predictions at monthly and inter-annual scales (Gudmundsson et al., 2012a, b; MacKellar et al., 2013) and at daily resolution (Dadson and Bell, 2010; Dadson et al., 2011; Zulkafli et al., 2013; Weedon et al., 2015; Martinez-de la Torre et al., 2019). JULES showed good performance in inter-model comparison with other LSMs for simulating the inter-annual variability of observed runoff in Europe (Gudmundsson et al., 2012a) and the mean annual runoff (Gudmundsson et al., 2012b). Especially, for the first case, JULES was ranked 2nd out of 10 LSMs that shows that JULES has significant predictive power compared to other models. In another intercomparison project (Schellekens et al., 2017), 10 global hydrological and LSMs were tested against benchmark datasets from The International Land Model Benchmarking Project (ILAMB) (Collier et al., 2018). The results showed that there is no single best model for all variables and the model performance is spatially variable based on climate conditions. Zulkafli et al. (2013) used JULES to simulate the hydrological fluxes in daily time step at a humid tropical mountain basin of the Peruvian Andes-Amazon (360,000 $km^2$), but the results were poor. According to their findings, JULES constantly underestimated the observed runoff and failed to simulate most of the flood events. Researchers attributed the low model performance in two reasons, firstly in the water balance of JULES, as they found cases in which runoff ratio was above one, and secondly in the inability of JULES to represent baseflow. Dadson and Bell (2010) compared the two river flow routing schemes of JULES, namely TRIP and RFM, for 10 large catchments using daily data. The model performance was poor with negative Nash Sutcliffe Efficiency (NSE) values even after calibration and only for one catchment the NSE was higher than 0.5, that is usually assumed as an acceptable performance (Moriasi et al., 2007). Dadson et al. (2011) coupled JULES with a

flow routing model (RFM) and they tested its performance in six large catchments in Europe. The study claims that adding some processes in JULES, like the soil moisture transport, the explicit representation of groundwater stores and a better model for the snowpack can lead to more accurate predictions.

Other studies evaluated the sub-grid parameterization schemes of JULES, namely PDM and TOPMODEL. These two options are used to account for subgrid heterogeneity of soil moisture and saturation excess runoff. MacKellar et al. (2013) examined the sensitivity of JULES in PDM and TOPMODEL and the impact when there is no sub-grid runoff parameterization. They tested three catchments in Africa, and they analysed average monthly values. They found that there was not a single parameterization that gives optimal results across all the catchments. Similar work was done by Clark and Gedney (2008) in three subcatchments of the Rhone basin using the Met Office Surface Exchange System (MOSES; a previous land surface model version which formed the basis for the JULES development). Clark and Gedney (2008) found that the model without the PDM or TOPMODEL scheme underestimated runoff (NSE ranged between 0.4 and 0.6), but the performance increased significantly after applying and calibrating one of the two sub-grid parameterizations (PDM and TOPMODEL) (NSE ranged between 0.7 and 0.9). Weedon et al. (2015) applied nine distributed hydrological models and LSMs, including JULES, to simulate the daily runoff at the Thames catchment, that is a chalk groundwater dominated catchment. The evaluation was based on the cross spectral analysis and they found that JULES performance depends on the configuration used to introduce sub-gridscale heterogeneity into the soil moisture (PDM, TOPMODEL), with TOPMODEL producing a slightly better performance (i.e., NSE for JULES-TOPMODEL found equal to 0.66, for JULES equal to 0.53 and for JULES-PDM equal to 0.38).

Finally, Largeron et al., (2018) and Martinez-de la Torre et al. (2019) made improvements in the JULES model parameterization of the runoff generation mechanisms. Largeron et al. (2018) tested different infiltration schemes to improve the performance of the model during intense rainfall events in UK catchments. Their proposed representation assumes that the infiltration rate is dictated by the hydraulic conductivity of the top layer at a specific time instead of the saturated conductivity that is a fixed value. This representation worked better for wet catchments with annual mean rainfall more than 1200 mm, because in dry catchments the hydraulic conductivity of the upper layer is low, so the scheme produces more runoff with a tendency of overestimation. Martinez-de la Torre et al. (2019) used another parameterization based on terrain slope to constrain surface runoff in order to improve the daily simulation of

runoff for 13 catchments in UK. In their method, they tried to constrain the *b* exponent of PDM that dictates the heterogeneity of soil moisture and the amount of runoff produced by the saturation excess mechanism, using the terrain slope as the determinant. Despite the improvement for almost all the catchments, it seems that for some groundwater dominated catchments the acceptable results were for the wrong reasons. In these catchments, the driving factor is the geology so the improvement of the model due to the different parameterization of saturation excess mechanism does not explain the groundwater component.

As mentioned above, JULES relies on the traditional free drainage assumption. Le Vine et al. (2016) made the first attempt to couple a groundwater model with JULES in a chalk groundwater dominated catchment. They coupled the ZOOMQ3D groundwater model (Jackson, 2001) with JULES to simulate the Kennet catchment, a tributary of the Thames River. They extended the soil depth from 3 to 6 meters and they used the recharge from JULES as the upper boundary condition to the groundwater model offline. After an extensive calibration, they managed to improve the water balance, the soil moisture, and the runoff simulation. The new model achieved NSE values above 0.8 for almost all the catchments against 10-day average flows, while it captured the groundwater levels in four positions and the soil moisture for depth up to 4 meters satisfactorily. However, with this type of coupling, groundwater dynamics are not explicitly represented in the model. Rather, recharge from standalone JULES is fed to the groundwater model directly in an uncoupled mode. In that case, there is no feedback between the groundwater model and JULES. Thus, the approach proposed by Le Vine et al. (2016) does not simulate the feedbacks between saturated and unsaturated zones because groundwater cannot affect the land surface processes.

## 2.2 Representing groundwater processes in Land Surface Models

There are recent developments of coupling groundwater models to an LSM (e.g., Gutowski et al., 2002; York et al., 2002; Gedney and Cox, 2003; Maxwell and Miller, 2005; Yeh and Eltahir, 2005a, b; Gulden et al., 2007; Niu et al., 2007; Fan and Miguez-Macho, 2011; Maxwell et al., 2011; Vergnes and Decharme, 2012; Tian et al., 2012; Fan et al., 2013;Guzman et al., 2015; de Graaf et al., 2015, 2017; Zeng et al., 2018; Ganji and Sushama, 2018; Koirala et al., 2018; Reinecke et al, 2019; Tian et al., 2019). Regarding their dimension, this can be either one-dimensional model (e.g. Gedney and Cox, 2003; Maxwell and Miller, 2005; Yeh and Eltahir, 2005; Niu et al., 2007; Huang et al., 2019), two-dimensional (e.g. Vergnes and Decharme, 2012, Fan and Miguez-Macho, 2011) or three-dimensional groundwater model (e.g.

Gutowski et al., 2002; York et al., 2002; Tian et al., 2012; Maxwell et al., 2011). In the coupling with one-dimensional groundwater model (Figure 2.1, left), the soil domain, which usually ranges between 2 and 5 m, is extrapolated for a few tens of meters. In that case, the free drainage flux, that is always downward and dries the soil domain, is replaced by a two-way flux that can retain or add water in the soil domain if the water table is inside it. However, lateral flux is not represented in this type of coupling. In the coupling with a two-dimensional groundwater model (Figure 2.1, middle), the total head of the aquifer or the recharge is calculated based on the total head or the recharge of the neighbouring cells solving the transient flow equation for two dimensions. In that case, the lateral flow is included in the calculation of the recharge flux that interacts between the aquifer and the soil domain of the LSM. However, the lateral flow is still not represented in the soil domain. The aquifer in that case is represented by a single layer. Finally, in the coupling with a three-dimensional groundwater model, the lateral flow and the vertical flux are calculated based on the neighbouring cells throughout the soil column, from the soil to the aquifer for both saturated and unsaturated zones. In that case, the aquifer is not represented only by a single layer as in the coupling with the two-dimensional groundwater model.

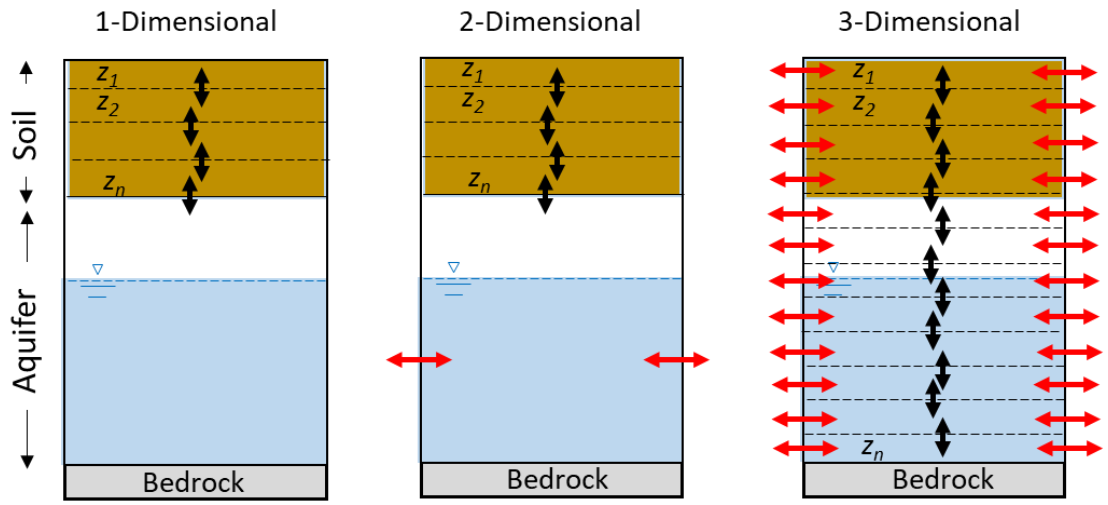


Figure 2.1: Schematic overview of 1-Dimensional, 2-Dimensional and 3-Dimensional groundwater models coupled with an Land Surface Model. Solid black arrows show the interaction between soil and aquifer and red arrows show the lateral flux.

The baseflow for all the types of groundwater models is usually extracted from the aquifer, as a percentage of the water table depth using exponential equations (e.g., Niu et al., 2007). Considering that the horizontal transport of groundwater is important on smaller spatial scales,

this makes the use of a three-dimensional model suitable for regional studies and the use of one-dimensional model suitable for global studies (Niu et al., 2007; Bierkens et al., 2015). Finally, some schemes were implemented in uncoupled mode, as recharge from LSMs was the input to groundwater models or the output flux of the groundwater model feeds the bottom boundary condition of the LSM (e.g., de Graaf et al., 2015, 2017; Fan and Miguez-Macho, 2011). Both models run independently, and, in this way, there is not real time interaction between saturated and unsaturated zones.

Starting with the one-dimensional model developments, Gulden et al. (2007) coupled the CLM LSM with three configurations of vertical water flow; (1) a shallow soil profile that is common on LSM; (2) a deeper soil profile; (3) and a lumped, unconfined aquifer coupled with the shallow soil profile. They sampled the soil texture parameters with reasonably range of values, and they applied Monte Carlo simulation to test the sensitivity of the model to represent the monthly terrestrial water storage change in Illinois. They showed that the explicit aquifer representation (3$^{rd}$ configuration) was more robust and less sensitive to the chosen parameters compared to the other two configurations. Other model configurations (Niu et al., 2007; Nie et al., 2018; Sutanudjaja et al., 2018; Huang et al., 2019) that incorporate a groundwater representation showed improvement of the simulated water storage variations when compared against the Gravity Recovery and Climate Experiment (GRACE) water storage dataset. Furthermore, better agreement with the water table observations was achieved by Liang et al. (2003) and Niu et al. (2007), while it was detected wetter soil moisture profiles in global scale and more evapotranspiration produced by the coupled scheme (Niu et al., 2007; Huang et al., 2019). The higher values of soil moisture and evapotranspiration led to more accurate results against remote sensing data (Maxwell and Miller, 2005; Danierhan et al. 2012; Nie et al., 2018). Interestingly, although the lateral flow is not represented in the one-dimensional model developments, a few studies showed improved performance of monthly runoff (Maxwell and Miller, 2005; Sutanudjaja et al., 2018; Huang et al., 2019) and daily runoff (Gedney and Cox, 2003). Schlemmer et al. (2018) used Richards equation with an extended soil column in TERRA_ML LSM in order to represent topography in a crude way, ignoring the impact of lateral flow. Their hypothesis was that the absence of topography and lateral flow is linked with biases of near-surface temperatures and temperature lapse rate because the mountain tops are represented too wet, as the precipitation is higher, and the valleys drier. They tested their model for three sites against annual soil water content data for the upper 50 cm with small improvement of the new scheme compared to the default one. Furthermore, they tested the

mean seasonal temperature for a 10-year period over Europe against observations showing significant improvement.

Looking at the two-dimensional applications, Ganji and Sushama (2018) examined the impact of using one-dimensional or two-dimensional groundwater model below an LSM. They coupled a modified version of a groundwater model developed by Niu et al. (2007) below the LSM CLASS. They compared three different model configurations, namely the default gravitational drainage against the coupled groundwater model with and without lateral flow. They applied this study for a domain analysis in Canada that consists of 21 catchments. Ganji and Sushama (2018) found that the aquifer affects the near surface temperature and the evapotranspiration with lower values for the free drainage case in the summer season. The changes of temperature and evapotranspiration against the free drainage case are more significant for the configuration that considers the lateral flow. Regarding the simulation of the streamflow, there is improvement for the case of the groundwater scheme with lateral flow, especially during the dry period in which the baseflow contribution is low for the groundwater scheme without the lateral flow and even lower for the free gravitational drainage case. Fan and Miguez-Macho (2011) developed a 2D groundwater model to capture wetlands using the groundwater table depth. The limitation of this study is that there is not a dynamic real-time coupling between LSM and groundwater model, but these models run independently. The recharge came from mass balance equation based on observed precipitation and simulated runoff by three LSMs in the Global Land Data Assimilation System (GLDAS) (Rodell et al. 2004). This means that there is no feedback between the groundwater model and LSMs. Vergnes and Decharme (2012) coupled a simple groundwater scheme in the TRIP river routing model using the fluxes for surface runoff and drainage from the Interactions between Soil-Biosphere-Atmosphere (ISBA) LSM. The new scheme was evaluated against 3,500 gauges globally and they found that groundwater representation improved the model predictions in 70% of the cases and led to deterioration at 15% gauges. Guzman et al. (2015) coupled the physical based model SWAT with the groundwater model MODFLOW for a catchment in Oklahoma (780 $km^2$) showing agreement of modelled runoff and groundwater depth with observations.

Finally, considering the three-dimensional model developments, York et al. (2002) coupled a single column vertically discretized atmospheric model with a three-dimensional distributed physically based soil-vegetation-aquifer model and tested in a catchment in Kansas for monthly timescale and for a nine-year period. The new scheme was able to simulate the trends of

monthly and annual runoff, groundwater depth, evapotranspiration, and soil moisture. Their important finding is that about 5-20% of evapotranspiration comes from the aquifer. Maxwell et al. (2011) coupled the three-dimensional ParFlow with the Weather Research Forecasting (WRF) atmospheric model finding improvement of rainfall-runoff predictions. In a similar project, Gutowski et al. (2002) developed a model that simulates hydrologic processes from the atmosphere to aquifer including a groundwater model. They claim that the importance of using a groundwater model is proved by the fact that in dry periods up to 33% of monthly evapotranspiration comes from the aquifer. This finding is consistent with York et al. (2002). Tian et al. (2012) developed a three-dimensional variable saturated groundwater model (AquiferFlow) and coupled it with an LSM (SiB2). They validated it in a catchment in China (about 10,000 $km^2$) against evapotranspiration and relative saturation data from three stations.

Overall, most of the papers that introduce another parameterization for groundwater are evaluated in large regions, usually in global scale (e.g., Fan et al., 2013; de Graaf et al., 2015, 2017; Zeng et al., 2018) against global products (e.g., Niu et al., 2007; Nie et al., 2018), in steady state (e.g., Fan et al., 2013; de Graaf et al., 2015, 2017; Zeng et al., 2018) or over a couple of catchments (e.g., Liang et al., 2003; Maxwell and Miller, 2005) in monthly timescale (York et al., 2002; Yeh and Eltahir, 2005a, b ). All studies lead to a more realistic representation of the hydrological processes, especially evaporation and soil moisture.

In the traditional LSMs that apply the free drainage case, the lack of groundwater representation leads to bias of soil moisture and evapotranspiration fluxes because the models are unable to capture the interactions between the unsaturated zone and the water table (Maxwell and Miller, 2005). The one-way downward flux (free drainage) can only dry the soil domain and especially in areas with sandy soil types (Rahman et al, 2019), as there is no feedback from the aquifer below that could retain the water from drying or make the soil domain wetter due to the groundwater discharge. This implies lower soil moisture values in the soil column and ultimately lower evapotranspiration fluxes (e.g., Niu et al., 2007; Huang et al., 2019). As a consequence, studies that apply LSMs with the free drainage assumption globally tend to underestimate soil moisture and evaporation (Zeng et al., 2018), while mountain tops are represented too wet, because the precipitation is higher, and the valleys drier due to the absence of lateral flow (Schlemmer et al., 2018). Gutowski et al. (2002) and York et al. (2002) showed that during dry periods up to one third of evaporation comes from groundwater. Thus, it is crucial to incorporate groundwater representation in LSMs in order to get an accurate water partition. However, there are many studies about climate change that apply these kind of models

(e.g., Hagemann et al., 2013; Dankers et al., 2014; Huijgevoort et al., 2014; Giuntoli et al., 2015; Vetter et al., 2017) and possibly their results could be questionable. Interestingly, even for the one-dimensional groundwater configurations that ignore lateral flow, many studies found improvement of the near surface soil moisture and evaporation (e.g., Maxwell and Miller, 2005; Danierhan et al., 2012; Nie et al., 2018). It seems that extrapolating the soil column a few meters is enough to improve the water partition in the soil domain, as it allows the feedback between aquifer and soil domain. In addition, the introduction of the lateral flow improves the model predictions of the surface soil moisture and the evapotranspiration fluxes significantly (de Graaf et al., 2015, 2017; Fan and Miguez-Macho, 2011), while configurations with lateral flow outperform those without (e.g., Ganji and Sushama, 2018). Finally, inaccurate representation of soil moisture and the lack of lateral flow representation is, also, linked with biases of runoff predictions (e.g., Maxwell and Miller, 2005; Vergnes and Decharme, 2012) and near-surface temperatures and temperature lapse rate (Schlemmer et al., 2018).

## 2.3 Climate change impacts on UK hydrology

A common approach to evaluate the impact of climate change in the UK is the implementation of future climate scenarios (e.g., UKCIP02, UKCP09) produced by the UK Climate Impacts Programme (UKCIP). UKCIP was established in 1997 to support decisions-makers' assessments (Hedger et al., 2006) and is fully funded by the Department for Environment, Food and Rural Affairs (Defra) (Harman et al., 2005). The UKCIP02 (Hulme et al., 2002) climate scenario compromises of four scenarios (low, medium-low, medium-high, and high), representing the four Intergovernmental Panel on Climate Change Special Report Emissions Scenarios (IPCC SRES) (Nakicenovic et al., 2000). Each emission scenario assumes different rate of future population and economic growth (B1, B2, A2, AlFl) (Arnell, 2004). The UKCP09 dataset (Jenkins et al., 2010) produces probabilistic projections that incorporate uncertainties from different sources (e.g., natural variability, climate models) (Christierson et al., 2012), instead of previous UK climate scenarios (e.g., UKCIP02). UKCP09 relies on the low, medium, and high emissions scenarios that correspond respectively to the IPCC SRES B1, A1B and A1FI emission scenarios (Murphy et al., 2009). The UKCP18 dataset reflects the current information and the current hypotheses for the climate change projection based on the 5th assessment report from IPCC. However, the literature review of using UKCP18 dataset in peer-reviewed papers is restricted. There are only two reports from Met Office (Met Office, 2018a, b).

Regarding the UKCP09 projections, Christierson et al. (2012) utilized them to assess the impact of climate change on river flows for 70 catchments in the UK for the 2020s under the A1B scenario (medium emission scenario) using two conceptual models (i.e., PDM and CATCHMOD (Environment Agency, 2005) and the Generalized Likelihood Uncertainty Estimation (GLUE) methodology (Beven and Binley, 1992). Their main finding is the significant decrease of summer flows, while a decrease in river flows all year round is detectable for all the catchments. Also, they highlighted that uncertainties in climate modelling, especially during summer, are large. Prudhomme et al. (2012) used the semi-distributed hydrological model CERF (Young, 2006) and applied the change factors method (Hay et al., 2000) using the UKCP09 dataset under the medium emission scenario to examine the changes in river flows in the 2050s. They found a mixed pattern of changes in winter flows from -20% to +40% change in England and Wales and ±20% in Scotland. In summer season, they found predominantly decreases in the UK that ranged from +20% to -80%. Charlton et al. (2014) examined the impact of climate change in six British catchments using the Cat-PDM (Arnell, 2011) conceptual model. Charlton et al. (2014) found that changes of high flows are driven by changes in winter precipitation with low sensitivity to changes in temperature and potential evaporation. However, changes of low flows are dictated by both summer precipitation and temperature. Cloke et al. (2010) used the A1B scenario of UKCP09 projections to identify the impact on river flow for the Medway catchment in UK, using the CATCHMOD model. The results showed decline of flows, especially for summer months. Coulthard et al. (2012) used the hourly rainfall from UKCP09 weather generator to simulate geomorphic changes until 2099 in the Swale catchment in UK, finding 100% increase of sediment yield. Kay and Jones (2012) compared the capability of three products from UKCP09 to assess the flood frequency in nine catchments in UK. They found good agreement between the products for most of the catchments. Finally, Bell et al. (2012) used the UKCP09 projections, as input to the Grid-to-Grid model, to simulate the river flow at the Thames catchment finding a change about 36% of the modelled 20-year return flood period by the 2080s. Overall, we could say that almost all the papers applied the UKCP09 dataset mainly focus on the surface hydrology and not on groundwater hydrology.

There are only a few studies in the UK that do not rely on UKCIP products (e.g., Pilling and Jones, 1999) or apply trend analysis only on observed rainfall and runoff data (e.g., Wilby, 2006). Pilling and Jones (1999) implemented the physical process-based hydrological model (HYSIM) (Manley, 1977). They used the climate change scenarios from the Hadley Centre's

high-resolution equilibrium global circulation model (UKHI) for 2050 and transient global circulation model (UKTR) for 2065. Winter runoff is expected to increase in northern Britain under both scenarios, while summer runoff will experience major reductions more than 15% in England and Wales under the latter scenario. Wilby (2006) utilised monthly precipitation and runoff data for the period between 1865 and 2002 for 15 rivers in England and Wales. Interestingly, they found significant long-term increase in winter flow in only three rivers.

Additional studies have attempted to understand the impacts of climate change on recharge in the UK (e.g., Herrera-Pantoja and Hiscock, 2008; Holman et al., 2009; Jackson et al., 2011). Herrera-Pantoja and Hiscock (2008) predicted decrease of potential groundwater recharge from 7% to 40% in three different catchments for the future "high" gas emissions scenario (UKCIP02) using the water balance equation. They found that during the drying season, recharge can decrease up to 88%. Holman et al. (2009) applied a daily soil-water balance model and climate-change weather projections from the IPCC 4th assessment report. They found about 30% of decrease of annual potential recharge for the "high" emission scenario in 2050s. Jackson et al. (2011) implemented the ZOODRM model (Mansour and Hughes, 2004) in a chalk-dominated area of 2,600 $km^2$ in UK. The use of 13 different Global Circulation Models (GCM) gave contradictory results from -26% to +31% for the mean annual potential recharge, while the ensemble mean predicted only 5% decrease of potential recharge. Cooper et al. (1995) investigated the effects of global warming on groundwater recharge, storage and river baseflow, using an idealized representation of the aquifer/river system (Wilkinson and Cooper, 1993) for the two principal aquifer in the United Kingdom, namely Chalk and Triassic sandstone aquifers. Two different climate change scenarios were used, based on the Climate Change Impacts Review Group (CCIRG (Climate Change Impacts Review Group), 1991). Under the high evaporation scenario, baseflow and recharge are expected to decrease up to 40% and 21%, respectively, in the Chalk aquifer, and by around 15% and 13% in the Triassic aquifer.

All these studies mentioned the uncertainty from the projections and the high sensitivity of their models to predict the impact of climate change. However, there is limited knowledge on how climate change would affect all the water balance components on groundwater dominated catchments, e.g., soil moisture and evaporation, because previous studies largely focused only on runoff and recharge. This thesis aims to identify the impacts of climate change on surface-subsurface interactions, as new groundwater parameterization in JULES model represents dynamically the feedback between soil domain and groundwater.

# 3 METHODS AND DATA

This chapter is part of a paper published in the Hydrological Processes Journal. This paper has been modified to enhance consistency all through this dissertation.

**Citation:** Batelis, S.C., Rahman, M., Kollet, S., Woods, R. and Rosolem, R., 2020. Towards the representation of groundwater in the Joint UK Land Environment Simulator. Hydrological Processes. https://doi.org/10.1002/hyp.13767

This chapter introduces the meteorological forcing data, the ancillary data and the data used for model evaluation. In addition, it is presented the current JULES LSM version with the needed forcing data, the main hydrological processes, and the limitations it has. Focus will be given on the soil hydrology parameterization and how JULES calculates the important fluxes related to water partition, namely infiltration, drainage, runoff generation and moisture movement in the soil column. Then, the river routing scheme (RFM) in JULES with the parameters and their optimal values are described. Furthermore, the new groundwater parameterization (JULES-GFB) is presented, with the fundamental equations and the main assumptions. In the last part, the metrics used in this study to evaluate the model performance are presented.

## 3.1 Data

### 3.1.1 Meteorological Forcing Data

The regional forcing data are obtained from the Climate, Hydrology and Ecology Research Support System (CHESS) dataset (Robinson et al., 2016) that is freely available. CHESS data are available at daily temporal resolution and 1 km spatial resolution covering the entire Great Britain. The CHESS data stem from the Met Office 40 km gridded MORECS dataset (Thomson et al., 1981; Hough et al., 1997), from the CRU TS 3.21 dataset (Harris et al., 2014) and from the WATCH Forcing Data dataset (Weedon et al., 2011). All the CHESS data were downscaled to 1 km grid based on topographic information (Robinson et al., 2016). The forcing data are downscaled from daily to the simulation time-step using the downscaling tool of JULES, described in Williams and Clark (2014), and the daily temperature range obtained from CHESS database, as well.

### 3.1.2 Ancillary Data

The elevation map is derived from HydroSHEDS (Hydrological data and maps based on Shuttle Elevation Derivatives at multiple Scales) (Lehner et al., 2008) with 1 km spatial resolution. The elevation map is post-processed in order to provide flow directions following the methodology presented in Maxwell et al. (2009). Flow direction is important for the river routing scheme (RFM) and for the groundwater flow in the new parameterization. The Land Cover Map 2007 (LCM2007) (Morton et al., 2011) is used to obtain spatially distributed land cover type information. We used the Land Cover Map 2007 as representative of reality considering that the simulation time was for the period between 2006 and 2015 for the fourth Chapter and for the period between 2008 and 2012 for the fifth Chapter. Land cover types are required by JULES to define parameters related to vegetation, such as the canopy capacity, infiltration enhancement factor or root depth (Best et al., 2011, Table 5). In addition, soil albedo values are derived from the land use map based on Houldcroft et al. (2009). Soil parameters are derived for the regional domain from the Land Information System provided by the Cranfield University. This dataset is based on field work, like 250,000 auger bore records, 8,000 soil profile descriptions and 15,000 filed records, and covers both England and Wales (Hallett et al., 2017). This dataset provides with all the necessary soil ancillary data required by the JULES model and are described in Section 3.2.1. Note that the Land Information System provides data up to 1.5 to 2 m depth. However, we use these data for the whole soil domain (3 m) in JULES model by assuming the same values.

### 3.1.3 Validation data

We evaluate runoff predictions in the analysis of JULES Default performance (Chapter 4) and in the regional analysis during the JULES-GFB testing (Chapter 5) using daily streamflow observations from the National River Flow Archive (NRFA). NRFA provides with openly available streamflow data for more than 1500 catchments in UK (https://nrfa.ceh.ac.uk/data last visited on 03/09/2020). Furthermore, it provides with catchment characteristics (e.g., catchment size, climate, geology, and land cover) and information about the hydrological regime (e.g., runoff percentiles, BFI, factors that affect runoff). This dataset has been extensively used in hydrological applications (e.g., Lane et al., 2019; Martinez-de la Torre et al., 2019).

Finally, we evaluate JULES-GFB against the default model in representing the evapotranspiration (Chapter 5) against the daily Global Land Evaporation Amsterdam Model (GLEAM) (Miralles et al., 2011; Martens et al., 2018) dataset (version 3.3a). GLEAM is a simple bucket-type soil hydrological model (with no groundwater) that calculates evapotranspiration from the water balance equation, based on satellite data to estimate the different components of the water cycle. Although remote sensing data are a large component of this product, ultimately the model dictates how the water balance components are distributed, and this is a limitation of this dataset. However, this product was validated against 43 FLUXNET stations for different canopies and climate conditions (Miralles et al., 2011), showing an average correlation of about 0.8 with -5% bias (20 mm $y^{-1}$). The dataset is freely available and widely used (e.g., Forzieri et al., 2017; Martens et al., 2018).

## 3.2 Joint UK Land Environment Simulator (JULES)

JULES is a community LSM and can simulate the energy cycle, the carbon cycle and the water balance between the land surface and the lower atmosphere (https://jules.jchmr.org/ last visited on 03/09/2020). A detailed description of the model can be found in Best et al. (2011) and Clark et al. (2011). The JULES model requires eight meteorological forcing data, namely the wind speed, air temperature, surface downwelling longwave and shortwave radiation, specific humidity, pressure, rainfall rate and snowfall rate. Ancillary data include land cover and soil types. Land cover data are used to link regional characteristics to vegetation parameters (Clark et al., 2011, Table 1, 2) and the soil type map is used to prescribe soil hydrology parameters, as described in Best et al. (2011, Table 3). JULES uses 9 categories of land use, namely Broadleaf forest, needleleaf forest, C3 grasses, C4 grasses, Shrubs, Urban, Inland water,

Bare soil and Snow/ice. For the soil data, JULES needs the hydraulic conductivity in saturation (mm $s^{-1}$), the volumetric soil moisture content at saturation, at critical point and at the wilting point ($m^3$ water per $m^3$ soil), dry heat capacity (J $m^{-3}$ $K^{-1}$), dry thermal conductivity (W $m^{-1}$ $K^{-1}$), soil albedo and two parameters of the van Genuchten model.

In relation to surface runoff generation, there are two mechanisms to produce runoff, namely the infiltration excess and saturation excess mechanisms. The infiltration excess runoff is calculated considering both the throughfall and the grid-box mean infiltration (Best et al., 2011) based on the equations by Johannes Dolman and Gregory (1992). Infiltration excess occurs when the rainfall rate is higher than the hydraulic conductivity of the topsoil. To account for the effect of a finite model timestep on the throughfall and the surface runoff, the hydraulic conductivity of the topsoil is multiplied by an enhancement factor which is fixed and equal to four for broadleaf and needleleaf trees land-use type and two for grasses and shrubs (Best et al., 2011).

JULES has two options to account for subgrid heterogeneity of soil moisture and saturation excess runoff, namely TOPMODEL and PDM. The first one, and the more complicated one, is TOPMODEL that describes heterogeneity throughout the soil column. In TOPMODEL (e.g., Beven and Kirkby, 1979), a topographic index, $\lambda$, is required based on the digital elevation model. Sub-grid variation in $\lambda$ is simulated using a gamma distribution (Best et al. 2011). The saturated fraction of a grid-box that is saturated is calculated from soil moisture conditions in the soil profile. The saturation excess runoff is calculated based on the saturated fraction (MacKellar et al., 2013). The second scheme is the Probability Distribution Model (PDM) (Moore, 2007) that does not need additional information (such as topography), like TOPMODEL, and it describes heterogeneity in the topsoil layer (1 m). The PDM scheme requires the prescription of two parameters, namely the depth of the topsoil ($dz_{PDM}$) and an exponent coefficient for the Pareto distribution of soil water holding capacity ($b_{PDM}$). The default values for $dz_{PDM}$ and $b_{PDM}$ are 1 m and 1 (unitless), respectively. According to the theory, higher values of $b_{PDM}$ increase the runoff, because for a given soil moisture storage, a higher $b_{PDM}$ leads to higher saturated fraction (Fig. 6a of Moore, 1985), which leads to higher surface runoff (Clark and Gedney, 2008). It is found $b_{PDM}$ to be more sensitive than the $dz_{PDM}$ (Bakopoulou, 2015), whereas $dz_{PDM}$ is usually ignored (Martinez-de la Torre, 2019) or constrained by data (Dadson et al., 2011; MacKellar et al., 2013). In case that JULES underestimates the runoff, that is a common thing, a higher value of $b_{PDM}$ is expected to give better performance (Martinez-de la Torre, 2019).

The default JULES (JULES Default) model has four soil layers defined with different thickness, namely 0.10, 0.25, 0.65 and 2 m. We have modified the vertical discretization of the default model to be more consistent with the new groundwater model. In this case, all JULES model versions discussed in Chapter 5 and 6 with or without the groundwater parameterization have been set up with evenly spaced 10 or 20 cm thick soil layers from surface to 3 m depth. We refer to this version of the JULES model as JULES-FD (free drainage) without the groundwater parameterization and as JULES-GFB with the groundwater parameterization. In Chapter 4, the JULES Default is used with the default four soil layers (i.e., 0.1, 0.25, 0.65 and 2.0 m thickness) in order to evaluate the model under different conditions.

The main reason why we changed the vertical discretization for the groundwater model was the numerical instabilities we found when the default discretization (i.e., 0.1, 0.25, 0.65 and 2.0 m thickness) or the CLM structure (i.e., soil thickness increases exponentially for deeper depth) were used. In the explicit equations (like in JULES), variables are calculated directly from the previous time step without iterations for convergence of the solution. To achieve stability in the solution, it is crucial to keep the timestep and the vertical discretization small, otherwise fluctuations will occur, and they will be amplified, as the model continues running. In the preliminary analysis, we tried evenly spaced 50 cm thick soil layers from surface to 3 m depth, but we did not avoid again numerical inconsistencies. Finally, we ended up with 10 and 20 cm evenly spaced soil layers that provided us with robust solution.

### 3.2.1 Soil Hydrology in JULES Default

At each node *n*, the soil water content $\theta_n$ is updated using the one-dimensional finite difference form of the Richards equation to estimate the transport of moisture. Richards' equation is the combination of continuity (3.1) and Darcy's law (3.2) that models the vertical fluxes.

$$\frac{d\theta_n}{dt} = W_{n-1}^{'} - W_n^{'} - E_n^{'} \tag{3.1}$$

$$W^{'} = K_h\left(\frac{\partial \Psi}{\partial z} + 1\right) \tag{3.2}$$

where $W_{n-1}'$ and $W_n'$ are the diffusive fluxes flowing in from the upper layer, and to the layer below, respectively. $E_n'$ is the evapotranspiration extracted by plant roots in the layer, $K_h$ is the hydraulic conductivity (mm $s^{-1}$), $\Psi$ is the soil water suction (m) and $z$ is the soil depth (m). In

Equation (3.1), the top boundary condition is the infiltration of water at the surface and the bottom boundary condition is the free drainage, which contributes to the subsurface runoff (Best et al., 2011). The water table and the lateral flow in the saturated zone are not explicitly represented in the default JULES that rely on the free drainage case as the bottom boundary condition (Figure 3.1a).

The amount of moisture accessed by vegetation at each layer in the soil is determined by root density, assumed to follow an exponential distribution with depth. Each vegetation surface has different root depth with C3 and C4 grasses, and Shrubs to have 0.5 m root depth, Needleleaf trees 1 m and broadleaf trees 3 m. Furthermore, leaf photosynthesis is linked to the potential (non-stressed) lead photosynthesis based on the beta ($\beta$) parameter (unitless) that defines the stress factor based on soil moisture concentration in the root zone, and the critical and wilting point concentrations (Best et al., 2011).

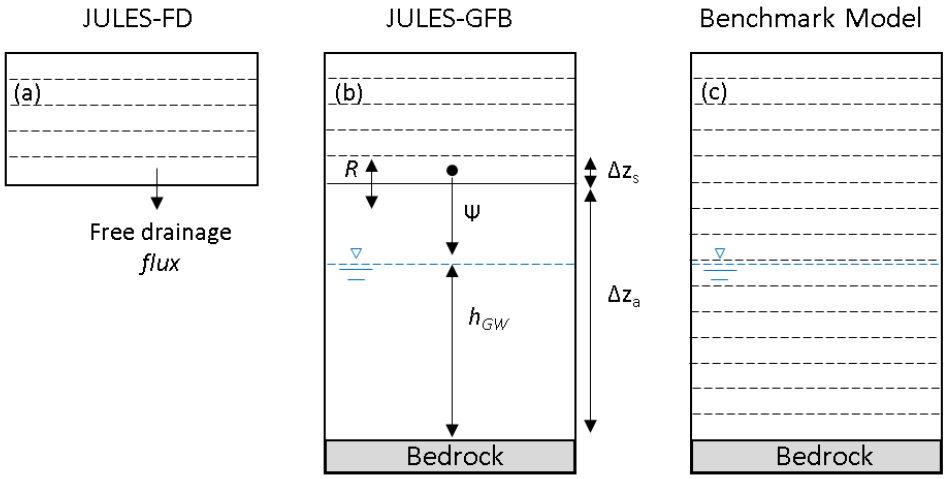


Figure 3.1: Schematic overview of JULES-FD with the free drainage assumption (a), JULES-GFB with the simplified groundwater representation (b) and a full 3-D hydrological model (shown here as a 2-D column for simplicity) (c).

The soil hydraulic parameters are calculated from soil texture information, using the van Genuchten (1980) hydraulic relationships. There is, also, the option of using the Brooks and Corey (1964) parameterization but in this study, it was chosen the first one, as it is scientifically more robust (Best et al., 2011) and it has been proven to increase the model performance (Dadson et al., 2011).

$$S_{\theta} = \frac{\theta - \theta_r}{\theta_s - \theta_r} = \frac{1}{[1+(\alpha_v \Psi)^n]^m} \quad (3.3)$$

where $\theta$ is the soil moisture ($m^3\ m^{-3}$), $\theta_r$ is the residual soil moisture ($m^3\ m^{-3}$),$\theta_s$ is the soil moisture at saturation ($m^3\ m^{-3}$), $S_\theta$ is the effective saturation, and $\alpha_v$, $n$ and $m$ are soil water retention curve fitting parameters computed based on the van Genuchten model. The hydraulic conductivity ($K_h$) is calculated from the following equation (Schaap et al., 2001).

$$K_h = K_{hS} S_\theta^{\xi} \left[1 - \left(1 - S_\theta^{\frac{1}{m}}\right)^m\right]^2 \quad (3.4)$$

where $K_{hs}$ is the hydraulic conductivity for saturated soil and $\xi$ is a parameter fixed at 0.5

JULES solves the subsurface heat and water transport equations based on finite difference approximation (Cox et al, 1999). To sum up, in JULES, every time step, the model calculates the fluxes and the derivatives d$W'$/d$\theta$ in the interfaces between all the layers. To do that, in each node, it uses the saturation from that node $\theta_1$ and the saturation from the node below $\theta_2$. Based on the Darcy equation, it estimates the difference of pressure heads between the two nodes and then the flux $W'$ and the derivatives d$W'$/d$\theta$. In the last layer, the model cannot use the soil moisture concentration from the node below. Thus, based just on the soil moisture concentration in the last node, it estimates the conductivity $K$, which is the flux for the last layer, using either the van-Genuchten (Equation 3.4) or the Brooks and Corey formula, and then it calculates the derivative d$K$/d$\theta$. Finally, the model creates a matrix based on all these fluxes and derivatives and solves this matrix using the Gaussian elimination method. The result is the increment of soil moisture concentration in each layer which is used to update the soil moisture concentration for the next timestep.

### 3.2.2 River Routing Scheme

The river routing scheme in JULES is based on the Rapid Flow Model (RFM) (Bell et al., 2007) and it estimates the approximation of the 1-D kinematic wave equation with lateral inflow. It uses 6 globally constant parameters, namely two surface and two subsurface wave celerities for river or land pixels, and two return flow fractions (Figure 3.2). Table 3-1 depicts the default values that stem from Bell et al. (2007).

Table 3-1: Parameter values of RFM Scheme (Bell et al., 2007).

| | Parameter Name | Description | Default Values |
|---|---|---|---|
| **RFM Parameters** | *cland* | Land wave speed | 0.2 (global), 0.4 (high-resolution) m $s^{-1}$ |
| | *criver* | River wave speed | 0.62 (global), 0.5 (high-resolution) m $s^{-1}$ |
| | *cbland* | Subsurface land wave speed | 0.1 (global), 0.05 (high-resolution) m $s^{-1}$ |
| | *cbriver* | Subsurface river wave speed | 0.15 (global), 0.05 (high-resolution) m $s^{-1}$ |
| | *retl* | Land return flow fraction | 0.0 |
| | *retr* | River return flow fraction | 0.005 |
| | *a_thresh* | The threshold drainage area | 1 (global), 10 (high-resolution) |

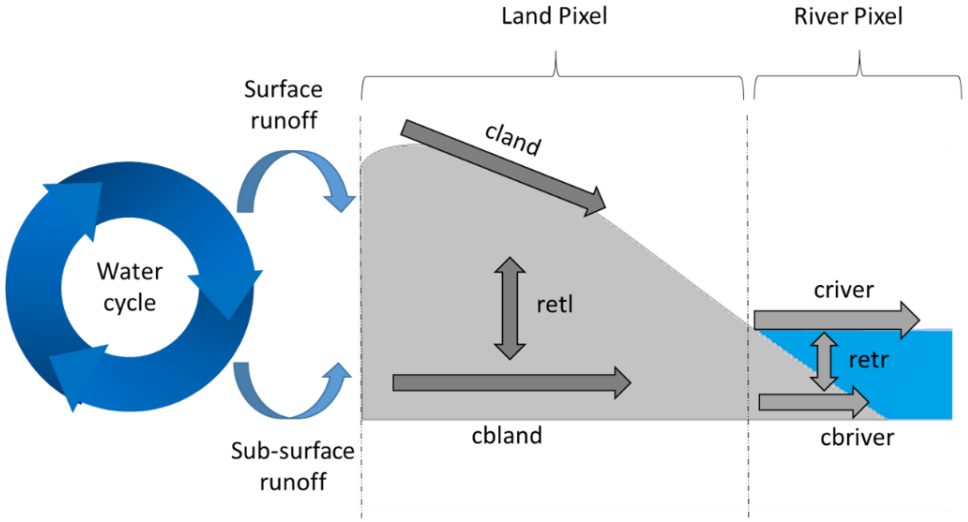


Figure 3.2: JULES River Flow Model comprises runoff production scheme, based on two runoff generation mechanisms, and flow routing parameterization where wave speeds are specified as universal constants.

### 3.2.3 New Groundwater Flow Boundary Parameterization

One of the key aspects of this study is to investigate the impact of explicitly adding a groundwater parameterization scheme into JULES (Batelis et al., 2020; this thesis). The results associated with this task is presented in Chapter 5. Here, we describe the methodology used to achieve this coupling. We adapt the concept of a free-surface groundwater flow boundary (GFB) condition, developed by Rahman et al. (2019) and targeted to large-scale hydrological modelling, to the JULES model (referred here as JULES-GFB). This is the first study that

applies this method in an LSM and tests the benefits and the limitations. The fundamental equation that governs the estimation of the groundwater flow in two horizontal dimensions can be written as (Pinder and Bredehoeft, 1968; Prickett and Lonnquist, 1971; Meenal and Eldho, 2011):

$$\frac{S\partial H}{\partial t} = \nabla(T\nabla H) + R \tag{3.5}$$

where $H$ is the total head [L], $T$ is transmissivity [$L^2\ T^{-1}$], $R$ is the recharge flux [$L\ T^{-1}$] and $S$ is the specific yield [-]. The coupling between the groundwater model and JULES-FD is based on recharge from the groundwater model. This is the flux that links the two models and interacts with them (flux $R$ in Figure 3.1b).

The GFB approach considers two major assumptions. The first assumption is the pressure and flux continuity at the interface between the lowest soil layer and the underlying aquifer. Following from this assumption, the saturated depth ($h_{GW}$) is calculated in JULES-GFB using the pressure head at the lowest soil layer as follows:

$$h_{GW} = \Delta z_\alpha + \Delta z_s/2 - \Psi \tag{3.6}$$

where $\Delta z_a$ is aquifer thickness [L], $\Delta z_s$ is the thickness of the last soil layer [L], and $\Psi$ is the pressure head [L] at the last soil layer (Figure 3.1b). Then, adding to $h_{GW}$ the datum from the bedrock, we can get the total head ($H$).

The second assumption is that the variability in the saturated depth ($\Delta h_{GW}$) is negligible compared to its absolute value of $h_{GW}$ (i.e., $\Delta h_{GW} << h_{GW}$). This assumption allows the calculation of transmissivity ($T$) at each model time step, as follows:

$$T = K_{hs} h_{GW} \tag{3.7}$$

Ultimately, Equation 3.5 takes the final format in order to calculate the recharge flux ($R$) based on the transient flow equation in an unconfined aquifer (Wang and Anderson, 1982):

$$R(x,y,t) = S\frac{\partial H}{\partial t} - \frac{\partial}{\partial x}\left(T_x \frac{\partial H}{\partial x}\right) - \frac{\partial}{\partial y}\left(T_y \frac{\partial H}{\partial y}\right) \tag{3.8}$$

Equations 3.5-3.8 show the formulation of the GFB described in Rahman et al. (2019), which serves as the basis for the representation of groundwater dynamics in JULES-GFB.

In addition to the approach described above, a few additional steps were necessary to incorporate the GFB concept into JULES. Firstly, we follow similar assumptions made by Niu et al. (2007) for the special case when the water table enters the soil domain in JULES. In this

case, we ignore the transient flow term ($S\frac{\partial H}{\partial t}$) and as a result, the recharge flux in the saturated portion of JULES soil domain is based only on the lateral flow (Equation 3.9), while the remaining unsaturated soil layers are updated as usual. The reason for this modification is to avoid numerical instabilities found in the original implementation of the model.

$$R = -\nabla(T\nabla H) \tag{3.9}$$

Secondly, a pseudo layer is introduced (acting as a "coupler" layer) between the soil domain and the aquifer, in order to calculate the derivative of flux with soil water content ($\frac{dW_n}{d\theta_n}$) that is needed for calculating the soil moisture increment every timestep. JULES does not calculate this derivative at the lowermost soil layer. Thus, a pseudo layer is needed to calculate the derivative in the last true layer of the soil domain, above the pseudo layer.

Our calculation of groundwater table depth is dependent entirely on the pressure head at the last layer of the JULES soil domain. Because of the way of our calculation, any abrupt change in the pressure head between two consecutive timesteps at the last layer of JULES could create oscillations in the groundwater table depth. Note that abrupt changes in the pressure head in the last layer are likely, especially due to the direct Gaussian solver of JULES. To reduce those numerical oscillations, a non-iterative Picard method (Paniconi et al., 1991; Tan et al., 2004) is introduced between timesteps (Equation 3.10).

$$R^{t+1,l+1} = \frac{R^{t+1,l} + R^{t}}{2} \tag{3.10}$$

where *t* is the timestep and *l* is the number of iterations. One iteration is sufficient to remove the oscillations and apply the new model explicitly. Note that if the pressure head is such that the water table falls below the bedrock for a given timestep (i.e., outside of the pre-determined aquifer domain), then the model reverts back to the free drainage parameterization (i.e., the effect of the water table on soil moisture dynamics becomes negligible). Furthermore, we should eliminate the subsurface flow from RFM, otherwise the new JULES-GFB model would estimate the subsurface flow twice, once in the groundwater model and once more in the river routing scheme. In summary, the potential advantages of this new scheme are the replacement of the often unrealistic free drainage assumption from JULES-FD; the lateral saturated flow interaction, as the calculation of the water table is based on the neighbouring cells; and the impacts on soil moisture content at the topsoil and the surface, where the water table intersects the surface and generates streams. The new scheme is extensively tested against a benchmark model, described in the next section, in two synthetic experiments in Chapter 5.

### 3.2.4 Benchmark Model

In this paragraph, the benchmark model, Parflow, is described briefly. For a more detailed description refer to other studies (e.g., Ashby and Falgout, 1996; Kollet and Maxwell, 2006; Maxwell, 2013). Parflow is used as a benchmark model in the synthetic experiments to validate the new model parameterization for groundwater. The reasons to choose Parflow as the benchmark model are that: (1) it has been extensively evaluated through model inter-comparison projects (Maxwell et al., 2014; Kollet et al., 2017) with satisfactorily results; (2) it has been used as benchmark in another model development study which tested the same groundwater theory implemented in JULES in this study (e.g., Rahman et al., 2019) and (3) it has been successfully tested in Continental US (CONUS) against observations (Maxwell et al., 2015) and against the National Water Model (WRF-HYDRO) (Tijerina, 2018). The results from the simulations of Parflow came from Kollet et al. (2017) and Rahman et al. (2019).

Parflow is a 3-Dimensional variably saturated groundwater surface water flow model that simulates the groundwater, vadose zone and the surface water as a single continuum based on the Richards and Saint Venant equations (Kollet et al., 2017). The model solves the Richards' equation in three spatial dimensions using the finite control volume approach with two-point flux approximation in space and implicit backward Euler scheme in time (Kollet and Maxwell, 2006; Rahman, 2015).

$$S_s\theta\frac{\partial\Psi}{\partial t}+\varphi\frac{\partial\theta(\Psi)}{\partial t}=\nabla\cdot R+Si \tag{3.11}$$

$$R=-K_{hs}K_r(\Psi)\nabla(\Psi-z) \tag{3.12}$$

where *Ss* is the specific storage coefficient [$L^{-1}$], $\varphi$ is porosity [-], *Si* is a general source/sink term [$T^{-1}$], $K_r$ is relative permeability [-] and $z$ is the depth below surface [L]. The kinematic wave equation is solved maintaining the continuity of pressure and flux at the boundary (Rahman, 2015). Parflow has been coupled to many LSMs and atmospheric models, such as CLM (Shrestha et al., 2014), WRF (Maxwell et al., 2011) and TerrSysMP (Kurtz et al., 2016).

It is worth noting that both Parflow and JULES-GFB solve the Richards' equation to estimate the transport of moisture. However, Parflow solves it in three spatial dimensions using the finite control volume approach from the surface up to the bedrock. On the other hand, JULES-GFB solves the Richards' equation in one spatial dimension using the finite difference approach from the surface up to 3 m depth (same depth with the soil domain) and solves the two-dimensional transient flow equation to estimate recharge in the aquifer. Parflow uses an

implicit backward Euler scheme to solve the differential equations, whereas JULES-GFB uses a forward explicit scheme. The implicit scheme of Parflow allows to have only five layers for 102 m total thickness of soil column (Maxwell et al., 2015), albeit in JULES-GFB a much finer vertical discretization is needed (i.e., 10 or 20 cm.) to run the model without numerical instabilities due to the explicit scheme. The runoff generation mechanisms are different. In Parflow, overland flow is calculated via an overland flow boundary condition assuming pressure and flux continuity at the surface-subsurface interface (Kollet and Maxwell, 2006). JULES-GFB uses the infiltration excess mechanism and the groundwater contribution to generate runoff.

## 3.3 Metrics

For the analyses, we use the following metrics to assess the model performance: First, we use the Pearson correlation (*r*) between observations and simulations (Equation 3.13)

$$r = \frac{\sum_{t=1}^{t=T}(obs(t) - \mu_{obs})\,(sim(t) - \mu_{sim})}{\sqrt{\sum_{t=1}^{t=T}(obs(t) - \mu_{obs})^2}\sqrt{\sum_{t=1}^{t=T}(sim(t) - \mu_{sim})^2}} \tag{3.13}$$

where *t* is the time, *obs* the observed timeseries, *sim* the simulated timeseries, $\mu_{sim}$ the simulation mean, and $\mu_{obs}$ the observation mean. Correlation ranges between -1 and 1 with zero value to suggest no correlation. According to Schober et al. (2018) values of correlation above 0.7 should be considered as strong, between 0.4 and 0.7 as moderate and between 0.1 and 0.4 as weak. Note that Pearson correlation calculates linear correlation between observations and simulations. In case that there is non-linear correlation, then Pearson will either not identify it or will underestimate the significance of it.

The Kling and Gupta Efficiency (KGE) (Gupta et al., 2009) (Equation 3.14) and Nash and Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970) (Equation 3.15) are two common metrics in hydrology. NSE and KGE = 1 indicate perfect agreement between simulations and observations.

$$\mathrm{KGE} = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_{sim}}{\sigma_{obs}} - 1\right)^2 + \left(\frac{\mu_{sim}}{\mu_{obs}} - 1\right)^2} \tag{3.14}$$

where *r* is the linear correlation, $\sigma_{obs}$ is the standard deviation in observations, $\sigma_{sim}$ the standard deviation in simulations. Looking in the literature about the acceptable range of values for KGE, we find different limits. Okello et al. (2018) considered values of KGE less 0.5 as

poor, between 0.5 and 0.7 acceptable and above 0.7 as good. Poméon et al. (2018) and Rajib et al. (2016) considered as acceptable models with KGE above 0.5. Gutenson et al. (2019) classified KGE as acceptable for KGE between 0 and 0.4, as good for KGE between 0.4 and 0.7 and very good for KGE values above 0.7, following Tavakoly et al. (2017).

$$\text{NSE} = 1 - \frac{\sum_{t=1}^{t=T}(sim(t) - obs(t))^2}{\sum_{t=1}^{t=T}(obs(t) - \mu_{obs})^2} \tag{3.15}$$

In term of NSE, negative values suggest that the model has less explanatory power from the mean of the observations (e.g. Schaefli and Gupta, 2007). Values of NSE equal to zero are used as benchmark to distinguish "good" and "bad" models (e.g., Moriasi et al., 2007; Schaefli and Gupta, 2007). Moriasi et al. (2007) considered NSE values above 0.5 as satisfactory. Taking into consideration the above classifications, we define KGE and NSE values below 0.5 as poor and above 0.5 as acceptable. Negative values are classified as "bad" performance.

Finally, mean bias was used to evaluate the new groundwater parameterization in the synthetic experiments (Equation 3.16). The reason why we do not use absolute values is to utilize the sign that suggests if the new parameterization overestimates or underestimate a particular variable.

$$\text{Mean Bias} = \frac{1}{T}\sum_{t=1}^{t=T}(sim(t) - obs(t)) \tag{3.16}$$

# 4 EVALUATING JULES IN GREAT BRITAIN

## 4.1 Introduction

Recently, LSMs have been applied for operational applications in global hydrology (e.g., Beck et al., 2017; Sutanudjaja et al., 2018; Givati et al., 2016). However, the performance of the models is not acceptable in all cases under different climate and hydrological conditions (e.g., Beck et al., 2017; Zeng et al., 2018). Furthermore, even though LSMs are used for hydrological purposes, the hydrological processes in these models need improvement, as the river routing and most importantly the groundwater are usually represented in a crude way or are neglected (Clark et al., 2015). Thus, evaluating an LSM under multiple catchments can give insight about the missing components and lead to model improvements.

The hydrological components of the JULES model have been tested for runoff predictions, but usually in a few catchments apart from Martinez-de la Torre et al. (2019). So far there is not an extensive study of JULES model in multiple catchments to evaluate the model and get insight on the missing hydrological processes in the model. Most of the previous studies used JULES to simulate hydrological processes in large time-steps (monthly, annual) (e.g., Gudmundsson et al., 2012a, b) or in a few catchments (Dadson and Bell, 2010; Dadson et al., 2011). Therefore, the aim of this chapter is to investigate the performance of the JULES model to simulate runoff at 47 catchments in Great Britain in a finer time step (daily). We want to understand the strengths and the limitations of the model to simulate streamflow, while we want to find the key hydrological processes that are missing in the current version of the model. Our two specific research questions in this chapter are:

i) How does JULES' performance to simulate runoff contrast in surface water and groundwater dominated catchments?

ii) What are the dominant factors (land use, geology, soil texture) that affect the performance of the model to simulate runoff? Are there any missing hydrological processes in JULES?

## 4.2 Evaluation methods

For the initial evaluation step, we use the default version of JULES (v4.4) referred here as JULES Default. We use the default parameterization and vertical discretization (i.e., 0.1, 0.25, 0.65 and 2.0 m.) of the model. Furthermore, we use the results of the performance metrics (Nash and Sutcliffe Efficiency) from two other studies (Bell et al., 2007; Martinez-de la Torre et al., 2019) to extend our analysis. Bell et al. (2007) applied the model Grid-to-Grid to Great Britain. Grid-to-Grid is a catchment-based rainfall-runoff model and it is used for operational purposes. JULES hydrological parameterization stem from Grid-to-Grid. It uses the same river routing model with Grid-to-Grid, and it relies on the PDM scheme, in a similar way with Grid-to-Grid. Bell et al. (2007) tested Grid-to-Grid model in 25 catchments in UK for a nine-year period. Although Grid-to-Grid has differences with JULES, we decided to include the results of the performance metrics in our analysis.

We select 47 catchments from the National River Flow Archive (NRFA) (Figure 4.1) that provides with daily streamflow observations. The selection criteria were to be larger than 50 $km^2$ with no artificial influences, like dams and abstractions. The selected catchments have different size, from 53 to 1480 $km^2$, and different land use and soil types (Table 4-1). Looking the histogram of the catchment size (Figure 4.1), we find 30 catchments with area less than 500 $km^2$ and 7 catchments with area more than 1000 $km^2$. The histogram is not uniform but skewed with many small catchments and less larger catchments. This is an inevitable weakness of this analysis, because the results will reflect mostly the behaviour of the smaller catchments. Note that usually catchments with size between 100 and 10,000 $km^2$ are considered as medium-large catchments (Solheim et al. 2019). However, the terms “small” or “big” catchments here reflect the current samples and the conditions in UK. Another limitation is the fact that our evaluation is limited to British catchments, knowing that JULES is a British model, which are very particular (relatively small compared to other basins worldwide and having narrow hydroclimatological characteristics). We select catchments without artificial influences, like dams and abstractions, because we want to assess only the impact of the model without adding more processes, like water management, that are not represented in the model. Most of these

catchments are relatively small, as most of the bigger catchments in UK (not selected here) are used for water management purposes.

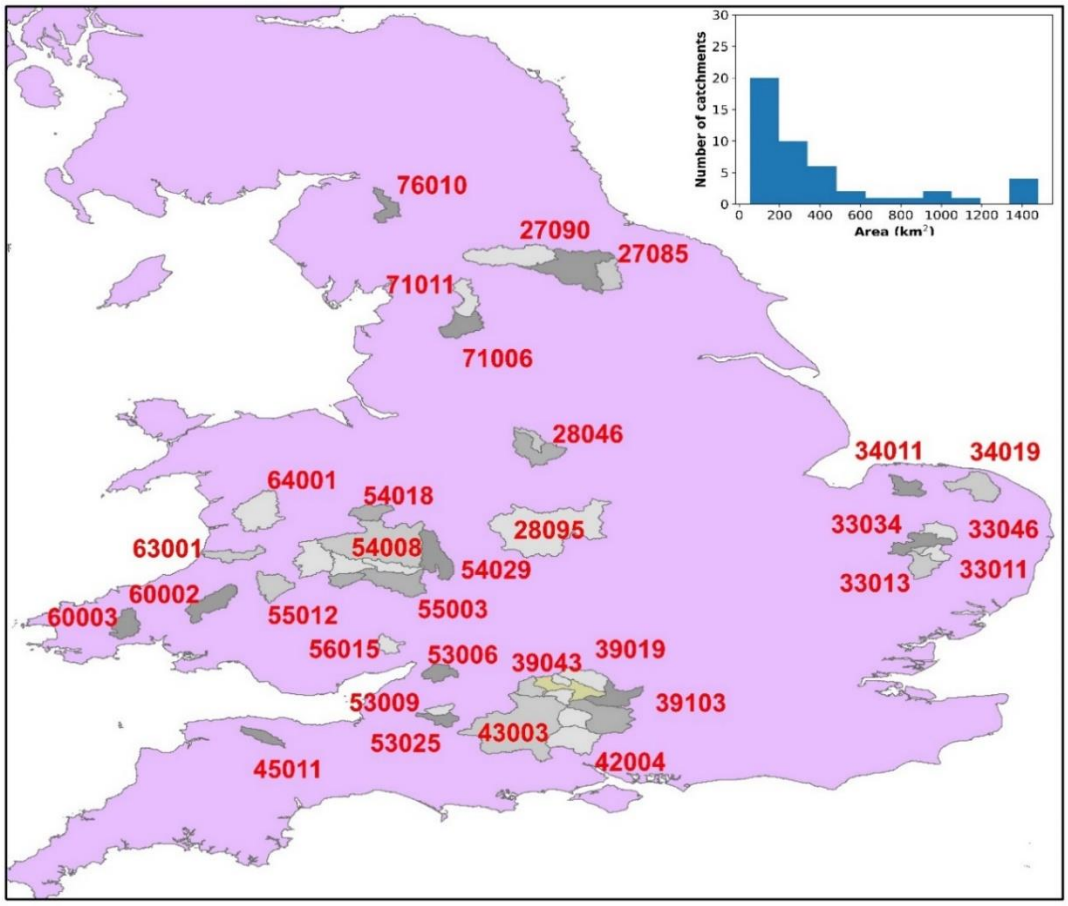


Figure 4.1: Map of 47 catchments used in this analysis. For clarity purposes, we applied different grey scale to distinguish the catchments. The inset plot shows the histogram of the catchment size.

Regarding the spatial data, in the East of England, the dominant land use type is Shrubs, whereas in Wales and in the North of England the dominant land use type is C3 grass (Figure 4.2a). Clay loam is the dominant soil texture in the Northern and West part of England and in the West part of Wales (Figure 4.2b). Rock is the dominant soil texture in South England and clay in the East part of England. Saturated conductivity is a key variable for the groundwater processes, as it dictates the rate of the drying of the soil domain. Figure 4.2c presents the spatial data of saturated conductivity with similar values for the biggest part of England and Wales (50-100 cm $d^{-1}$), except a few small parts in North and West England with higher values (150-200 cm $d^{-1}$).

We decided to use open data from the British Geological Survey to acquire information related to the geology (bgs.ac.uk/datasets/bgs-geology-625k-digmapgb/ last visited on

18/11/2020). Land Information System provides information for depth up to 2 meters. In our simulation, we define all the hydrogeological parameters for the soil domain (3 m) using only this dataset. However, we think that using geological data will help us firstly to define the area of chalk and secondly to evaluate the impact of geology in the model performance, even if we haven't used this information in the model parameterization.

In the South East part of England, the dominant type of geology is the chalk with the macroporosity behaviour and very high values of conductivity up to 500 cm $day^{-1}$ (Lewis et al., 2006) (Figure 4.2d). Chalk is a fine-grained porous medium traversed by fractures (Price et al, 1993) and has dual porosity medium with flow occurring from both the matrix and fractures (Mathias et al., 2005). Our goal is to develop a parameterization for the groundwater for JULES (JULES-GFB) that will be applicable in different locations, even globally. Chalk is a different case with complex hydrology (Rahman and Rosolem, 2017), so it is out of the scope of this study.

In Wales and in Central England, the dominant types of geology are Mudstone, Siltstone and Sandstone (Figure 4.2d). In the north part of England, the dominant type is Limestone. In general, Limestone has the highest saturated conductivity values (0.01 to 100 cm $day^{-1}$) (McCall and Marker, 1989) compared to the Sandstone (0.001 to 10 cm $day^{-1}$), Siltstone ($10^{-4}$ to 0.1 cm $day^{-1}$) and Mudstone geology type ($10^{-5}$ to 0.001 cm $day^{-1}$) (Freeze and Chery, 1979). Interestingly, looking the map of saturated conductivity from the Land Information System (Figure 4.2c) and the geology from BGS (Figure 4.2d), we find that there are no different saturated hydraulic conductivity values for the chalk region, against of what the literature suggests. Hence, the use of BGS data can enlighten the reason that JULES performs well or bad in a few areas, even if we don't use them in the soil parameterization, as the Land Information System cannot provide us with information about the aquifer and the groundwater processes.

Table 4-1: Information about the selected catchments in UK with the model performance for JULES v4.4. KGE and NSE values are calculated against daily streamflow observations from NRFA. Note that negative KGE and NSE values are shown qualitatively as "< 0".

| **Name** | **Station** | **BFI*** | **Area ($km^2$)** | **Geology** | **Land Use** | **KGE** | **NSE** |
|---|---|---|---|---|---|---|---|
| Swale at Crakehill | 27071 | 0.48 | 1363 | Mudstone, Sandstone | Shrubs | 0.34 | 0.40 |
| Cod Beck at Dalton Bridge | 27085 | 0.48 | 209 | Mudstone, Sandstone | Shrubs | 0.47 | < 0 |
| Swale at Catterick Bridge | 27090 | 0.37 | 499 | Limestone | Shrubs | 0.15 | 0.24 |
| Dove at Rocester Weir | 28008 | 0.62 | 399 | Limestone | C3 Grass | 0.59 | 0.17 |

Table 4-1: (continued).

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Dove at Izaak Walton | 28046 | 0.79 | 83 | Mudstone, Sandstone | C3 Grass | 0.04 | < 0 |
| Tame at Hopwas Bridge | 28095 | 0.7 | 1422 | Mudstone, Sandstone | C3 Grass | 0.36 | < 0 |
| Little Ouse at County Bridge Euston | 33011 | 0.7 | 129 | Chalk | Shrubs | < 0 | < 0 |
| Sapiston at Rectory Bridge | 33013 | 0.65 | 206 | Chalk | Shrubs | < 0 | < 0 |
| Little Ouse at Abbey Heath | 33034 | 0.8 | 689 | Chalk | Shrubs | < 0 | < 0 |
| Thet at Redbridge | 33046 | 0.6 | 145 | Chalk | Shrubs | 0.17 | < 0 |
| Little Ouse at Knettishall | 33063 | 0.64 | 101 | Chalk | Shrubs | < 0 | < 0 |
| Wensum at Fakenham | 34011 | 0.82 | 162 | Chalk | Shrubs | < 0 | < 0 |
| Bure at Horstead Mill | 34019 | 0.81 | 313 | Chalk | Shrubs | < 0 | < 0 |
| Kennet at Theale | 39016 | 0.88 | 1033 | Chalk | Shrubs | < 0 | < 0 |
| Lambourn at Shaw | 39019 | 0.97 | 234 | Chalk | Shrubs | < 0 | < 0 |
| Enborne at Brimpton | 39025 | 0.53 | 147 | Chalk | Shrubs | < 0 | < 0 |
| Dun at Hungerford | 39028 | 0.95 | 101 | Chalk | Shrubs | < 0 | < 0 |
| Lambourn at Welford | 39031 | 0.98 | 176 | Chalk | Shrubs | < 0 | < 0 |
| Kennet at Marlborough | 39037 | 0.94 | 142 | Chalk | C3 Grass | < 0 | < 0 |
| Kennet at Knighton | 39043 | 0.95 | 295 | Chalk | Shrubs | < 0 | < 0 |
| Aldbourne at Ramsbury | 39101 | 0.96 | 53 | Chalk | C3 Grass | < 0 | < 0 |
| Kennet at Newbury | 39103 | 0.93 | 548 | Chalk | Shrubs | < 0 | < 0 |
| Test at Broadlands | 42004 | 0.94 | 1040 | Chalk | Shrubs | < 0 | < 0 |
| Anton at Fullerton | 42012 | 0.96 | 185 | Chalk | Urban | < 0 | < 0 |
| Test at Chilbolton Total | 42024 | 0.97 | 453 | Chalk | C3 Grass | < 0 | < 0 |
| Avon at East Mills Total | 43003 | 0.91 | 1478 | Chalk | C3 Grass | < 0 | < 0 |
| Barle at Brushford | 45011 | 0.52 | 128 | Sandstone, Conglom. | C3 Grass | 0.44 | 0.16 |
| Frome at Frenchay | 53006 | 0.38 | 149 | Mudstone, Sandstone | Urban | 0.54 | 0.23 |
| Wellow Brook at Wellow | 53009 | 0.62 | 73 | Mudstone, Sandstone | Shrubs | 0.21 | < 0 |
| Mells at Vallis | 53025 | 0.59 | 119 | Limestone | Shrubs | 0.2 | < 0 |
| Teme at Tenbury | 54008 | 0.56 | 1134 | Mudstone, Sandstone | Shrubs | 0.66 | 0.46 |
| Rea Brook at Hookagate | 54018 | 0.51 | 178 | Mudstone, Sandstone | C3 Grass | 0.64 | 0.24 |
| Teme at Knightsford Bridge | 54029 | 0.55 | 1480 | Mudstone, Sandstone | Shrubs | 0.62 | 0.47 |
| Lugg at Lugwardine | 55003 | 0.64 | 886 | Mudstone, Sandstone | Shrubs | 0.54 | 0.01 |
| Irfon at Cilmery | 55012 | 0.37 | 244 | Mudstone, Sandstone | C3 Grass | 0.62 | 0.55 |
| Arrow at Titley Mill | 55013 | 0.56 | 126 | Mudstone, Sandstone | Shrubs | 0.54 | 0.16 |
| Lugg at Byton | 55014 | 0.65 | 203 | Mudstone, Sandstone | C3 Grass | 0.57 | 0.11 |
| Ithon at Disserth | 55016 | 0.39 | 358 | Sandstone, Conglom. | C3 Grass | 0.45 | 0.48 |
| Lugg at Butts Bridge | 55021 | 0.65 | 371 | Mudstone, Sandstone | Shrubs | 0.61 | 0.40 |
| Olway Brook at Olway Inn | 56015 | 0.43 | 105 | Mudstone, Sandstone | C3 Grass | 0.42 | 0.32 |
| Cothi at Felin Mynachdy | 60002 | 0.43 | 298 | Sandstone, Conglom. | C3 Grass | 0.48 | 0.49 |
| Taf at Clog-y-Fran | 60003 | 0.55 | 217 | Mudstone, Sandstone | C3 Grass | 0.57 | 0.31 |
| Ystwyth at Pont Llolwyn | 63001 | 0.4 | 170 | Mudstone, Sandstone | Needleaf Forest | 0.57 | 0.53 |
| Dyfi at Dyfi Bridge | 64001 | 0.39 | 471 | Mudstone, Sandstone | C3 Grass | 0.36 | 0.43 |
| Ribble at Henthorn | 71006 | 0.31 | 456 | Mudstone, Sandstone | C3 Grass | 0.54 | 0.48 |
| Ribble at Arnford | 71011 | 0.25 | 204 | Sandstone, Conglom. | C3 Grass | 0.54 | 0.54 |
| Petteril at Harraby Green | 76010 | 0.46 | 160 | Mudstone, Sandstone | C3 Grass | 0.73 | 0.50 |

* BFI was computed by NRFA based on the archived record of gauged daily mean runoff.

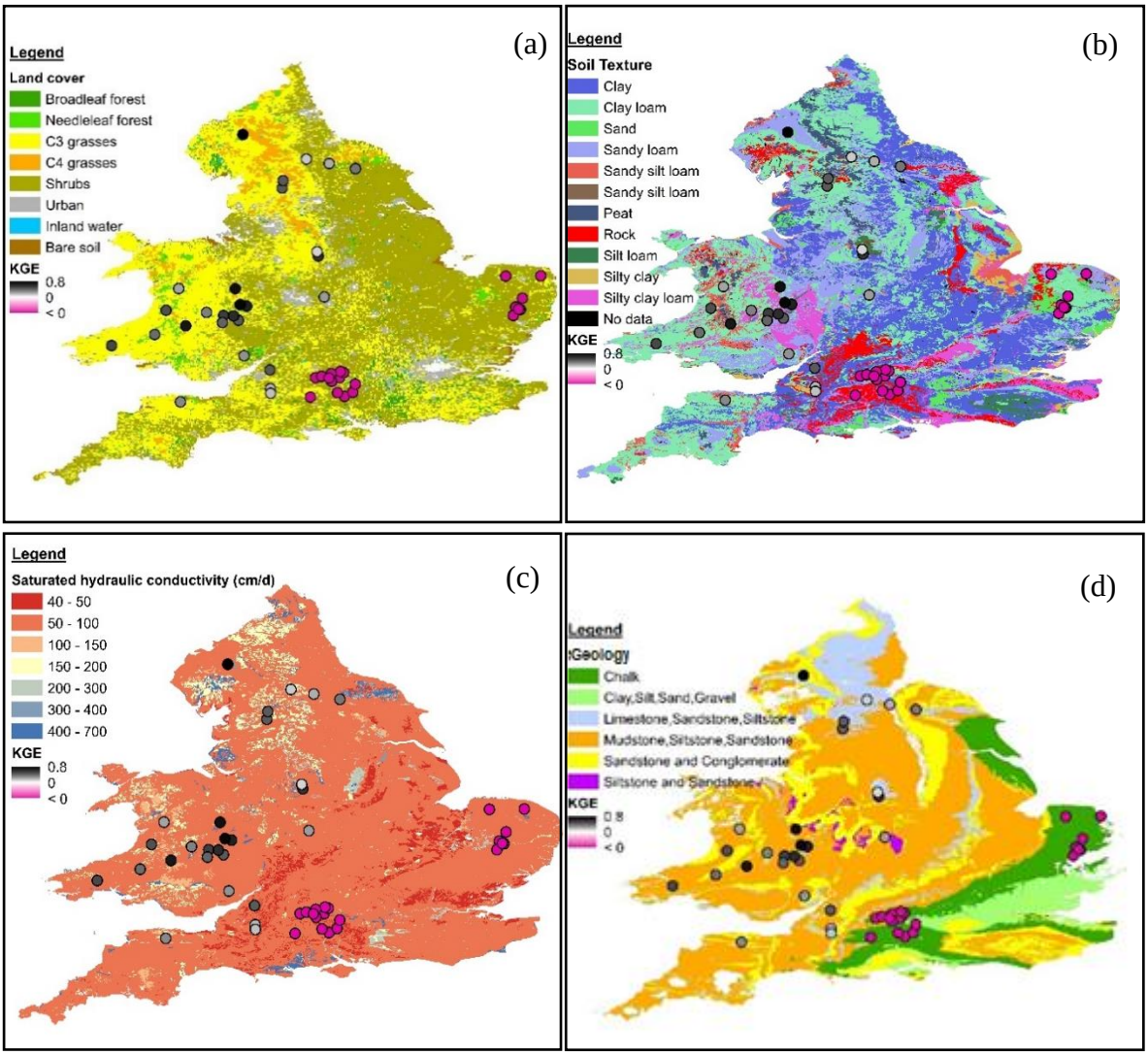


Figure 4.2: Physical characteristics of the experiment: (a) land cover (Land Cover Map 2007); (b) soil texture (Land Information System); (c) saturated hydraulic conductivity (Land Information System) and (d) geology (British Geological Survey), overlaid by model performance (KGE) of JULES Default for 47 catchments.

The atmospheric forcing data are obtained from the Climate, Hydrology and Ecology research Support System (CHESS) dataset (Robinson et al., 2016) that is available at daily temporal resolution and 1 km spatial resolution covering the entire Great Britain. The model set-up has spatial resolution equal to 1 km and temporal resolution equal to 30 minutes. The forcing is disaggregated from daily to 30 min timestep using a technique described in Williams and Clark (2014). In particular, the daily forcing is disaggregated to model timesteps imposing a diurnal cycle on temperature and radiation while allocating a constant value of precipitation rate for the entire day. The wind and the pressure are unchanged by the disaggregation scheme

(Williams and Clark, 2014). The simulation time consists of a ten-year period [2006-2015]. This period includes both drought [e.g., 2009-2010] and flood periods [e.g., 2013-2014]. In this case, the model will be assessed in both dry and wet conditions. This can be seen in Figure 4.3, in which the forcing for three different catchments, that represent Central, North and East England, are shown. The temperature shows almost identical behaviour and similar patterns in the three different areas compared to precipitation which has higher deviation from area to area. There is no significant change of the monthly temperature for the 10-year period. For the three catchments, the years of 2009 and 2010 showed the lowest minimum temperature in the winter with negative values during December. In the following years, there is increase of the winter temperatures, while this is not the case for the mean monthly temperatures during the summer. Regarding the precipitation, we find differences, as we expected, because the distance between the domains is more than 400 km. The monthly rainfall of the Irfon at Cilmery (55012) catchment has higher variability from the other two catchments. However, looking the Figure 4.3, we can identify the extreme events of flood in 2009 and of extreme drought in 2010-11 for all the catchments.

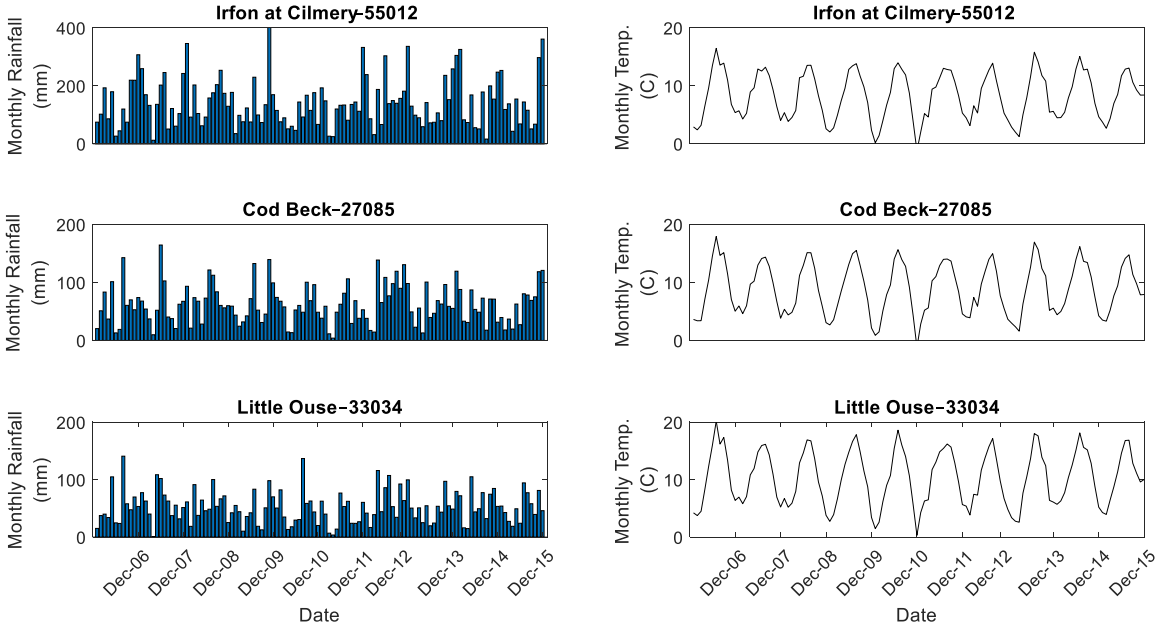


Figure 4.3: Monthly rainfall (left) and temperature (right) for three different catchments that represent Central (top), North (medium) and East England (bottom).

We run the ten-year period for simulation [2006-2015] twice, as we disregard the first cycle as the spin-up period. The model is evaluated against daily runoff data from NRFA using the

Kling and Gupta Efficiency (KGE) (Gupta et al., 2009) and Nash and Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970) as evaluation metrics. Finally, in our analysis we use the Base Flow Index (BFI) that represents the proportion of baseflow runoff to the total annual streamflow. Values of BFI lower than 0.5 usually represent flashy catchments with low contribution of groundwater. Values higher than 0.5 represent groundwater dominated catchments. High values close to one may represent chalk catchments dominated by macroporosity and remarkably high saturated hydraulic conductivity values.

Note that the benefit of an LSM, like JULES, is that it should rely mostly on data and on explicit representation of processes, and not on calibration. We assume that the physics of the model is close to reality. Thus, if the forcing data (soil data, land use, forcing) have high quality, then we expect that the physics should be represented as expected in a region. In our case, we hypothesize that our spatial data have high quality and JULES simulates the hydrological processes accurately, so we expect good agreement between the observed and simulated runoff. In case that in some regions, the model fails to simulate the processes correctly, then it seems that there are a few missing parameterizations that the model needs to include in order to behave properly.

Following this, we do not apply any calibration to any parameter during the analysis. We use the default PDM scheme to incorporate the saturation excess mechanism, as is a common approach within JULES community (e.g., MacKellar et al., 2013; Martinez-de la Torre et al., 2019). We use the default values for the PDM scheme (i.e, $b_{PDM}$ (1) and the depth (1 m)).

## 4.3 Results

Comparing the daily runoff simulated by JULES Default against observations from NRFA, we distinguish three different patterns, namely for the flashy (BFI<0.5), groundwater dominated (BFI>0.5) and chalk catchments (overlay chalk). The first group (flashy catchments) comprises of 13 catchments, the second group (groundwater dominated) 14 and the last one (chalk) 20 catchments. Figure 4.4 depicts on the left the runoff timeseries for JULES Default against observations for three different catchments that represent the three distinct behaviour we find during this analysis for a subset of the simulation period. Figure 4.4 on the right depicts the scatter plot of observed daily streamflow against the simulated streamflow by JULES Default. In these scatter plots, we distinguish the three aforementioned behaviours. For one case

JULES Default underestimates runoff and for two cases, both of them groundwater dominated, JULES overestimates the runoff component.

The first catchment (Irfon at Cilmery - 55012) is flashy (i.e., BFI = 0.37) with very low runoff values during the dry season, which ultimately means low baseflow contribution. In this catchment, JULES Default underestimates runoff, especially high flows (Figure 4.4 top right). This is a common issue in the model (e.g., Clark and Gedney, 2008; Zulkafli et al., 2013; Martinez-de la Torre, 2019). In particular, JULES reduces the high flows up to 60%. From the timeseries, we can see that JULES Default captures the timing of the floods, but not always the magnitude, while it overestimates a few medium flood peaks. The model shows a flashy behaviour and fails to simulate the recession curve after a flood event. During the dry season, although the observed runoff is close to zero, in JULES Default there is a constant flow coming from the bottom of the soil domain (free drainage) that contributes to the total runoff. This means that the model may underestimate the impact of a drought on the streamflow. The performance metrics of KGE and NSE are 0.62 and 0.55, respectively. These values are generally considered as acceptable (e.g., Moriasi et al., 2007; Poméon et al., 2018).

Tame, on the other hand, (28095) is a groundwater dominated catchments with BFI=0.7. In this case, the linear fit is close to the 1:1 line with $R^2$ equal to 0.59 and slope close to unity (1.15), meaning a good fit (Figure 4.4 middle right). However, carefully looking the scatterplot we find good agreement for high flows, that influence the linear fit mostly, since the samples overlay on the 1:1 line. On the other hand, there is poor agreement for simulating medium flows, as JULES Default overestimates them up to four times. From the observed daily runoff timeseries (Figure 4.4 middle left), we see a constant baseflow around 30 $m^3$ $s^{-1}$, many low floods, but only a few significant flood events. JULES Default fails completely to capture the baseflow dynamics and it is generally more responsive to the rainfall events, while it overestimates the peaks from 50 to 200%. The low performance metrics (KGE=0.36 and NSE=-0.1) show the poor performance of the model to simulate runoff in this catchment.

Finally, the last catchment (Test at Broadlands - 42004) represents a chalk landscape with very high BFI (0.94). Looking the daily timeseries, it is obvious that the dominant component of the streamflow is the baseflow, whereas the flood peaks during flood events are very low compared to the baseflow. JULES Default fails completely to capture the macroporosity behaviour of the chalk catchment. The scatter plot visualizes better this failure, with JULES Default to overestimates the daily flows up to six times. The range of daily simulated flows is from 0 to 150 $m^3$ $s^{-1}$, whereas the range of the observed runoff is from 0 to 30 $m^3$ $s^{-1}$. Both

metrics have negative values (i.e., KGE=-1.36, NSE=-8.3), meaning that the model performs worse than assuming a mean value across the entire time-series. Chalk is a different case with complex hydrology (Rahman and Rosolem, 2017), so it is out of the scope of this study.

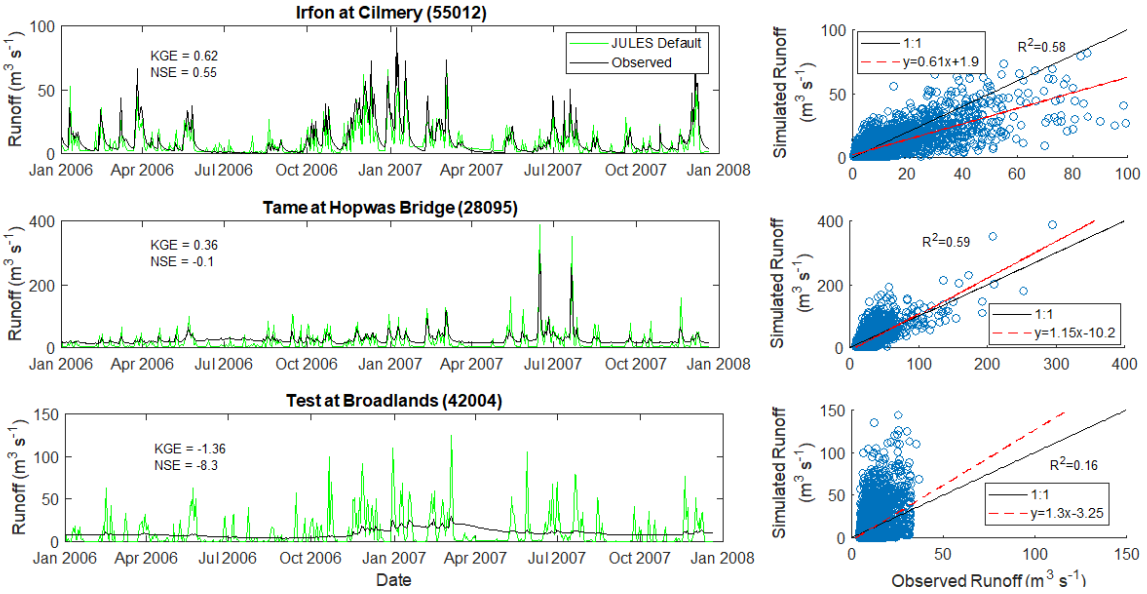


Figure 4.4: Daily runoff timeseries (left) for JULES Default (green) against observations (black) and scatter plot (right) for three different catchments that represent flashy (top), groundwater dominated (middle) and chalk catchment (bottom).

In the next step, we examine if the performance of the model is similar for the dry (June-August) and wet season (December-February) for each cluster of catchments (Figure 4.5). We find two distinct patterns for the overland and groundwater dominated catchments. In the flashy catchments, we find a better performance of the model during the summer season. In these catchments, the runoff comes only from the rainfall, since the baseflow is insignificant. During the dry season, the streamflow is usually close to zero with rare flood events. JULES constantly overestimates the runoff due to the contribution of the groundwater through the free drainage flux, but the difference is low, and it does not affect the components of KGE. During the winter, there are significant flood events that the model fails to capture the magnitude. The bias during these events affects the value of KGE and the total performance during this season. Consequently, the model has superior performance during the summer. We find the opposite behaviour in the groundwater dominated catchments, that is more profound in the chalk cluster. The model has better performance during the winter. The reason is that during the summer, the runoff is heavily dominated by the baseflow, while the peaks are relatively low compared to the baseflow. JULES model does not have a groundwater parameterization, so the baseflow is

very low. This increases the bias and decreases the KGE. During the winter, there are a few significant flood events that the model achieves to capture. The fact that the influence of baseflow during the winter decreases, as there are more flood events, leads to better performance of the model.

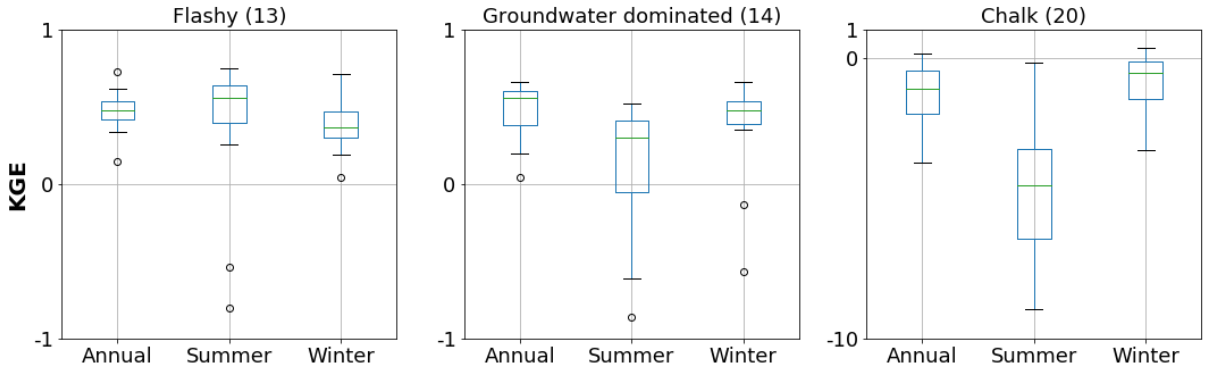


Figure 4.5: KGE values of daily runoff against observations for the three different clusters of catchments in annual, summer and winter season. Note vertical scale for the Chalk subplot is different from the other two subplots.

Table 4-1 shows the values of KGE and NSE for the 10-year period for all the catchments. Furthermore, the dominant geology, the soil texture and the land use type are included to identify the factors that control the model performance. We find three different land cover types, seven different soil textures and four different types of geological formations. Figure 4.2 includes four maps of the physical characteristics of England and Wales. The circles depict the KGE performance of the model from high (0.8-black) to low (0-white). The purple circles depict catchments with KGE lower than zero and represent chalk catchments. Considering the geological map, we find catchments with negative KGE performance to overlay chalk regions at South and South East of England. It is something we expect by analysing the daily runoff timeseries (Figure 4.4). Catchments with higher KGE (above 0.5) overlay the Mudstone, Siltstone and Sandstone, whereas catchments with KGE between 0 and 0.5 overlay the Limestone. Here, there is a strong indication that the performance of the model is linked to the geology of the catchments. Considering the importance of saturated conductivity in the processes within the soil domain and especially the free drainage flux, we include the map of it provided by the Land Information System. Catchments with KGE less than 0.5 have higher saturated conductivity (about 150-200 cm $d^{-1}$), whereas catchments with better performance,

have lower saturated conductivity (about 100-150 cm $d^{-1}$). This means that catchments with lower saturated conductivity perform better. The map of soil texture is patchy and is difficult to draw conclusions from it. Catchments with higher KGE (above 0.5) overlay loam, silty clay loam, clay, or sandy loam. However, clay loam and clay are the most common soil textures in catchments with KGE between 0 and 0.5, as well. It seems that the soil texture does not dominate the model performance. Finally, in the land use map, we find that catchments with poor performance overlay with Shrubs that is the dominant land use in East, South and Central England. Overall, we find that lower performance of the model is associated with Chalk and Limestone areas that have higher saturated conductivity values and the land use type is Shrubs. There is not a unique soil texture that leads to better model performance. However, from all these factors that may or may not affect the model performance, it should be found which is the dominant one. Having found the dominant factor, we can link it with a missing process in the model parameterization that affects the model to perform satisfactorily.

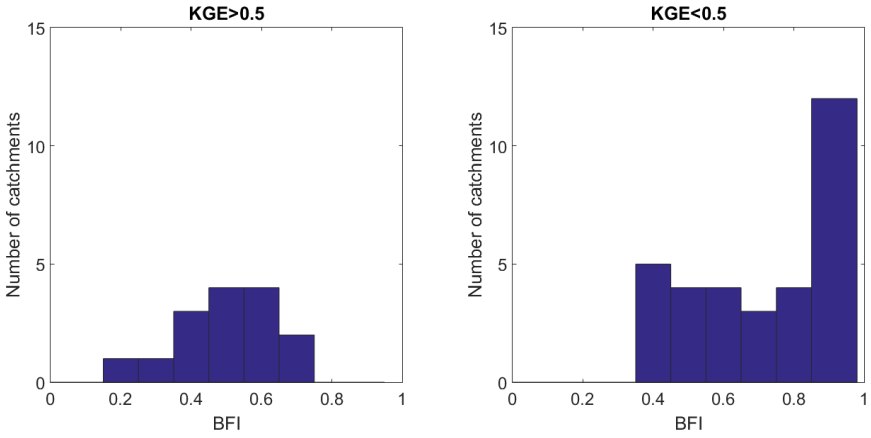


Figure 4.6: Histogram for BFI of catchments with KGE higher than 0.5 (left) and lower than 0.5 (right).

First, we analyse if the geology is the dominant factor of the model performance by using the Base Flow Index (BFI). BFI is widely used in combination with the Hydrology of Soil Types (HOST) classification to develop the BFIHOST catchment descriptor in UK (Griffin et al., 2019). The HOST classification uses groups of soil classes based on physical and hydrogeological properties of soil. We plot a histogram (Figure 4.6) that compares the model performance and BFI indexes. We plot separately the catchments with high values of KGE (KGE > 0.5) (Figure 4.6 left) and low KGE values (KGE < 0.5) (Figure 4.6 right) following the performance classification from Rajib et al. (2016) and Poméon et al. (2018). We can see that

all the catchments with BFI less than 0.4 give KGE values above 0.5. Furthermore, most of the groundwater-dominated catchments with high values of BFI produce low KGE performance. We can see the strength of the model to simulate more accurately flashy catchments (i.e., low BFI) compared to groundwater-dominated ones.

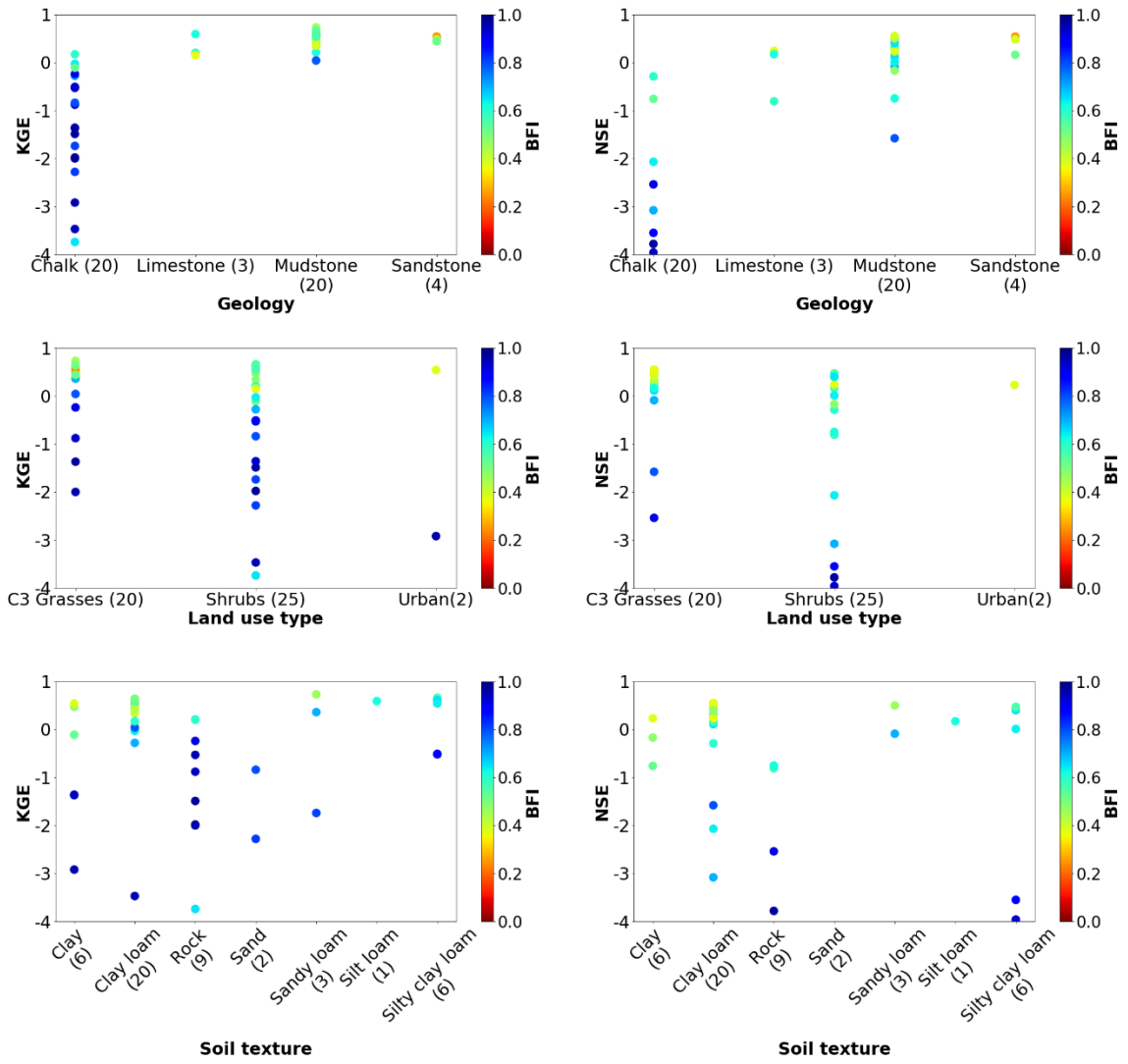


Figure 4.7: Scatter plot of Geology (upper), land cover type (middle) and soil texture (bottom) against KGE (left) and NSE (right). Base Flow Index (BFI) is used as colorbar. The number in the parenthesis shows the number of catchments per category.

In order to extend this analysis, we plot the model performance (i.e., KGE and NSE) against catchment characteristics (geology, soil texture and land cover type) (Figure 4.7). We use the BFI values to colour the dots, blue for high BFI values and red for low BFI values. The two metrics show similar behaviour. Interestingly, all the types of geology, soil textures and land

use types give higher performance for low BFI values. We do not find a significant improvement of the model for a specific type of geology or land use type. We find better performance for a few soil textures (e.g., clay, clay loam) that usually have lower saturated conductivity, but again as the BFI increases, the performance deteriorates. Overall, we can deduce that it is the BFI or in other words the contribution of groundwater to the total regime that dominates the model performance. This finding proves our initial hypothesis that the missing groundwater represention on JULES affects the model performance and especially catchments in which groundwater has a significant proportion to the total runoff.

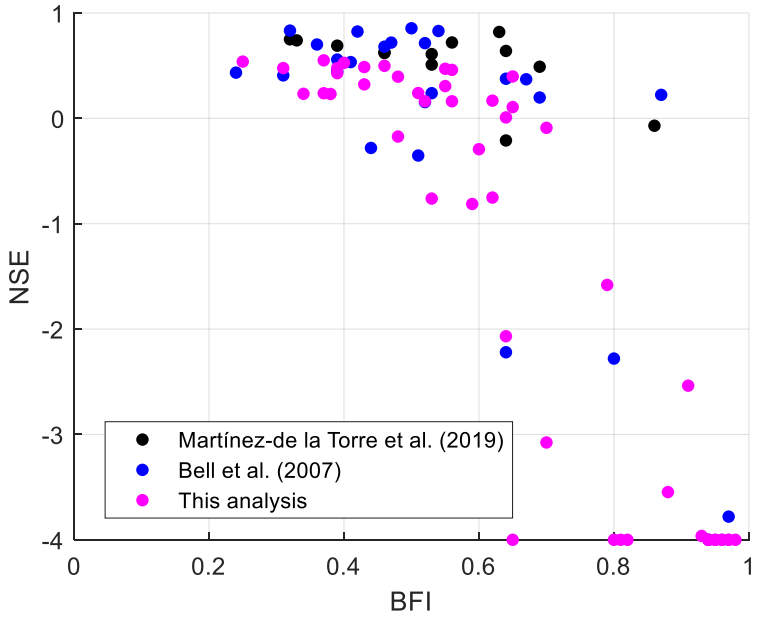


Figure 4.8: Model performance of JULES Default from two studies (Bell et al., 2007 and Martinez-de la Torre et al., 2019) and our analysis for 85 catchments. For visual purposes, the performance of JULES that showed NSE values of less than -4 are presented here as equal to -4.

In the last step, we want to investigate if our results are in line with previous studies. To this end, we include results of the performance metrics from two other research papers that applied Grid-to-Grid and JULES in Great Britain (Bell et al., 2007; Martinez-de la Torre et al., 2019). Both used Nash and Sutcliffe (NSE) (Nash and Sutcliffe, 1970) metric for evaluation. We plot the NSE values for 47 catchments from this analysis, 13 catchments from Martinez-de la Torre et al. (2019) and 25 catchments from Bell et al. (2007) (i.e., 85 catchments in total). Figure 4.8 depicts the comparison of the BFI against the NSE metric for all the 85 catchments. Due to visual purposes, NSE values less than -4 are depicted as equal to -4. We should note that points

from Martinez-de la Torre et al. (2019) represent values after calibration to improve the saturation excess mechanism of PDM. Applying calibration in an LSM, that in theory should rely on data and on explicit representation of processes, implies that there is a missing component, or the physics of the model fail to simulate the processes accurately. Thus, improving the model through calibration does not mean improvement of the processes and the realism of the model, but probably an indirect way to address a missing parameterization. Looking at the big picture, we can infer that flashy catchments with BFI less than 0.6 can achieve high performance with NSE up to 0.8. For catchments with BFI greater than 0.6, the performance of JULES Default starts to decline with NSE values less than 0.5, even for cases after calibration. Therefore, it is evident that the model fails to achieve an acceptable performance in groundwater-dominated catchments, even after applying calibration.

## 4.4 Discussion

Here we present an evaluation of the operational version of JULES (JULES Default) in Great Britain. We selected 47 catchments with different soil and land use characteristics, and minor artificial influences to examine the behaviour of the JULES model under different conditions. By doing this, we can highlight the limitations of the free drainage assumption that the model uses to simulate groundwater. We use the default parameters for river routing and saturation excess mechanism (PDM), following the guidelines from the JULES community and other similar studies (e.g., Bell et al., 2007). We have organised this section to answer the two specific research questions highlighted in the Introduction (Section 4.1). First, we discuss the performance of JULES Default to simulate runoff in different catchments. Then, we discuss about the impact of soil and land use on the model performance.

### 4.4.1 Performance of JULES Default on simulating runoff

First, we can highlight from the differences between observed and modelled streamflow values that there are three distinct patterns. JULES Default shows acceptable performance on runoff prediction for most of the flashy catchments with KGE values higher than 0.5 (Table 4-1). In flashy catchments, the dominant runoff generation mechanism is the infiltration excess mechanism and secondarily, the saturation excess mechanism. Both are represented in JULES-Default; thus, the model can predict the daily runoff satisfactorily. Note that the contribution of groundwater in these catchments is minor and consequently the performance of the model is high. The model usually underestimates the peaks, but it captures the timing. Also, it fails to

represent the recession curves and the baseflow when there is no rain (Figure 4.4 top left). The latter implies that by overestimating the runoff in dry seasons, JULES Default will underestimate the impact of drought and could possibly mitigate the impact of climate change in a region.

The second pattern comprises of the groundwater dominated catchments. In this kind of catchments, the soil has high saturated conductivity thus, a high proportion of rainfall is infiltrated in the soil domain. Then, the catchment discharges the groundwater with a lag response (Boorman et al., 1995) In these catchments, the baseflow during dry periods is high and sometimes dominates the annual water potential (Wheater et al., 2007). JULES Default overestimates the medium magnitude peaks more than 100% in a few cases (Figure 4.4 middle left), but it captures the significant floods (Figure 4.4 middle right). However, JULES Default fails completely to capture the baseflow, especially during dry periods. The model predicts streamflow values near zero during dry season, whereas the observations show much higher runoff values. This behaviour can be seen clearly in the scatter plot (Figure 4.4 middle right) in which JULES Default constantly underestimates the low flows. The weakness of the model to effectively simulate the runoff during the dry season caused worst performance during that period when the baseflow component has more significant influence on the total runoff (Figure 4.5). The implication of this behaviour is that the model may fail to capture the impact of climate change, as by underestimating significantly the low flows would alert the signal of seasonality and exaggerate the impact of climate change.

The missing groundwater parameterization in combination with the missing parameterization to capture the macroporosity behaviour of chalk catchments decreases even more the performance of JULES in this region. In these catchments, the observed runoff is usually constant with a few minor floods (Figure 4.4 bottom left). The baseflow strongly dominates the daily runoff and JULES Default fails entirely to capture this behaviour. In addition, the saturated conductivity values provided by the Land Information System for the chalk region are much lower from the literature. This deteriorates even more the performance of the model because the model reacts in the same way with an overland-dominated catchment. As we expected, the performance of the model (KGE) (Figure 4.5) is considerably worse during summer, because in this period there are no flood events and the total regime is dominated by groundwater, a process that JULES does not represent.

In JULES Default, the baseflow contribution to the total runoff comes directly from the free drainage flux that is linked with the saturated conductivity and the relative saturation in the last

layer of the soil domain. The model could partially capture this behaviour of a groundwater dominated catchment only after an extensive calibration of the soil characteristics or of the PDM parameters that is a common approach within JULES community (e.g., MacKellar et al., 2013; Martinez-de la Torre et al., 2019). Nonetheless, the necessity of calibration is a sign of a missing component in the model (assuming the soil property observations are representative to the 'true' values) that currently is addressed using the calibration process. However, if the benefit of an LSM is that it relies on data and on explicit representation of physics, then using calibration to improve realism in the model is contradictory to the scope of the LSM.

The purpose of this Chapter is to test the operational JULES Default in a series of catchments with different characteristics. The simulations allow us to select key factors (e.g., geology, soil texture, land use) that affect JULES performance. In doing so, we have limited the number of catchments to 47, all with minor artificial influences. We do not use the whole archive from NRFA that would need the simulation of the whole UK due to technical and time constraints. Furthermore, we pick only a 10-year period for the simulation of our evaluation performance. In this period, we include periods with droughts [2009-2012] and floods [2007, 2013] to get a clear picture. We think that 10 years of simulation is more than enough to evaluate the model performance.

### 4.4.2 Dominant factors that affect JULES model performance

Here, we try to find the dominant factors that influence model performance. First, we analyse the impact of geology (Figure 4.2a). It seems that there is a cluster of points with bad performance in South and South East of England, a cluster of points with poor performance in South of Wales and two clusters with acceptable performance in the East part of Wales and in North England. The first cluster of catchments overlay with Chalk and this can explain this behaviour of the model. Representing the groundwater to groundwater dominated regions may not solve the chalk regions that have different hydrology. Maybe later we could extend the application of JULES-GFB to include the chalk hydrology. The current version of JULES Default cannot yet simulate the macroporosity behaviour of chalk, despite recent efforts (Le Vine et al., 2016; Rahman and Rosolem, 2017). Looking the other clusters in the map (Figure 4.2), we find that catchments with acceptable performance overlay Mudstone, whereas catchments with poor performance overlay Limestone. This observation cannot be very informative if it is not linked to the saturated conductivity map (Figure 4.2c). It is found that catchments with acceptable performance have low values of saturated conductivity (100-150 cm $d^{-1}$), while catchments with poor and bad performance have saturated conductivity higher

than 150 cm $d^{-1}$. Soil texture map (Figure 4.2b) does not help to draw any specific conclusion. Catchments with clay and clay loam soil textures usually have better performance, but we can find catchments with the same type of soil texture that give negative values of NSE or KGE. Finally, considering land use, we find two dominant land use types, mainly Shrubs and C3 grasses. Both have catchments in which the performance is assessed as bad or acceptable. This is an indication that probably the land use has lower impact on the model performance. Note that C3 grasses and Shrubs have the same parameters in the model apart from the Snow-covered albedo for large leaf area index, set to 0.4 for Shrubs and 0.6 for C3 grasses, respectively (Best et al., 2011, Table 5). Considering that England and Wales usually are not covered by ice, we could say that we do not expect the land use type of C3 grasses and Shrubs to affect the model performance. In general, land use type affects mostly the evapotranspiration flux. Thus, the differences in the hydrograph would be less significant between observations and JULES Default, if the evapotranspiration was the dominant factor of the discrepancies. However, looking the geology and the land use without considering the hydraulic characteristics of the soil and aquifer is not enough to interpret the limitations and the strengths of the model.

We introduce the impact of BFI in our analysis, in a way to include the impact of aquifer properties indirectly through the BFI index. We use a histogram (Figure 4.6) to present the KGE values against BFI index, separating the plot for KGE above and below 0.5. In this way, we separate the overland and groundwater-dominated catchments. Interestingly, all the catchments with BFI less than 0.4 (flashy catchments) give acceptable KGE values above 0.5. On the other hand, most of the groundwater-dominated catchments give low KGE performance. Despite the small sample size, we can see the strength of the model to give better performance in flashy catchments compared to groundwater-dominated ones. The above statement can be substantiated by Figure 4.7 in which, the performance of the model is plotted by geology, soil texture and land cover type. We use the BFI value to colour each catchment/point, with blue to represent significant groundwater influence and red to represent flashy catchments with no significant groundwater influence in the streamflow. We do not find improvement or deterioration of the model horizontally that would imply that a specific geology or land cover type dominates the model performance (exception for chalk catchments). However, we find changes vertically with BFI to be the dominant factor. This implies that the percentage of the contribution of groundwater to the total regime dominates the model performance. The reason for this behaviour is the missing groundwater representation on JULES Default that affects the model in groundwater dominated regions.

Finally, we increase the sample size by including results from two other research studies that applied JULES (Martínez-de la Torre et al., 2019) or the hydrological parameterization of JULES (Bell et al., 2007) in Great Britain. Overall, we use 85 catchments [47 catchments from this analysis, 13 catchments from Martinez-de la Torre et al. (2019) and 25 catchments from Bell et al. (2007)], and we compare the performance of each catchment using the Nash and Sutcliffe (NSE) (Nash and Sutcliffe, 1970) metric against the BFI for each one. We could consider this plot as a multi-objective optimisation problem with BFI and NSE to be the two objectives. We want to find the causality between the cause (groundwater contribution described by BFI) and the effect (model performance described by the NSE metric). If we could "draw" an imaginative line to depict the Pareto frontier with solutions/catchments that are not dominated, we could find an interesting line that summarizes the above findings. This line would be flat for BFI values lower than 0.6 with NSE around 0.8. This part describes the flashy catchments with low groundwater contribution in which the model performs satisfactorily, as the missing groundwater representation has minor effect in this kind of catchments. However, the line for BFI values higher than 0.6 would have a steep slope to depict the deterioration of the model performance as the contribution of groundwater (BFI) increases. These catchments are heavily groundwater dominated, thus JULES without a groundwater parameterization cannot simulate them accurately even when calibration is applied, like in the results from Martinez-de la Torre et al. (2019). This is an evidence that JULES cannot perform in a robust way everywhere and it is needed the development of a new groundwater scheme to extend the applicability of the model.

## 4.5 Conclusions

In this Chapter we evaluate the operational version of JULES (JULES Default) in 47 different catchments in Great Britain with different size, geology, and land use type. The aim is to find limitations and weaknesses of the model and then, identify the dominant factors that affect the model performance. This will help us to identify missing components that are not represented in the current model. In order to extend our analysis, we use results from two other studies (Bell et al., 2007; Martínez-de la Torre et al., 2019). For the model evaluation, daily observed runoff from NRFA is used. Two metrics, namely KGE and NSE, are applied to assess the model performance.

Results from the experiment demonstrate that JULES Default shows three distinct behaviours, namely for flashy, groundwater dominated and chalk catchments. In particular, the

model gives acceptable performance (KGE > 0.5) in flashy catchments with low contribution of groundwater. In these catchments, the infiltration excess and the saturation excess mechanisms are sufficient to represent the runoff generation mechanisms. However, the model fails to simulate catchments in which the groundwater has significant role. The missing groundwater representation of the model affects the model performance by constantly underestimating the daily flows during dry season.

The dominant geology, land use, soil texture, saturated hydraulic conductivity and BFI for each catchment are analysed to find the dominant factor. Land use has minor influence in the model performance, as the two dominant types we find in this area, namely C3 grasses and Shrubs, have similar parameters in the model. Mudstone with lower saturated conductivity is linked with higher model performance. Overall, the performance of the model is associated with groundwater flow that is largely controlled by geology in the UK. Areas with high conductivity values (higher BFI values) showed bad model performance due to the missing representation of groundwater dynamics. Our major conclusion, which aligns with two previously published similar studies, is that JULES Default fails to simulate runoff accurately in groundwater dominated areas even after applying calibration (Martínez-de la Torre et al., 2019), while it performs accurately in overland flow dominated areas. However, in both overland and groundwater flow dominated catchments, the model fails to simulate low flows. This could possibly affect the response of the model under different climate change scenarios by alerting the impact of a drought. This analysis was our motivation to develop a new groundwater representation in JULES. In Chapter 5, we implement and test a new groundwater scheme in JULES, and we apply the new scheme in Central Britain to identify the impact of climate change on the water components in Chapter 6.

# 5 IMPLEMENTING THE NEW GROUNDWATER FLOW BOUNDARY SCHEME

This chapter is part of a paper published in the Hydrological Processes Journal. This paper has been modified to enhance consistency all through this dissertation.

**Citation:** Batelis, S.C., Rahman, M., Kollet, S., Woods, R. and Rosolem, R., 2020. Towards the representation of groundwater in the Joint UK Land Environment Simulator. Hydrological Processes. https://doi.org/10.1002/hyp.13767

## 5.1 Introduction

This chapter explains the experimental procedure we followed to test the new groundwater parameterization. We test JULES-GFB in comparison to JULES-FD, using two synthetic experiments: (1) single-column, and (2) tilted-V catchment, against an independent 3-D hydrological model (Parflow), used as our benchmark. In addition, we apply our new JULES-GFB model to regional domain in the UK, where groundwater has a key importance for runoff generation.

The use of synthetic experiments to test a new model is a common approach (e.g., Kim et al., 2012; Lee and Chang, 2005; Sulis et al., 2010; Clark et al., 2015) that allows to control better the variables and the forcing. In this work, we reproduce the same synthetic cases used in Rahman et al. (2019) and in Kollet et al. (2017) in order to compare the new JULES-GFB model with a widely used hydrological model (ParFlow), used as benchmark in this work.

ParFlow is a three-dimensional variably saturated hydrological model that can simulate the water cycle between the bedrock to the top of the plant canopy (Ashby and Falgout, 1996; Kollet and Maxwell, 2006; Maxwell, 2013). However, the amount of computational resources needed are not easily available everywhere, so a simpler model is needed to run the coupled model for large region (e.g. the whole UK) and for a long-time period for applications, like flood analysis. Coupling a complicated model, like JULES, with an even more complicated groundwater model, would increase the computational time and it would reduce its applicability. Once the model evaluation is done for both synthetic experiments, the test will be expanded to a regional modelling case using runoff observations and remotely sensed products for actual evapotranspiration (i.e., GLEAM product).

Every single experiment has a different purpose and is applied to test a different mechanism. The success all of them will give us confidence about the applicability and the correctness of the new model. The infiltration experiment evaluates how relative saturation dynamics are simulated by JULES including the newly implemented groundwater parameterization ignoring the impact of topography and lateral flow. This allows us to focus on processes related to potential recharge of the aquifer. In addition, the tilted-V catchment experiment will assess the magnitude and dynamics associated with both storage and flux terms focusing on the groundwater contribution to discharge generation. Finally, the regional experiment evaluates both versions of JULES (JULES-FD and JULES-GFB) against real streamflow observations and evapotranspiration products from remote sensing. It is important to mention that these series of experiments help identify advantages and limitations of the new approach when compared to a previous version of JULES. Note that a full application of this new version at larger (national to continental scales) is beyond the scope of this Chapter, which focuses on the model development.

Note that in the following experiments, we have modified the vertical discretization of the default JULES to be more consistent with the groundwater model. In this case, all JULES model versions discussed in this study have been set up with evenly spaced 20 cm thick soil layers from surface to 3 m depth for the infiltration experiment and the regional analysis, and 10 cm layers for the tilted-V experiment. We refer to this version of the model as JULES-FD (Free Drainage).

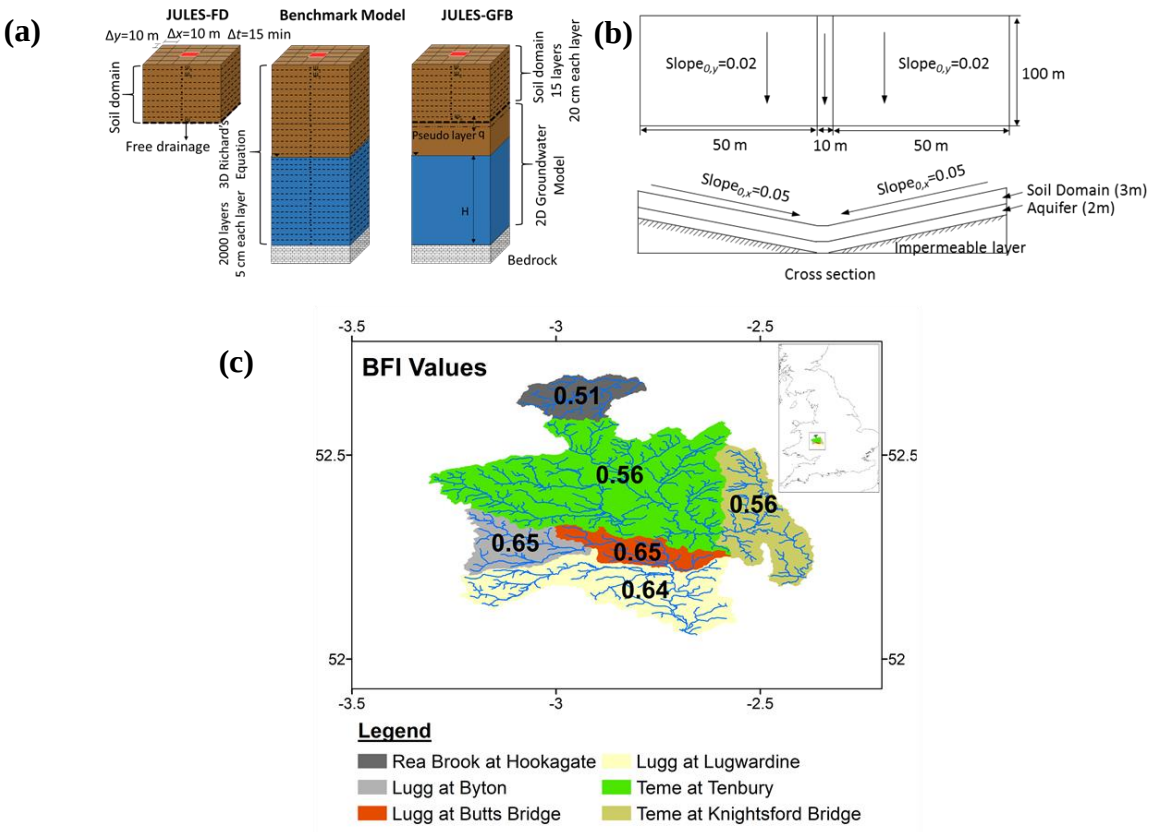


Figure 5.1: Schematic overview of experiments in the study: (a) infiltration single-column synthetic experiment, (b) synthetic groundwater discharge experiment, and (c) UK regional domain.

## 5.2 Experimental design

### 5.2.1 Synthetic infiltration experiment

We use a column experiment to test and compare the infiltration mechanisms in each model without the impact of lateral flow (Figure 5.1a). We evaluate how the wetting front changes after an applied rainfall pulse interacts with the aquifer (in both JULES-GFB and JULES-FD), in comparison to the benchmark ParFlow model. Each experiment comprises a 5-day simulation with an applied constant rainfall rate of 5 mm $h^{-1}$ in the first 10 hours of the experiment, assuming no surface water loss through evaporation. In our case, the zero contribution from evaporation is achieved by setting both the shortwave radiation and the wind speed to zero; longwave radiation is set to a constant value of 219 W $m^{-2}$, while temperature and specific humidity are set to 293.15 K and 0.00757 kg $kg^{-1}$, respectively. The temporal resolution is set to 15 minutes. The setup is the same used by Rahman et al. (2019).

In order to evaluate model performance, we use the mean bias of relative saturation (unitless) which is defined as the fraction corresponding to the ratio of actual moisture content to maximum soil moisture possible for each soil layer (ranging from 0 to 1 which corresponds to fully dry and fully wet, respectively). In JULES model, relative saturation is defined as soil wetness. The bias is calculated from the soil profiles (i.e., surface to 3 m depth) for each of the two JULES versions (JULES-FD and JULES-GFB) relative to the results from the ParFlow benchmark model obtained from Rahman et al. (2019). By assessing mean bias, the results will indicate conditions where models underestimate (overestimate) the benchmark model because of drier (wetter) relative saturation conditions. Our comparisons use a combination of 12 different soil types (parameters defined according to Table 5-1) and 12 different initial conditions of water table depth (from shallow, located at 0.25 m, to deeper, located to maximum depth of 30 m below the surface, following Rahman et al., 2019).

Table 5-1: Hydraulic properties for different soil texture classes (source: Rahman et al., 2019).

| | van Genuchten parameter $b$ | van Genuchten parameter $1/\alpha_v$ | Saturated conductivity ($K_{hs}$) | Volumetric soil moisture content at saturation | Volumetric soil moisture content at the critical point | Volumetric soil moisture content at the wilting point | Specific yield $S$ |
|---|---|---|---|---|---|---|---|
| Units | (-) | ($m^{-1}$) | (mm $s^{-1}$) | ($m^3$ water per $m^3$ soil) | ($m^3$ water per $m^3$ soil) | ($m^3$ water per $m^3$ soil) | (-) |
| Clay | 4.00 | 0.67 | 1.70E-04 | 0.459 | 0.396 | 0.272 | 0.02 |
| Clay loam | 2.44 | 0.63 | 9.40E-04 | 0.442 | 0.318 | 0.197 | 0.04 |
| Loam | 1.03 | 0.67 | 1.40E-03 | 0.399 | 0.27 | 0.117 | 0.11 |
| Loamy sand | 1.35 | 0.29 | 1.20E-02 | 0.39 | 0.125 | 0.055 | 0.22 |
| Sand | 0.46 | 0.28 | 5.80E-02 | 0.375 | 0.091 | 0.033 | 0.25 |
| Sandy clay | 5.00 | 0.30 | 1.30E-03 | 0.385 | 0.339 | 0.239 | 0.07 |
| Sandy clay loam | 3.03 | 0.48 | 1.50E-03 | 0.384 | 0.255 | 0.148 | 0.1 |
| Sandy loam | 2.27 | 0.38 | 4.40E-03 | 0.387 | 0.207 | 0.095 | 0.12 |
| Silt | 1.49 | 1.54 | 5.10E-03 | 0.489 | 0.3 | 0.06 | 0.08 |
| Silty clay | 3.13 | 0.62 | 1.10E-03 | 0.481 | 0.387 | 0.25 | 0.01 |
| Silty clay loam | 1.92 | 1.20 | 1.20E-03 | 0.482 | 0.366 | 0.208 | 0.03 |
| Silt loam | 1.52 | 2.00 | 2.10E-03 | 0.439 | 0.33 | 0.133 | 0.05 |

### 5.2.2 Synthetic groundwater discharge experiment

In the previous experiment, it was examined the impact of a rainfall event in the infiltration mechanism and in the development of the wetting front. However, lateral flow was prevented from happening, due to lack of changes in topography. For this experiment, we introduce the effects of topographic slope in driving the subsurface flow in a tilted-V shaped synthetic catchment (Figure 5.1b). Our specific objective here is to evaluate the contribution of the lateral flow to the simulated discharge at the outlet. We follow the same set up used in the Integrated Hydrologic Model Intercomparison Project 2, IH-MIP2 (Kollet et al., 2017), which also includes the benchmark ParFlow model. In addition, the IH-MIP2 includes the simulation results from additional integrated hydrological models, which allows us to evaluate the overall performance of JULES-GFB more robustly with respect to a range of similar models. In addition to ParFlow, IH-MIP2 models include the Advanced Terrestrial Simulator (ATS) (Coon et al., 2016; Painter et al., 2016), Cast3M (Weill et al., 2009), CATchent Hydrology (CATHY) (Bixio et al., 2002; Camporese et al., 2010), GEOtop (Endrizzi et al., 2014; Rigon et al., 2006), HydroGeoSphere (HGS) (Aquanty, 2015), and MIKE-SHE (Abbott et al., 1986; Butts et al., 2004). For details about the IH-MIP2 and its models, please refer to Kollet et al. (2017).

The tilted V-shaped topography is perfectly symmetrical in both directions with spatial resolution defined as 10 m. The width of the V-shaped catchment is 110 m, corresponding to two 50 m slopes separated by a 10 m wide river channel. The length of the catchment is 100 m (Figure 5.1b) and the soil column is 5 m deep at all locations. The slope in the x direction is $Slope_{0,x}$=0.05 and in the y direction is $Slope_{0,y}$=0.02. The water table depth is initialized at 2 m below land surface with hydrostatic conditions vertically. We simulate 120 hours with no rainfall and no loss from evapotranspiration and run with timestep of 5 seconds. We specify the soil properties based on a sandy soil following Kollet et al. (2017). Our JULES-GFB and JULES-FD simulations are run without the River Flow Model parameterization due to the small size of the catchment. We, therefore, assume that the surface storage from all the cells within the domain contribute to the outlet of the catchment in the same timestep. The evaluation of JULES-GFB and JULES-FD will be based on bias and linear correlation computed against the benchmark and other models used in IH-MIP2 for saturated and unsaturated storages, as well as for the outlet discharge. In JULES-FD, we extended the soil domain from 3 to 5 meters to directly compare the same initial conditions in terms of saturated and unsaturated storages to other IH-MIP2 models.

### 5.2.3 Regional analysis

Synthetic experiments can be useful tools to validate new model developments. However, they pose the limitation of using ideal conditions and forcing, often assuming homogeneous spatial characteristics (soil type, land use). Here, to test the model further, we compare our new JULES-GFB model against the JULES-FD in a real-world application. In this case, the experiment is carried out over a regional domain in the UK, which consists of six neighbouring catchments (with individual areas greater than 100 $km^2$ and near natural conditions) characterized by being groundwater dominated (i.e., relatively high Base Flow Index, BFI) (Figure 5.2c and Table 5-2). The study area described here is used during the testing of the model (Chapter 5), while we utilise the same area to assess the impact of climate change (Chapter 6).

The elevation ranges from 0 to 980 m within the domain (Figure 5.2a). The elevation map is post-processed in order to provide flow directions following the methodology presented in Maxwell et al. (2009). Dominant vegetation types within the domain are shrubs (50%) and C3 grasses (43%), as shown in Figure 5.2b. Soil classes are shown in Figure 5.2c with dominant clay loam (31%) and silty clay loam (30%) classes within the domain. Note that we did not apply any parameter calibration to JULES-GFB. Instead, we used an ad hoc sensitivity analysis of the aquifer conductivity to identify the impact on the model performance.

In order to simplify our initial test with JULES-GFB, we define key aquifer properties with spatially uniform information from domain-average parameters also obtained from the Land Information System. Furthermore, we assume a homogeneous depth of aquifer equal to 100 m, which is a common approach for large-scale groundwater applications (e.g., Condon and Maxwell, 2015; Keune et al., 2016). We set a constant specific yield to 0.2, following a similar approach by previous studies (Niu at al., 2007; Ganji and Sushama, 2018). The focus of this study is to initially test the introduction of the groundwater parameterization into the JULES model and setting homogeneous properties is sometimes a common strategy (e.g., Niu et al., 2007; Condon and Maxwell, 2015). The impact of introducing heterogeneous aquifer properties and depth to bedrock is beyond the scope of this study and can be further tested in the future.

The vertical discretization of JULES-FD and JULES-GFB is 0.2 m, similar to the column-scale experiment with both model versions extending their soil domain depth to 3 m. We used a timestep of 10 minutes which ensures satisfactory runoff predictions without adding too much computational time.

Table 5-2: Catchment characteristics for the regional analysis.

| Name | Catchment area ($km^2$) | Mean flow ($m^3 s^{-1}$)* | BFI** | Rainfall (mm $year^{-1}$)* |
|---|---|---|---|---|
| Lugg at Byton | 203 | 3.9 | 0.65 | 1039.1 |
| Lugg at Butts Bridge | 371 | 6 | 0.65 | 981.7 |
| Lugg at Lugwardine | 886 | 10.8 | 0.64 | 845.4 |
| Teme at Tenbury | 1134 | 14.6 | 0.56 | 755.1 |
| Teme at Knightsford Bridge | 1480 | 18.2 | 0.56 | 736.4 |
| Rea Brook at Hookagate | 178 | 1.7 | 0.51 | 822.9 |

* Mean flow and Annual Mean Rainfall are based on NRFA and CHESS datasets from 1970-2015, respectively.

** BFI was computed by NRFA based on the archived record of gauged daily mean runoff.

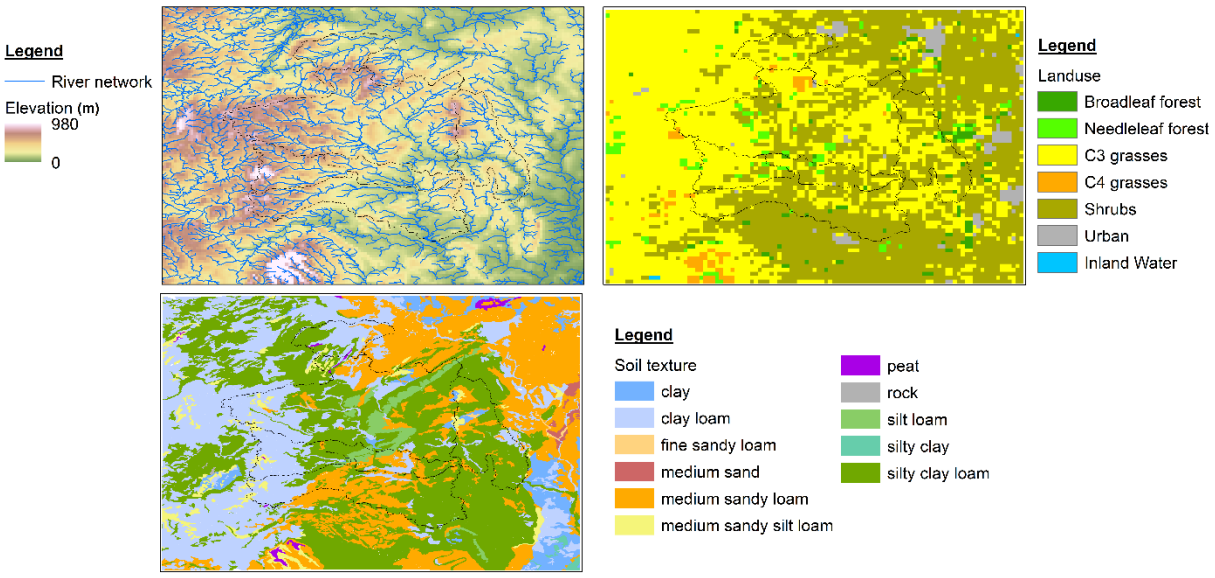


Figure 5.2: Physical characteristics of the regional domain experiment: (a) river network and topography (HydroSHEDS); (b) land cover (Land Cover Map 2007); (c) soil texture (Land Information System).

The regional forcing data are obtained from the Climate, Hydrology and Ecology Research Support System (CHESS) dataset (Robinson et al., 2016). The regional simulation encompasses the year between 2008 and 2012. The different JULES versions are each individually spun up following common procedures. Particularly to our case, we cycle the 2008-2012 period repeatedly until the mean monthly soil moisture does not deviate by more than 0.1% from the previous year, following the same protocol used for the Large-Scale Biosphere-Atmosphere

Experiment in Amazônia, Model Intercomparison Project (LBA-MIP) (de Goncalves et al., 2008). The initial soil moisture initialize for the first cycle is set so that all soils layers across the domain have a relative saturation of 0.95. The meteorological conditions during selected 2008-2012 period show a mean annual temperature of 9.4 °C and mean annual precipitation of 871 mm (Figure 5.3) which are very similar to the 1981-2015 climatological record for the region calculated from the CHESS (9.3 °C and 886 mm, respectively).

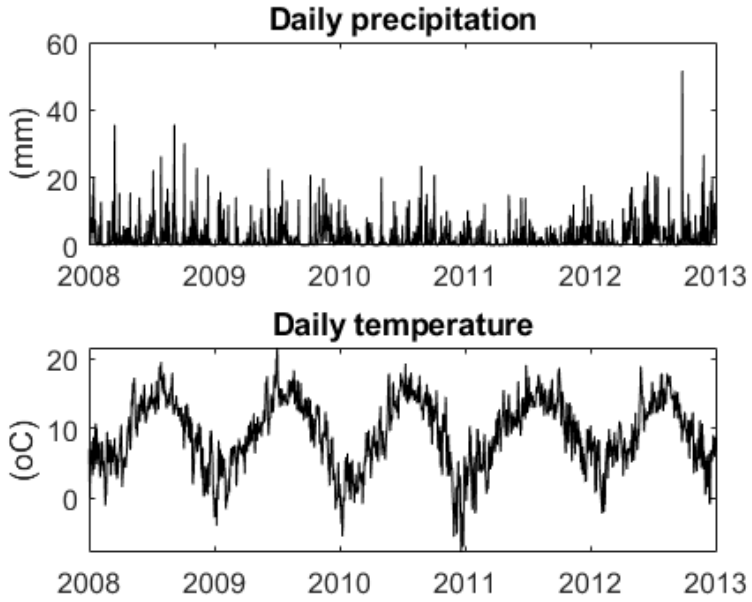


Figure 5.3: Daily time-series of domain average input meteorological forcing for precipitation (m) and air temperature (°C) used in the regional experiment.

As it was described in Section 2.1, PDM is the subgrid variability model based on statistical parameterization to simulate saturation excess runoff. For JULES-FD, we tested different values of $b_{pdm}$, ranging from 0.2 to 1. In JULES-GFB, PDM is not used, because JULES-GFB can directly simulate saturation excess in those cells which have the water table at the surface.

In this experiment, we evaluate all versions of JULES against daily streamflow data from individual catchments (Figure 5.2 and Table 5-2) provided by the National River Flow Archive (NRFA), covering 2008-2012 period. We choose Mean Bias, the Pearson correlation coefficient, and the Kling-Gupta Efficiency (KGE) (Gupta et al., 2009), as model performance metrics.

In addition, we test these versions of JULES in representing the evapotranspiration flux within the domain and we compare model estimates against the daily Global Land Evaporation

Amsterdam Model (GLEAM) (Miralles et al., 2011; Martens et al., 2018) dataset (version 3.3a). This dataset calculates evapotranspiration from the water balance equation, based on satellite data to estimate the different components of the water cycle.

## 5.3 Results

### 5.3.1 Synthetic infiltration experiment

Figure 5.4 depicts the soil moisture profile from the land surface to 3 m depth for JULES-FD (on the left), the ParFlow Benchmark Model (middle) and JULES-GFB (right) for three different combinations of soil type and initial water table depth (WTDi). The simulation is carried out for five days with a constant rainfall rate of 5 mm $h^{-1}$ during the first ten hours only and with no water loss through evaporation.

We observe two distinct behaviours from the simulations. First, for the two cases where the water table is initialized within the soil domain of JULES (i.e., shallower than 3 m), JULES-FD is unable to retain the soil moisture, resulting in much drier soil moisture profiles when compared to the Benchmark Model. This behaviour is even more pronounced for relatively coarser textured soil types (e.g., sand), despite the initial water table being very close to the surface (i.e., at 0.5 m depth). The quick drying behaviour observed at the beginning is a result of initially high hydraulic conductivity as a result of the gravity flow driving the vertical water flow. This drying behaviour is sustained until the end of the simulation in JULES-FD due to a lack of physical constrain at the bottom of the domain (i.e., gravity flow drives loss of water leaving the bottom of the soil domain). Unlike the JULES-FD results, the new JULES-GFB shows remarkably good agreement with the Benchmark Model in both shallow water table cases shown in Figure 5.4, except for some minor differences close to the surface that are related to the different water partitioning in ParFlow, compared to the two JULES models.

Secondly, we notice that when the water table is initialized outside of the soil domain in the model (i.e., deeper than 3 m), there are only minor differences between JULES-FD and JULES-GFB in the soil moisture dynamics within the profile. Additionally, we observe some differences for both JULES versions from the Benchmark Model. This suggests that improvements from using JULES-GFB instead of JULES-FD are expected to occur more often for cases where the water table is shallow, and for relatively coarse soil types.

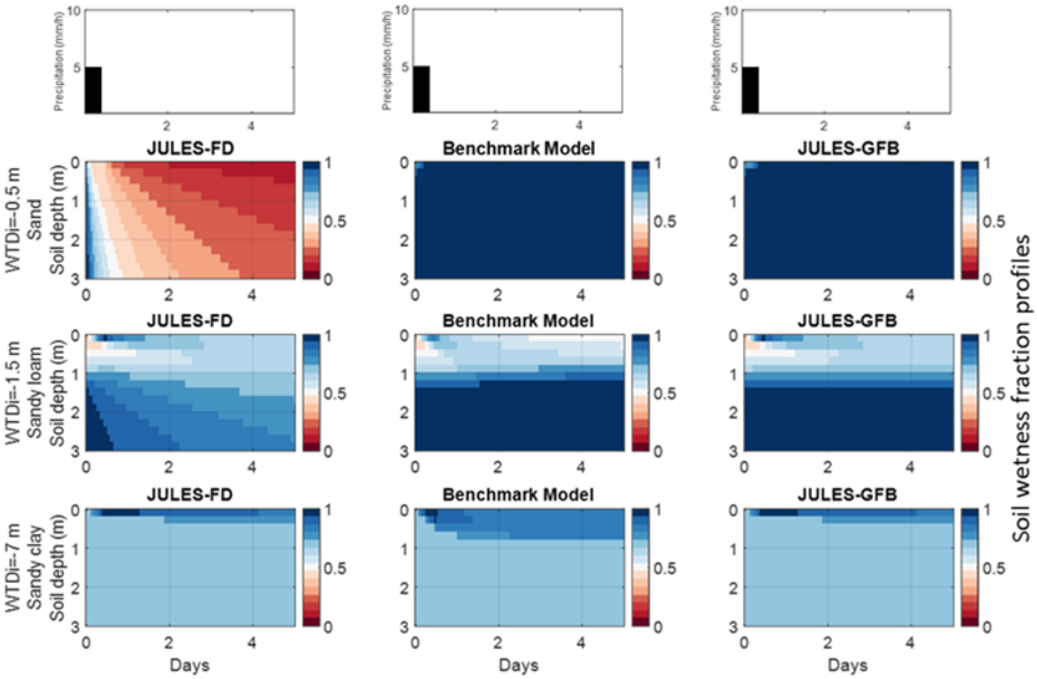


Figure 5.4: Examples of relative saturation profiles within the JULES soil domain comparing the JULES-FD (left) and the new JULES-GFB approach (right) against the Benchmark Model simulations (centre), for the infiltration experiment.

In order to investigate this result more broadly, we carried out a number of simulations using different combinations of soil types and initial water table depths. We test typical 12 different soil types with 12 different initial Water Table Depths, resulting in a total of 144 combinations. For each combination, we calculate the Mean Bias between each of the JULES versions (i.e., FD or GFB) against the Benchmark model for the entire 5-day simulation period (Figure 5.5). In this case, we use the soil moisture profile corresponding to the domain defined between the surface and down to 3m depth (like those shown in Figure 5.4). The results in Figure 5.5 corroborate the selected example cases shown in Figure 5.4. JULES-FD errors suggest systematically drier conditions within the top 3 m of the domain with slightly worse performance for relatively coarse soils. Conversely, the new JULES-GFB model is characterized by very low Mean Bias across different combinations when the water table is initialized within the 3m soil depth. However, the deviation from the Benchmark Model, when the water table is initialized below 3m depth, reveals no differences between JULES-FD and JULES-GFB. As expected, most of the impact of having the new JULES-GFB model is observed for shallow water table cases.

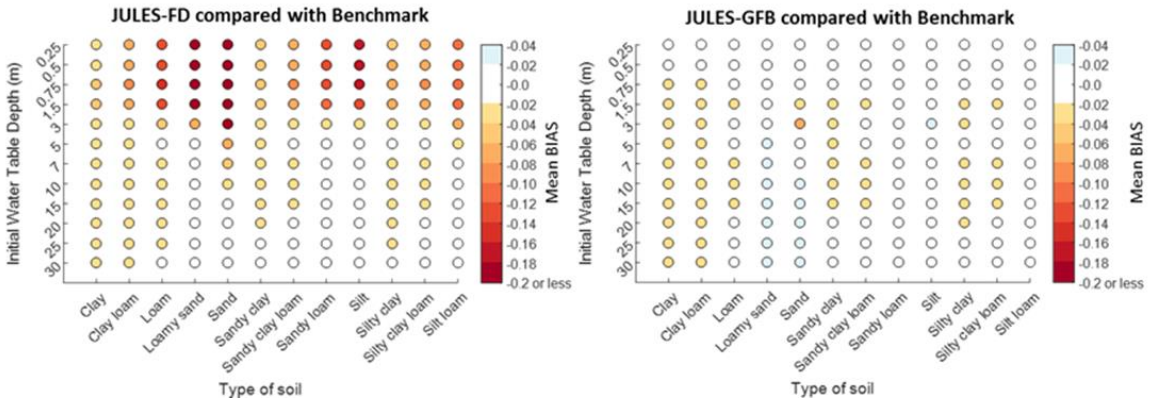


Figure 5.5: Overall model performance for the infiltration experiment computed within the soil domain as mean bias (Bias) with respect to the Benchmark Model: (left) JULES-FD and (right) JULES-GFB model.

### 5.3.2 Synthetic groundwater discharge experiment

Figure 5.6 compares JULES-GFB with the IH-MIP2 models in terms of storage dynamics of the saturated and unsaturated zones (averaged over the entire model domain), as well as outlet discharge rates. The behaviour of JULES-GFB (red line) is broadly consistent with the other IH-MIP2 models (grey lines) and the benchmark model in terms of dynamics (linear correlation) and overall magnitude (mean bias) with both computed relative to the benchmark model (Table 5-3). On the contrary, JULES-FD shows a very distinct behaviour from all the other models due to the different assumption of the free drainage parameterization.

The initial reduction in saturated storage of JULES-GFB is consistent with water moving downwards and slowly contributing to the outlet streamflow (Figure 5.6a). All models are able to capture this initial behaviour with some expected model-to-model differences. Interestingly, the JULES-GFB model shows the highest amplitude in change of saturated storage initially, when compared to other models, although the overall magnitude appears to be consistent with the average behaviour of all models. The reason for this steep behaviour may be because in JULES-GFB the water from the soil domain instantly reaches the water table. As discussed by Rahman et al. (2019), this behaviour will tend to happen even faster for relatively coarse soils given the intrinsic limitations of linearly extrapolating pressure head values from bottom of soil domain to where the water table is located. Towards the end of the first day, we observe in all IH-MIP2 models a consistent small increase in the saturated storage followed by a more gradual

decrease of storage in the following days. JULES-GFB shows similar behaviour with slightly larger amplitude than other models, suggesting a faster decrease in storage from approximately day three. JULES-FD has a very distinctive behaviour, with saturated storage almost completely draining within the first timestep. This fact drying behaviour is consistent with the physical processes highlighted previously in the column experiment for JULES-FD (Figure 5.4).

Furthermore, the initial increase in unsaturated storage of JULES-GFB is consistent with the initial drop of the water table (Figure 5.6b). At the end of the first day, the water table has returned to the initial condition and starts to contribute to the outlet streamflow. After the first day, there is increase of unsaturated storage that is linked to the contribution to the runoff for almost all the models. JULES-GFB shows similar behaviour with slightly larger amplitude than the other models, suggesting a faster increase in storage from approximately day three. On the other hand, JULES-FD simulation deviates drastically from the other models resulting in very high bias and low correlation against the benchmark model (Table 5-3).

Finally, Figure 5.6c presents the outlet discharge produced by the contribution to the surface runoff, as there is no rainfall in this experiment and no other source of runoff generation rather than the groundwater. At the end of the first day, all the models, except JULES-FD, start to produce surface runoff almost at the same time. The models (except JULES-FD) follow the same behaviour and dynamics with different magnitude. JULES-GFB produces lower values of outlet discharge than any of the IH-MIP2 models. It appears that JULES-GFB produces more unsaturated storage compared to the benchmark (Figure 5.6a, 5.6b and Table 5-3), possibly indicating that JULES-GFB retains more water in the soil domain compared to the other models which generate more runoff. Overall, JULES-GFB shows similar magnitude and dynamics compared to other models with metrics of bias and linear correlation to be within the model envelope. JULES-FD does not produce any runoff, as would be expected. Soil water fluxes are exclusively vertical and hence the water leaves the bottom of the soil domain without contributing to runoff generation.

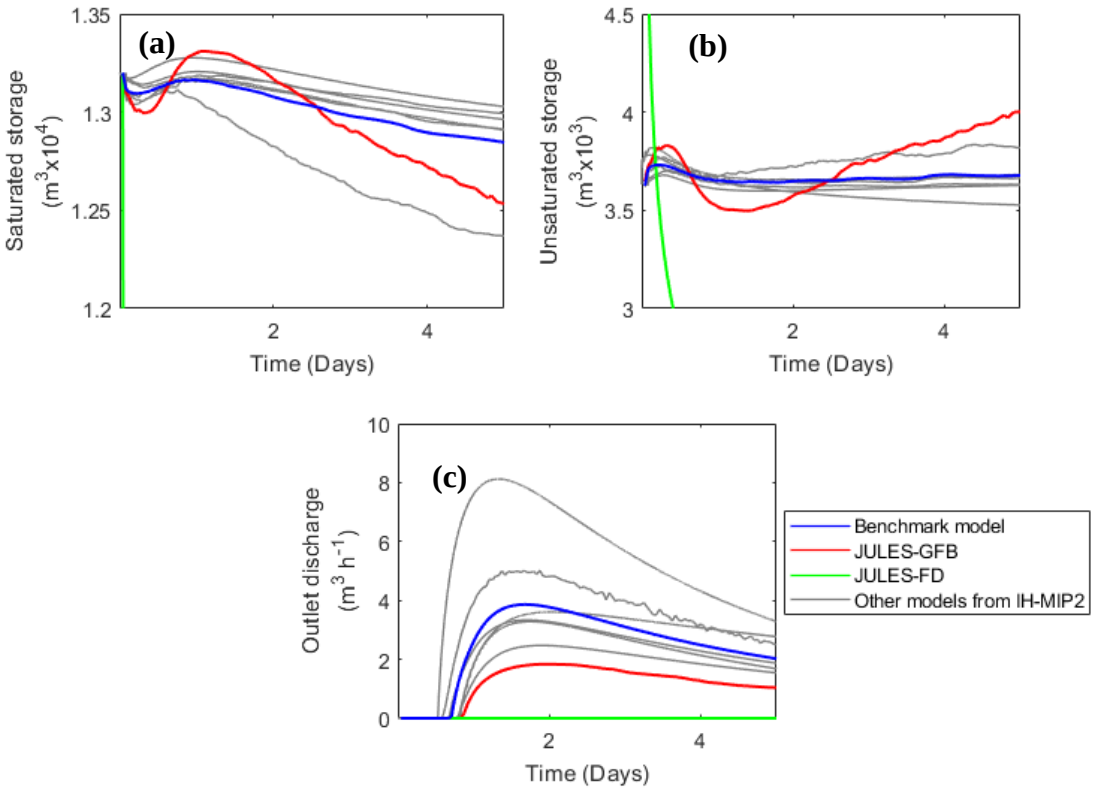


Figure 5.6: Comparison of (a) saturated storage, (b) unsaturated storage and (c) outlet discharge: JULES-GFB shown as red line, JULES-FD as green line and benchmark model as blue line. The grey lines correspond to other six Integrated Surface-Subsurface Hydrological Models (IHSSMs) investigated in the IH-MIP2 study (see Kollet et al., 2017).

Table 5-3: Performance metrics for the tilted-V synthetic experiment computed against benchmark model (ParFlow). The column in the right (Models in IH-MIP2) shows the metrics in terms of observed ranges from all models.

| **Saturated Storage** | **JULES-FD** | **JULES-GFB** | **Models in IH-MIP2** |
|---|---|---|---|
| Bias ($m^3$) | -12911 | -53 | -280 to 145 |
| Correlation | 0.157 | 0.965 | 0.965 to 0.996 |
| | | | |
| **Unsaturated Storage** | **JULES-FD** | **JULES-GFB** | **Models in IH-MIP2** |
| Bias ($m^3$) | -1542 | 64 | -67 to 83 |
| Correlation | 0.34 | 0.56 | 0.15 to 0.94 |
| | | | |
| **Runoff** | **JULES-FD** | **JULES-GFB** | **Models in IH-MIP2** |
| Bias ($m^3\ h^{-1}$) | -252 | -1.3 | -0.8 to 2.6 |
| Correlation | - | 0.964 | 0.86 to 1 |

### 5.3.3 Regional analysis

We carry out a comparison between both JULES models by setting up a regional domain experiment in the UK, where groundwater contributes significantly to streamflow. We assess the results in three ways: (1) by analysing spatial patterns of soil moisture, (2) by comparing modelled and simulated domain-average evapotranspiration, and (3) by comparing modelled and observed runoff at selected catchments within the domain. For JULES-FD, we tested different values of *b* exponent ($b_{pdm}$) from zero (the "JULES-FD-noPDM" case) to 1 (the "JULES-FD-PDM" case), following the discussion at Section 5.1.3. The test of multiple $b_{pdm}$ values represents an extra parameter calibration step in JULES-FD which is not needed in JULES-GFB.

**Spatial Patterns of Relative saturation**

We first begin by assessing the behaviour of JULES in reproducing spatial patterns of soil moisture within the regional domain. Figure 5.7a and 5.7b show the annual average relative saturation for the 5-year period (2008-2012) at the upper soil layer (0-20 cm) for JULES-FD-noPDM and JULES-FD-PDM $b_{pdm}$=1, respectively. Both simulations suggest a much smoother spatial pattern without being able to capture similar patterns from the river flow network and topography (Figure 5.2; top left). The results from both JULES model simulations suggest relatively drier conditions in the northeast region of the domain which is likely due to this region being characterized by relatively coarser soil (Figure 5.2; bottom) and also receiving less precipitation when compared to the west region of the domain (i.e., the west region receives annually about 1200 mm $yr^{-1}$ compared to about 700 mm $yr^{-1}$ in the east region of the domain; data not shown). This could explain the heterogeneity of Figure 5.7a and b.

The spatial pattern of soil moisture for JULES-GFB differs drastically from both JULES-FD simulations (Figure 5.7c), showing a distinct river network of saturated areas as a result of the convergence of groundwater in JULES-GFB. The pattern resembles quite satisfactorily the observed river network (Figure 5.2; top left). Soil type (coarse versus fine) continues to play a role in controlling soil moisture dynamics with the north-eastern region of the domain (relatively coarser soils as in Figure 5.2; bottom) suggesting drier conditions compared to the rest of the domain. We found no strong control by the land cover (Figure 5.2; top right) on annual soil moisture conditions, which is expected in more humid (cold) regions such as in Wales.

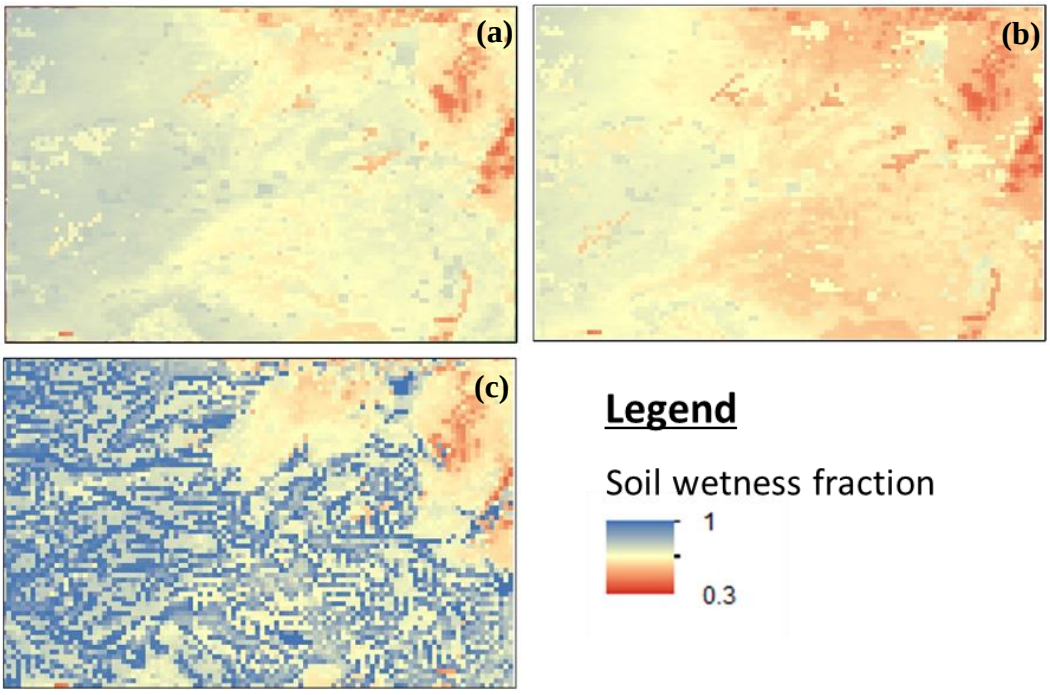


Figure 5.7: Annual average relative saturation at first soil layer (0-20cm) from JULES-FD-noPDM (a), JULES-FD PDM ($b_{pdm}$ = 1) (b), JULES-GFB (c).

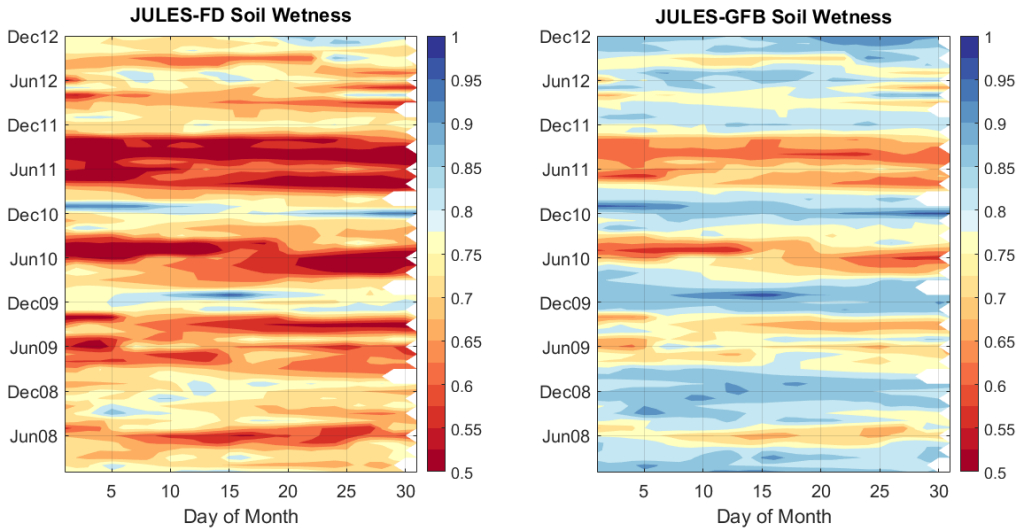


Figure 5.8: Domain-average daily soil moisture for upper 20 cm for JULES-FD $b_{pdm}$=1 (left) compared to JULES-GFB (right), for the regional experiment. The figure shows day of the month in the horizontal axis and months increasing from 2008 to 2012 in the vertical axis to better highlight winter and summer seasons.

Figure 5.8 depicts the daily relative saturation, as a domain average for the upper 20 cm for JULES-FD $b_{pdm}$ =1 against JULES-GFB. We present the JULES-FD $b_{pdm}$ =1 case, because after the calibration of $b_{pdm}$ in JULES-FD against observed runoff, the optimal value for $b_{pdm}$ was equal to 1 for 5 out of six catchments. Note that the results do not change dramatically for relative saturation if we use another configuration for PDM. The first observation is that JULES-GFB is much wetter, especially during the summer months (June - July - August) with relative saturation value about 0.7, when for JULES-FD is less than 0.5. The second observation is about the low values of JULES-FD during the winter months (December – January - February). For JULES-FD, it is unusual to have relative saturation values above 0.7, even for months with high precipitation. Finally, the annual range of relative saturation seasonality is much lower for JULES-FD compared to JULES-GFB.

**Domain-average Evapotranspiration**

Since this part of the UK is not a water-limited area, the impact of groundwater model on the simulation of evapotranspiration (ET) is expected to be low, particularly on the annual basis. Nevertheless, we evaluate JULES simulated ET against the GLEAM ET product (Figure 5.9). Monthly time-series of ET for the study period indicate that more pronounced differences are expected to occur over the summer period with little or no differences observed in the wintertime (Figure 5.9; left). During summertime, more realistic and relatively wetter soil moisture conditions within the domain (Figure 5.8) obtained with JULES-GFB result in a slightly increase in ET rates over the summer approaching GLEAM ET's estimates. The overall bias obtained with daily ET estimates computed for JULES-FD relative to GLEAM is -0.22 mm $d^{-1}$, whereas the mean bias for JULES-GFB relative to GLEAM reduces to -0.01 mm $d^{-1}$.

In previous sections, we highlighted the strong link between soil type and performance of JULES-GFB given the underlying assumptions of the groundwater parameterizations discussed in section 3.2.3 and in more detail by Rahman et al. (2019). In order to further understand this behaviour, we have computed the mean bias for 4 individual categories of soil types found within the domain (Figure 5.9; right). In this case, we focus on the summertime period defined to be between April and October. Our results suggest that for all types of soils within the domain, JULES-GFB biases are relatively lower than those obtained with JULES-FD when compared against GLEAM ET product. Notice that the regional domain does not necessarily show a large range of predominant soil types (Figure 5.2; bottom). Despite the similarities between the boxplots obtained for JULES-FD and JULES-GFB, we notice that the relative improvement (i.e., bias reduction here taken from the median) in JULES-GFB compared to

JULES-FD tend to be larger for the two "coarser" soil types (silty clay loam and medium sandy loam) and smaller for the two "finer" soil types (clay loam and clay). As we expect the limitations of free drainage assumptions to be more pronounced in relatively coarse soil types, the performance improvement by adding the groundwater parameterization in those soils is expected to result in more impact than in clay soils, where soil water dynamics are relatively slow and the limitations of assuming free drainage are reduced.

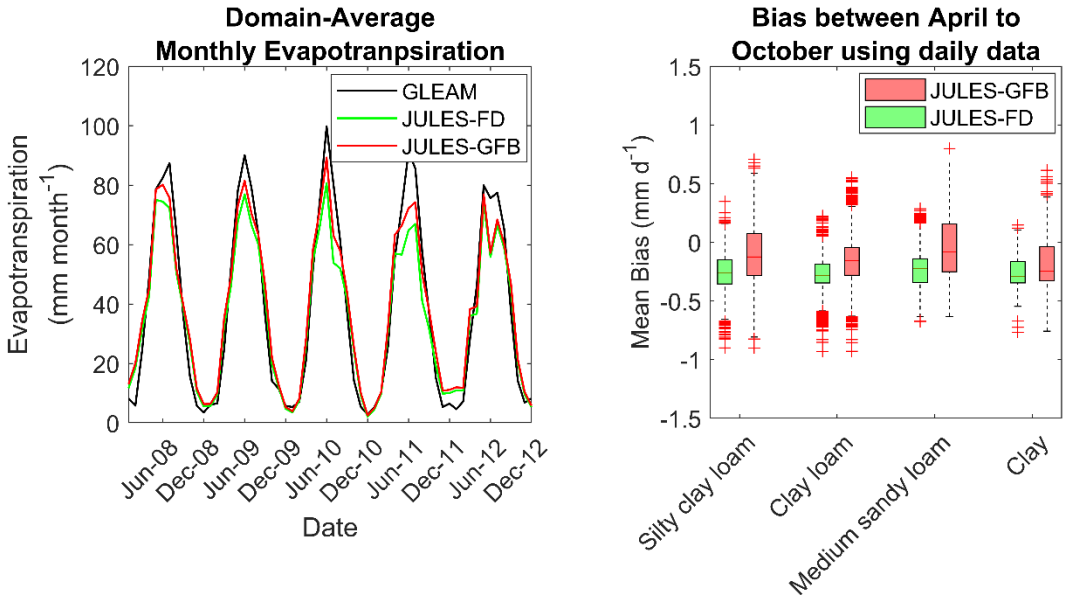


Figure 5.9: Left: Comparison of domain-average monthly evapotranspiration from GLEAM (black), JULES-FD (green) and JULES-GFB (red). Right: Mean bias compute relative to daily evapotranspiration estimates from GLEAM between April and September ("summer" season) and combined into four dominant soil types in the regional domain.

### Total Runoff at Catchment Scale

Figure 5.10 presents the runoff time-series for one catchment for a subset of the simulation period. The upper subplot compares the performance of JULES-GFB against JULES-FD-noPDM. The performance of JULES-FD-noPDM in reproducing the total runoff is very week with consistently missing runoff peaks and eventually capturing the large events only, while still showing lower peak magnitudes. The reason is that only the infiltration excess mechanism is activated, so the rainfall rate should be higher than the infiltration rate to produce surface runoff. On the other hand, JULES-GFB tends to capture more accurately the dynamics of the total runoff with some remaining limitations in reproducing low river flows. It seems that the contribution of the baseflow to the total runoff improves the runoff predictions. The bottom

panel of Figure 5.10 depicts the same comparison between JULES-GFB and JULES-FD against observations, but now JULES-FD is run with PDM enabled and with $b_{pdm}$ set to 1 (i.e., assuming an extra calibration step is taken with JULES-FD). Notably, the use of PDM improves remarkably the performance of JULES-FD suggesting similar behaviour as seen in JULES-GFB. The KGE metric computed relative to the observations was found to increase from 0.22 to 0.51 for the JULES-FD before and after introducing the PDM with the calibrated parameter, respectively; while the KGE obtained for JULES-GFB is 0.74. Notice that in the case of JULES-FD with PDM, a further calibration step is required hence adding an extra level of complexity in the model which is not needed in JULES-GFB, also limiting JULES applications in places where observations are not necessarily readily available.

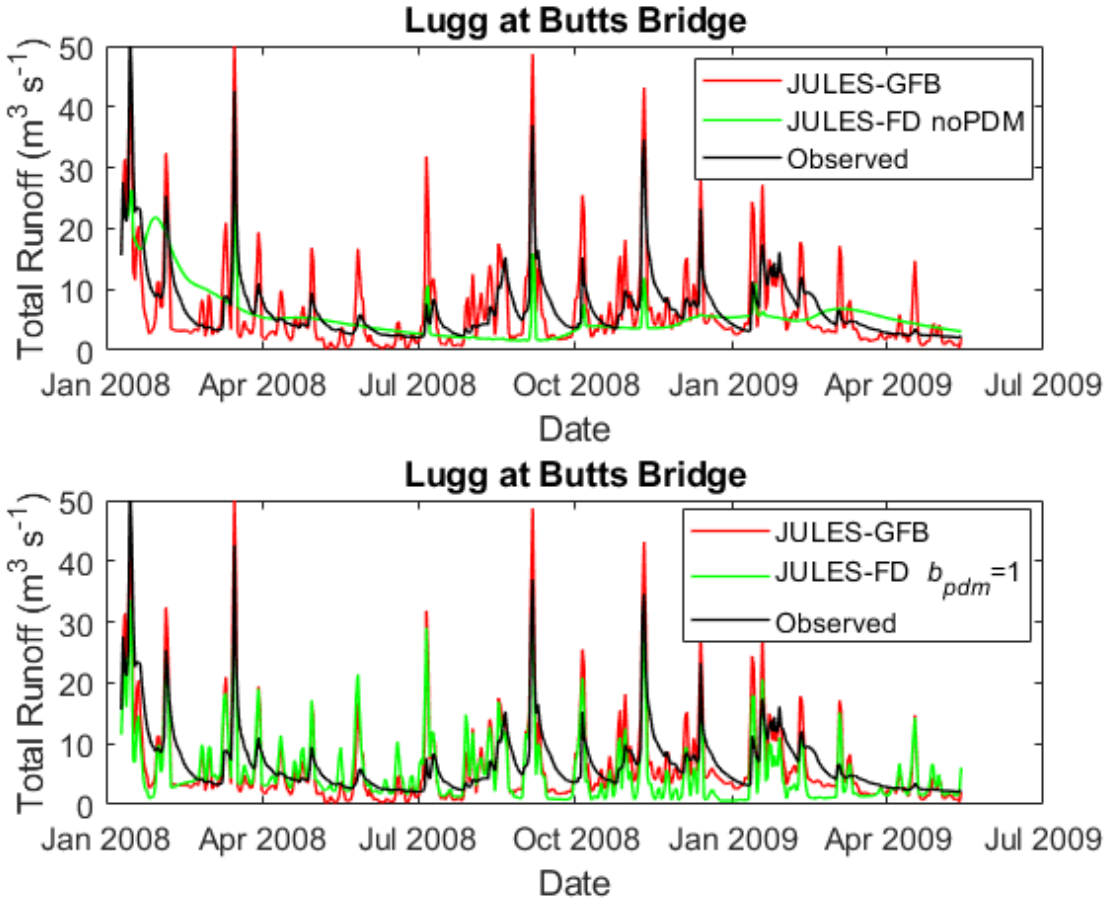


Figure 5.10: Comparison of simulated runoff time-series from JULES-GFB (red), and JULES-FD-noPDM (green) against observed runoff for the Lugg at Butts Bridge catchment for a subset of the simulation period. The difference between top and bottom figure relates to the version of JULES-FD used: JULES-FD-noPDM is shown in the top panel while and JULES-FD with PDM parameter $b_{pdm}$ set to 1.

The overall performance of the different JULES versions in reproducing total runoff at the selected catchments within the regional domain is summarized in Figure 5.11 using multiple performance metrics. In this case, the comparison is carried out against observations for three distinct model versions: JULES-FD without PDM, JULES-FD with PDM testing several values for $b_{pdm}$, and JULES-GFB. The decision to show JULES-FD with PDM with multiple $b_{pdm}$ parameters is to highlight (1) the overall uncertainty range associated with the feasible range of $b_{pdm}$ (e.g., if the model is applied at a data-poor catchment without calibration), and (2) if a calibration step (despite being depicted very simply here) can be achieved (e.g., in catchments where supporting observations can be used for calibration). Our results are focused on three metrics: (1) KGE which aims to show the overall model performance using a typical performance metric adopted in streamflow statistics, (2) linear correlation which will focus on the ability to reproduce the dynamics of the observations, and (3) mean bias as an indicator of systematic underestimation or overestimation by the different models. Notice that these metrics are intrinsic relates (Gupta et al., 2009) but there are benefits in studying them separately.

In all catchments investigated within the domain, JULES-GFB performs best (with the exception for two catchments with respect to KGE metrics) (Figure 5.11). In addition, JULES-FD without PDM (i.e., without any calibration attempt) has the overall worst performance across all six catchments and metrics. KGE values for JULES-FD-no-PDM are low around 0.20-0.30 range (Figure 5.11a). A possible calibration step taken with JULES-FD using PDM suggests improvements are achievable and can reach KGE values up to 0.5-0.70 range at different catchments. The KGE values obtained by JULES-GFB have similar moderate values compared to JULES-FD with PDM (i.e., around 0.5) with the top KGE values around 0.80. Looking in the literature if these KGE values are within the accepted boundaries, we find different limits. Gutenson et al. (2019) classified KGE as acceptable for KGE between 0 and 0.4, as good for KGE between 0.4 and 0.7 and very good for KGE values above 0.7, following Tavakoly et al. (2017). Considering the above classifications, we could say that JULES-GFB is very good for four out of six catchments and acceptable or good for the other two catchments.

In relation to reproducing the streamflow dynamics from observations, JULES-GFB shows the highest computed linear correlation coefficients at all catchments ranging from 0.60 to 0.80 (Figure 5.11b). The “best” JULES-FD with PDM version shows slightly lower coefficients around 0.55-0.75 while the linear correlation coefficients obtained with JULES-FD without PDM are around 0.50. Finally, an important aspect in understanding model performance is to

be able to identify any systematic conditions for over- or under-predicting hydrological fluxes such as the runoff. Systematic biases vary from catchment to catchment (Figure 5.11c), however JULES-GFB is the only model version which consistently predicts streamflow with lowest biases when compared to JULES-FD without PDM which appears to highly underestimate river flows (except for one catchment with the lowest BFI). If JULES-FD utilizes the PDM parameterization, biases are reduced across the six catchments but not to the same magnitude as seen with JULES-GFB. In summary, Figure 5.11 indicates that JULES-GFB is capable of predicting river flows satisfactorily (high KGE and NSE) with realistic dynamics (high linear correlation coefficients) and low systematic biases with respect to observations and when compared with the other JULES versions without explicitly representing groundwater processes.

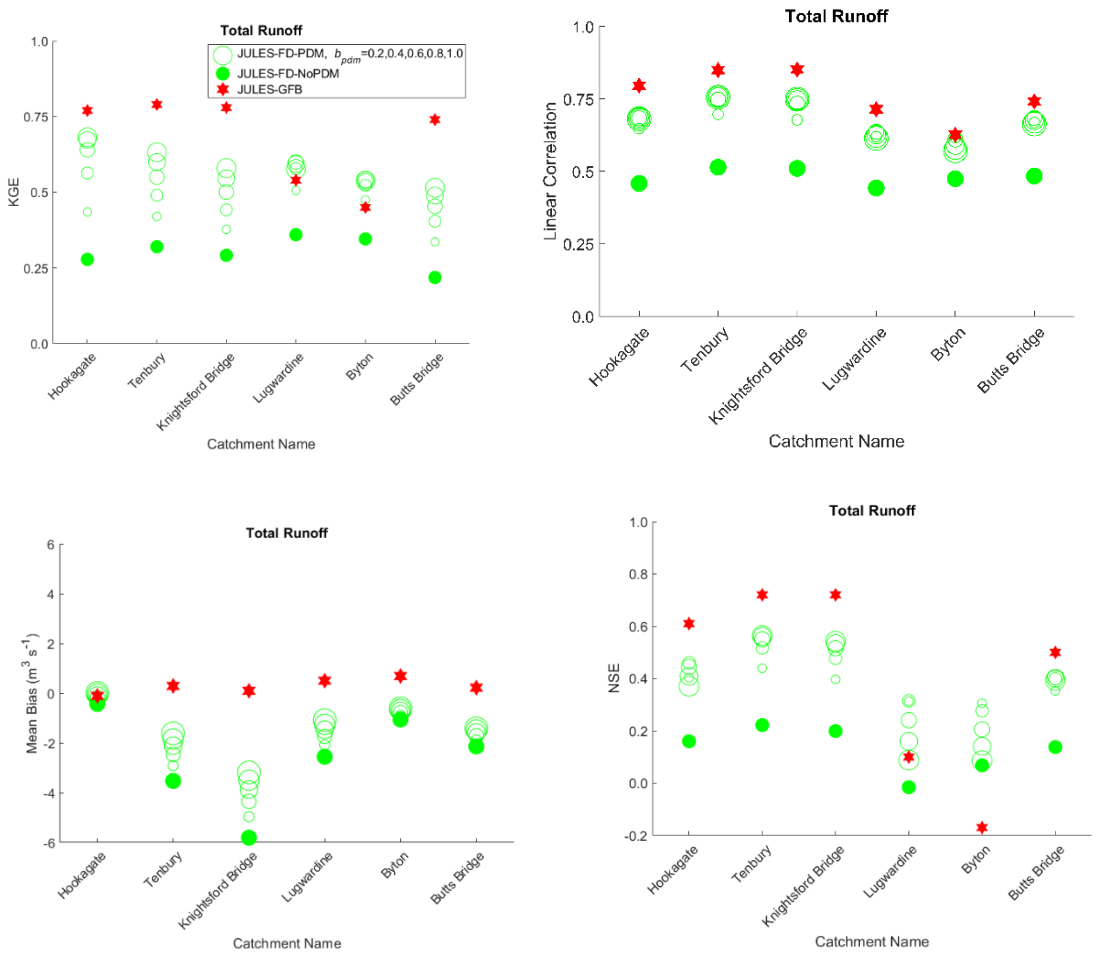


Figure 5.11: Model performance summary across four selected metrics: KGE (top left), linear correlation (top right), Mean Bias (bottom left), and NSE (bottom right) computed from simulated total runoff against observations. The metrics obtained for JULES-FD without PDM are shown as the closed

green circles while the open circles correspond to results obtained from JULES-FD with PDM with varying $b_{pdm}$ (relative to size of the circle); JULES-GFB metrics are shown as red stars.

## 5.4 Discussion

We have organized this section to answer the two specific research questions highlighted in the Introduction (Section 1.2). First, we discuss the direct impacts of introducing this new groundwater parameterization in JULES focusing on understanding the conditions which result in more or less pronounced improvements in the simulated soil water dynamics (Section 5.4.1). Then, we discuss the overall impact of such implementation on other hydrological fluxes that are related to soil water dynamics (Section 5.4.2).

### 5.4.1 Direct impact of new groundwater scheme on simulated soil water dynamics

We examined the behaviour of the JULES model under different conditions using the classic free drainage approach and with our newly proposed groundwater parameterization, respectively. These distinct conditions are tested by altering two physical controlling factors (soil type and position of water table) that are expected to exert some impact on simulated water dynamics in the soil (e.g., Niu et al., 2007; Rahman et al., 2019). From the column scale experiment (Section 5.2.1), we can highlight two significant findings which are more directly summarized in Figure 5.5. First, we found substantial improvements in representing groundwater table dynamics when explicitly incorporating groundwater processes into JULES for all soil types when the water table is initialized closer to the surface and within the soil domain (i.e., from surface to 3m deep). In this case, fine soil types tend to result in slightly better improvements when compared to the simulations from relatively coarse soils. This behaviour is expected given the nature and assumptions made with the proposed groundwater parameterization. On the other hand, the classic free drainage approach simulates systematic drier conditions for all soil types because it promotes relatively faster dynamics that are solely controlled by the gravity flow downwards without any physical constraint at the bottom of the soil domain. Shallow water regions account for 22-32% of the land globally (Fan et al., 2013), therefore we expect that the limitations of the free drainage can lead to substantial impact over large regions in global applications when using the default JULES model, those being for hydrological services directly or to account for the interactions between the subsurface-surface-atmosphere.

Our second interesting result relates to the fact that despite promoting substantial improvements in representing soil water dynamics when the water table is initialized at relatively shallow depths, we found little differences between the classic free drainage and the new groundwater scheme for those cases where the initial water table is located relatively deeper and outside of the soil domain (below 3m). In this case, the complexity of adding a new groundwater scheme into JULES may not be fully justified if the purpose of the model application is to track or depend only on knowing the soil moisture conditions within the soil domain. This is because both JULES-FD and JULES-GFB results agree satisfactorily with the benchmark model used in this study. Studies in the past have suggested groundwater may not exert strong controls to the near-surface soil dynamics when it is far from the surface (Maxwell and Kollet, 2008; Ferguson and Maxwell, 2010; Lo and Famiglietti, 2010). However, the use of JULES-GFB even for deep aquifer systems may be justified if the purpose is to use JULES applied to water resources studies where a representative volume of subsurface water is a key component of the study (Fan et al., 2013 and de Graaf et al., 2017).

The purpose of this study is to introduce this new approach by combining a series of synthetic and real-case experiments as recently supported by the hydrological community (e.g., Lee and Chang, 2005; Sulis et al., 2010; Kim et al., 2012; Clark et al., 2015). It is beyond the scope of this study to test the JULES model using the single column experiment under several different combinations of factors. The experiments we have selected allowed us to select key factors (soil type and initial water table depth) in order to compare directly against previously published studies (Kollet et al., 2017; Rahman et al., 2019). In doing so, we have limited the number of potential combinations to test, for example, by not investigating the role of rainfall intensity and duration on the model simulations. This was partially mitigated by introducing a real-case experiment in a region with groundwater-dominated catchments in addition to the synthetic cases.

### 5.4.2 Indirect impact of new groundwater scheme on other hydrological fluxes

Implementing the new groundwater parameterization into JULES has led to changes in soil water dynamics, especially when the water table is initialized closer to the surface and within the model soil domain. It is also important to understand how those changes in soil moisture dynamics affect other hydrologically relevant fluxes. Here, we focus on two specific fluxes, namely the total runoff and evapotranspiration. We first investigated changes in simulated total runoff with a synthetic catchment by comparing our results against the benchmark model, as well as other integrated hydrological models (Figure 5.6 and Table 5-3). In summary, we found

that JULES-GFB can reproduce the characteristics of multiple storage components and the outlet discharge satisfactorily. Performance metrics from JULES-GFB computed against the benchmark model are well within the range obtained with all other models (Table 5-3). The exception being only the fact that JULES-GFB had the lowest mean bias for outlet discharge (-1.3 $m^3 h^{-1}$) than the lowest value obtained with the other evaluated models (-0.8 $m^3 h^{-1}$). This suggests that JULES-GFB underestimated the discharge when compared to all other IH-MIP models (Figure 5.6c), possibly due to relative higher water retention in its soil domain. Nevertheless, the fact that the JULES-GFB model incorporates a mechanism to deal with lateral redistribution of water resulted in the displacement of water and ultimately leading to discharge of the outlet. This mechanism is non-existent in JULES-FD, hence only vertical water flow is promoted at individual columns (grid points) in the synthetic catchment which results in no occurrence of discharged (note: discharge occurs solely by groundwater contribution, as there is no net rainfall forcing imposed at the surface).

In addition to the simulations with the exact specifications from Kollet et al. (2017), we also ran the tilted-V catchment for the 12 different soil types using the same hydraulic properties with the infiltration column-scale experiment. Within the simulation time, only three out of 12 soil types could produce runoff, namely the sand, loamy sand and the silt loam. These three soil types have high value of saturated conductivity; thus, they react fast to the groundwater contribution by producing runoff in the outlet. It seems that the higher value of saturated conductivity gives faster and greater peak discharge. It is something we expected, as the timing is linked with the speed of groundwater contribution (saturated conductivity) and the magnitude of discharge is linked with the storage of the aquifer (specific yield). Soil types with lower saturated conductivity values respond much slower to the groundwater table, so they cannot produce runoff within the simulation time (5 days). This proves that sandy soil types are more sensitive to the groundwater parameterization compared to other soil types. Considering the saturated and unsaturated storage curves for different soil types, we find that sand, silt loam and the simulation from the tilted-V experiment (Section 5.3.2) have linear curves with small fluctuations that show that they change slowly; whereas, the silt loam soil type has a more responsive behaviour. It produces runoff only for a few hours, whereas the response of saturated and unsaturated storages shows a dynamic response.

Synthetic experiments have the benefit of allowing nearly full control of the results by comparing model simulations against a benchmark (reference) model taken to be the truth. These experiments allow us to isolate relevant processes more easily for analysis usually

representing ideal conditions, which do not necessarily correspond to reality. In order to get a better sense of how our proposed JULES-GFB perform under more realistic conditions, we performed a 5-year simulation over a regional domain in the UK characterized by groundwater-dominated catchments. We investigated the ability of JULES-GFB to accurately represent the spatial patterns of wetness associated with the river network in the region (Figure 5.7), as well as potential improvements to simulated evapotranspiration (Figure 5.9) and river discharge (Figure 5.11) when compared against independent observations.

Our initial analysis within the domain highlights the importance of groundwater process and more specifically the role of lateral flow now being resolved in JULES-GFB in order to allow more realistic patterns of relative saturation within the domain (Figure 5.7). These patterns are consistent with two key characteristics of the region, namely topography and soil type. This is consistent to our initial findings using the synthetic experiments highlighting the various model simulations for different soil types, initial water table depths, and with a synthetic catchment whose processes are driven entirely by groundwater dynamics (i.e., no net rainfall forcing). Unlike JULES-GFB, the JULES versions with the classic free drainage are unable to represent such patterns simply because they do not resolve lateral flow and hence all soil water dynamics happen vertically, meaning water eventually leaves the soil domain from the model without allowing for the low elevation areas to properly show convergent patterns. The more realistic water dynamics and ability to retain the water within the soil domain due to a newly imposed bottom boundary condition in the soil domain was also reported by (Niu et al., 2007).

The analysis from Figure 5.7 allows us to have some confidence in the model performance. However, they do not directly quantify potential improvement in model performance. Therefore, we further investigate whether this arguably more realistic behaviour of the model results in improved evapotranspiration (ET) fluxes. As expected in this humid but energy limited region, the impacts of introducing the new JULES-GFB scheme indicates more pronounced changes during the summer season ET (Figure 5.9) with slightly increase in simulated ET compared to JULES-FD resulting in slightly more consistent estimates when compared with the GLEAM ET product. Our overall results also indicate that simulated ET from JULES-GFB shows reduced biases compared to JULES-FD which tends to underestimate ET fluxes. This behaviour is consistent across the main soil types observed in the regional domain. Similar results were also previously suggested by a number of studies (York et al., 2002; Yeh and Eltahir, 2005; Niu et al., 2007; Nie et al., 2018; Huang et al., 2019) by

comparing against observations from the water balance equation, ground based measurements and remote sensing products.

Finally, we look at the results summarizing JULES-GFB's ability to reproduce river discharge in comparison with discharge observations at six catchments within the domain (all strongly dominated by groundwater processes, given their relatively high Base Flow Indices – BFI). We highlight the performance of the model by (1) employing a metric typically used for evaluating model performance of streamflow predictions (KGE), (2) analysing the ability of the model to reproduce similar dynamics from the observed record (linear correlation coefficient), and (3) identifying any unwanted systematic biases in the model (mean bias). Particularly for JULES-FD, we also tested an additional calibration step (which is not applied to JULES-GFB) by enabling the PDM scheme in the model that empirically tries to resolve any limitations related to soil saturation distribution at subgrid spatial scales. In summary, our results suggest that JULES-GFB performance is superior for nearly all metrics and catchments for total runoff (Figure 5.11). The majority of KGE values obtained are above 0.75 which suggest satisfactory performance whereas JULES-FD without PDM shows KGE values on the order of 0.25. Similar conclusions about the performance of JULES-GFB in comparison to JULES-FD without PDM can be drawn for the linear correlation coefficient and the mean bias. Previous studies reported improvements on the same order of magnitude (Maxwell and Miller, 2005; Yeh and Eltahir, 2005; Vergnes and Decharme, 2012; York et al., 2012). Interestingly, if JULES-FD is applied with a calibrated version of PDM, its performance improves substantially, but still does not exceed the performance of our JULES-GFB. We argue, however, that any additional calibration step introduced in this case will likely require a priori estimation of the parameters associated with PDM. This will likely require the availability of independent observations used for the calibration period, hence undermining the performance of JULES-FD particularly for hydrological applications in data-poor regions and eventually continental to global applications. It is important to stress that despite the observed improvements with JULES-FD with PDM, this version of the model still lacks the explicitly representation of groundwater processes.

Note that our approach of groundwater parameterization is different from other previous studies. For instance, the developed approach by Niu et al. (2007) and Huang et al. (2019) extracts the baseflow from the aquifer, as a percentage of the storage or the water table depth using an exponential function. In our study the flux of baseflow comes from the contribution of aquifer to the soil domain by an additional upward flux component. Other models employ an uncoupled representation of groundwater (e.g., de Graaf et al., 2015, 2017; Fan and Miguez-

Macho, 2011). According to this approach, the output of the groundwater model feeds the bottom boundary condition of the LSM. Both models run separately, and, in this way, there is not a real time feedback between saturated and unsaturated zones, whereas the groundwater model is not dynamically updated. In our case, we implement a dynamic model that can give us a more accurate insight for the surface subsurface interactions. Furthermore, our model adopts a simplified approach to represent groundwater in two dimensions. Three-dimensional groundwater coupling, such as Maxwell et al. (2011) and Tian et al. (2012), require substantial computational resources, especially for large-scale applications. Our approach is extensively evaluated against similarly more complex 3D hydrological models to ensure much of its realism is maintained during the development phase, while identifying potential limitations and shortcomings.

This approach and this methodology have some limitations. Firstly, we tested the model over some regional domain characterized by groundwater dominated catchments. We recognize that the region is still relatively small compared to potential applications of the model at continental or global scale. Furthermore, the selected regional domain has climatic conditions known to have weak interactions with the atmosphere (i.e., weak soil moisture-ET interactions). Regarding the characteristics of the new approach, our simplified 2-D approach has limitations compared to the more realistic 3-D approach. In order to examine the magnitude of this limitation, both the theoretical development study (Rahman et al., 2019) and this study employed synthetic experiments to evaluate how much this approach impacts the overall realism of the simulation by benchmarking our simulations with a 3-D hydrological model ParFlow. We did not find significant differences between the 2-D (JULES-GFB) and the 3-D (ParFlow) approach that proves the realism of our approach. Finally, despite these limitations, our study suggests substantial improvement in the JULES model, which is a typical LSM that currently does not represent soil-aquifer interactions in its operational version.

## 5.5 Conclusions

In this Chapter, we incorporate a simplified groundwater representation in the Joint UK Land Environment Simulator (JULES) and investigate the impacts of this implementation on land surface hydrological processes. We consider two synthetic (a single-column and a tilted-V) and one real-world (regional domain including six catchments in the UK) test cases to demonstrate the impacts of representing groundwater explicitly in JULES-GFB. The performance of JULES-GFB is evaluated by considering a three-dimensional groundwater flow model ParFlow

as the benchmark for the synthetic test cases. For the real-world test case, observed runoff and evapotranspiration data are used for the model evaluation.

Results from the synthetic test cases demonstrate that JULES-GFB improves soil moisture dynamics and runoff generation process compared to the default JULES-FD, especially for fine-textured soils. The real-world test case demonstrates that JULES-GFB improves the prediction of runoff and evapotranspiration compared to JULES-FD. From this test case, it can also be concluded that representing groundwater hydrology explicitly can supersede the advantage of implementing a calibrated saturation excess runoff generation scheme in JULES.

JULES-GFB shows some limitations in reproducing soil moisture dynamics and runoff generation compared to the benchmark model for coarse-textured soils. This is expected because of the simplifying assumptions considered for coupling the groundwater parametrization with JULES. These assumptions are necessary to achieve a high computational efficiency of the model (Rahman et al., 2019). In this regard, our proposed approach of representing groundwater hydrology aligns with the objectives of the recent development efforts of Hydro-JULES (Dadson et al., 2019) and other land surface and hydrological models for large-scale applications (Clark et al., 2015; Gleeson et al., 2019).

The real-world test case presented here is limited in term of hydrogeological and climatic conditions, because this is the first study that evaluates the newly developed JULES-GFB model. This model requires further testing under different hydrogeological and climatic conditions, which should be the subject of further research. Subsequently, JULES-GFB could potentially be used as a numerical tool to assess water resources under e.g., future climate change and land use/cover change scenarios.

# 6 REGIONAL CLIMATE PROJECTIONS BASED ON UKCP18 IN CENTRAL GREAT BRITAIN

## 6.1 Introduction

Groundwater is an important component of water resources in United Kingdom with about 30% of public water to be supplied by groundwater (Environment Agency, 2006). Thus, it is crucial to know the impact of climate change on groundwater dominated catchments. It is known that climate change will affect recharge (e.g., Scibek and Allen, 2006; Konikow, 2011; Cuthbert et al., 2019). However, there is limited knowledge on how climate change will affect all the water components on groundwater dominated catchments, like soil moisture, evaporation, because studies usually focus only on runoff and recharge. In this Chapter, we assess the impact of groundwater representation on climate predictions. A common approach to assess the impact of climate change is the implementation of an LSM that rely on the traditional free drainage approach (e.g., Hagemann et al., 2013; Dankers et al., 2014; Huijgevoort et al., 2014; Giuntoli et al., 2015; Vetter et al., 2017). JULES has also been used in climate change studies (e.g., Betts et al., 2018; Koutroulis, 2019; Koutroulis et al., 2019). However, we showed in the two previous Chapters, that the free drainage assumption can alert the signal of climate change, while it exacerbates drying and underestimates the evapotranspiration flux. Furthermore, in groundwater dominated catchments, JULES with the free drainage shows substantial shortcomings in representing runoff, especially during the dry season, when the baseflow has higher contribution. Now that we have developed a new version for the JULES model (Chapter 5), which incorporates groundwater and impacts other components of the water balance at the surface, we will apply this to future climate change

scenarios in the same region. Thus, the new model can increase our understanding about the impact of climate change on the hydrological processes, while the previous version with the free drainage case created bias in the model predictions. Furthermore, the new groundwater parameterization on JULES can help us to identify the impact of climate change on surface-subsurface interactions, as it represents dynamically the feedback between the soil domain and aquifer.

UK Climate Projections (UKCP) is a climate analysis tool that forms part of the Met Office Hadley Centre Climate Programme (www.metoffice.gov.uk, accessed here https://catalogue.ceda.ac.uk/?q=ukcp18&results_per_page=5&sort_by=relevance last visited on 04/09/2020). The previous version, named UKCP09, was based on the 4th Assessment Report from the Intergovernmental Panel on Climate Change (IPCC) and it used the Special Report on Emissions Scenarios (SRES) (Fung and Gawith, 2018). These products have gained a lot of attention to predict the future (e.g., Cloke et al., 2010; Bell et al., 2012; Coulthard et al., 2012; Christierson et al., 2012; Kay and Jones, 2012). The new ensemble data of the UK Climate Projections 2018 (UKCP18) reflect the current information and the current hypotheses for the climate change projection based on the 5th Assessment Report from IPCC. This report follows the Representative Concentration Pathway (RCP) scenarios, which is a greenhouse gas concentration trajectory. Four scenarios are adopted based on the amount of greenhouse gases that will be emitted in the future, namely RCP 2.6, RCP 4.5, RCP 6 and RCP 8.5. In RCP 2.6 scenario, it is assumed that the global annual greenhouse gas emissions (measured in $CO_2$-equivalents) will peak between 2010-2020 and then, the emissions will decline substantially. In RCP 4.5 scenario, the peak will occur in 2040s, and then, the emissions will decrease. In RCP 6, the peak will occur around 2080, and then, the emissions will decrease. Finally, for RCP 8.5 scenario, the emissions will continue to increase during the 21st century (Meinshausen et al., 2011). UKCP18 provides a suite of datasets. In a nutshell, this product provides probabilistic projections for all emission scenarios, global projections for RCP 8.5 emission scenario, regional projections for RCP 8.5 emission scenario and for finer spatial resolution, and derived projections, which is a set of climate futures derived from global projections, for a lower emission scenario (Fung et al., 2018). Refer to Fung et al. (2018, Table 1) for more information.

The literature review of using UKCP18 dataset in peer-reviewed papers is restricted compared to the previous UKCP09 product, because only recently this dataset was published. There are two reports that focus on Great Britain and are based on UKCP18 projections. According to UKCP18 overview (Met Office, 2018a, b) for the UK, it is expected the rainfall

to increase in the winter and decrease in the summer. However, more intense storms are expected during the summer. The range of the mean annual temperature increase, compared to 1980 in Central England, is between no change to +3.3 ºC and +1.1 to +5.8 ºC, for the low and high emission scenario, respectively. Hot summers, like 2018, will occur more likely, with probability about 50%. South and Southeast part of UK are more likely to experience changes.

Our motivation in this Chapter is to understand the impact of climate change on hydrological fluxes (soil moisture, runoff, recharge, ET) on groundwater-dominated catchments using the newly developed JULES model with groundwater parameterization. The study area covers a part of Central England and a part of Wales in Great Britain (See Section 5.2.3). Here, we address two main research questions:

(1) What is the impact of future climate change scenarios on individual water components of water cycle at the surface? Which key meteorological forcing variable (i.e., temperature or precipitation) has the greatest impact on each water component?
(2) Are there particular combinations of climate scenarios that result in severe extreme weather?

## 6.2 Setting up climate scenarios impact study

### 6.2.1 UK Climate Projection 2018 dataset (UCKP18)

We use the UK Climate Projections 2018 (UKCP18) dataset to simulate the climate projection and assess the impact of climate change on groundwater dominated catchments. In the following subsections we describe this dataset and discuss the methods we apply for data conversion to JULES requirements, bias correction, and the sample selection procedure. The UKCP18 suite of datasets is freely available at the following website: https://catalogue.ceda.ac.uk/?q=ukcp18&results_per_page=5&sort_by=relevance&record_types=ObservationCollection (last visited on 04/09/2020). We use the regional projections that cover Great Britain and Europe for the higher emission scenario (RCP 8.5) and for the period [1980-2080]. The regional projections are a set of 12-high resolution (12 km grid sizes and daily) climate future scenarios over Europe, downscaled from the global projections using Hadley Centre model HadREM-GARA11M (Fung and Gawith, 2018). This selection of this dataset has a few advantages and disadvantages, but in the end the advantages outweigh the disadvantages, and this dataset can give us a useful insight for the future climate and the impact on water components. Firstly, the advantages of using UKCP18 dataset are that it is compatible with the forcing data needed to

run JULES and the dataset covers United Kingdom. Furthermore, the fact that UKCP18 has been produced by Met Office and the previous product (UKCP09) has been extensively used in United Kingdom, gives credibility to UKCP18. On the other hand, the high resolution UKCP18 product is based only on RCP8.5, which represents the worst emission scenario. Furthermore, there is not an extensive literature on this product yet, except a few reports from Met Office discussed in the Introduction.

### 6.2.2 Making UKCP18 projections compatible to JULES forcing

As we discussed in Chapter 3, JULES needs eight meteorological forcing variables, namely the precipitation, wind speed, air temperature, surface downwelling longwave and shortwave radiation, specific humidity, pressure, and daily temperature range. These data come directly from UKCP18. The only difference is that UKCP18 provides with the net surface downwelling longwave and shortwave radiation. However, JULES needs the total shortwave and longwave radiation. We convert the net shortwave downwelling radiation to total shortwave downwelling radiation using the following equation:

$$SW = SWnet/(1-a) \quad (6.1)$$

where $SW$ is the total shortwave downwelling radiation (W $m^{-2}$), $SWnet$ is the net shortwave downwelling radiation (W $m^{-2}$) and $a$ is the albedo (unitless). Soil albedo is calculated based on previous study (Table 1 in Houldcroft et al., 2009) that links JULES land use type with surface albedo. The longwave downwelling radiation is calculated as follows:

$$LW = LWnet+\sigma\varepsilon_s T^4 \quad (6.2)$$

where $LW$ is the longwave downwelling radiation (W $m^{-2}$), $LWnet$ is the net longwave downwelling radiation (W $m^{-2}$), σ is the Stefan-Boltzmann constant $5.67x10^{-8}$ W $m^{-2}$ $K^{-4}$, $\varepsilon_s$ is the emissivity of the body (unitless) and $T$ is the temperature of the body (K). For the temperature of the body, it was used the air temperature coming from UKCP18 data. The emissivity values come from Shuttleworth (2012, Table 5.2). The temporal resolution of the data is daily, but this is downscaled to the computational time-step, based on the downscaling tool of JULES following the same procedure with the previous Chapters.

### 6.2.3 Bias Correction of UKCP18 forcing data

It is widely recognised that bias is present in the output from climate models (e.g., Christensen et al., 2008; Haerter et al., 2011). To address this issue, we apply bias correction during the period when the two datasets (i.e., current, and future conditions) overlap. Bias

correction is a common method to correct climate projections, based on observations (Hay et al., 2000; Leander and Buishand, 2007; Hempel et al., 2013). This technique compares the observations and the projections for an overlapping period and then, applies this correction to the future projections. The most common methods are linear scaling (e.g., Hay et al., 2000; Fowler and Kilsby, 2007; Lafon et al., 2013; Guillod et al., 2018), variance scaling (e.g., Leander and Buishand, 2007), quantile mapping (Wood et al., 2004; Prudhomme et al., 2012) and trend-preserving quantile-mapping (e.g., Hempel et al., 2013). Linear scaling is a simple method, but it corrects only the mean. The variance scaling method is simple, and it retains the climate change signal. However, it follows the same variability with the observed dataset, and it is restricted to the range of the observed anomalies. The quantile mapping takes into consideration the entire distribution and can be used for cases with changes in extreme values or where the variability is important. The disadvantages of this method are that it can alter the climate change signal, it assumes that the correction increments are equal as in the current climate, whereas the extreme values are restricted to the observed ones. Finally, the trend-preserving quantile-mapping has the same advantages with the quantile mapping plus that preserves the climate change signal. However, it is very complicated with many steps and it can lead to physical inconsistency (Met Office, 2018c). On the other hand, there are many doubts about the usefulness and the reliability of bias correction. Haerter et al. (2011) note that most of the bias correction methods do not preserve the conservation laws of energy and water balance. Furthermore, the correction is usually applied separately for each variable and consequently the links and feedbacks between the meteorological states and fluxes are destroyed (Ehret et al., 2012). One important reason for criticism is the assumption of stationarity of the methods that is not true in a climate change case (Ehret et al., 2012). Hagemann et al. (2011) found that applying a bias correction method can modify the climate change signal.

Considering the variables in which researchers apply bias correction methodologies, it seems that most of the cases (e.g., Hempel et al., 2013; Sorribas et al., 2016; Zulkafli et al., 2016) tend to correct all of them (i.e., precipitation, temperature, radiation, humidity, pressure, humidity and wind), while in some cases (e.g., Hay et al., 2000; Teutschbein and Seibert, 2012) choose to correct only precipitation and temperature. For the case of the simpler linear scaling bias correction method, all the variables need a multiplicative correction, except the temperature that is corrected additively (Hempel et al., 2013, Table 1; Hutchins et al., 2018). The selection of the bias correction method should be based on the statistical properties that someone wants

to correct (Teutschbein and Seibert, 2012; Lafon et al., 2013). To understand the magnitude of bias, we analyse the UKCP18 dataset against the CHESS dataset for an overlapping period between 1981 and 2015. We apply the comparison between the two products for the domain average. We consider CHESS as the benchmark against the UKCP18 product. We apply bias correction to all the variables following most of the studies (e.g., Hempel et al., 2013; Sorribas et al., 2016; Zulkafli et al., 2016).

We decided to use the linear scaling bias correction method (e.g., Hay et al., 2000; Fowler and Kilsby, 2007) because we want to correct the mean without changing the other moments. We are aware that this method will not correct higher-order moments of rainfall distribution, like the frequency of wet and dry days. However, we believe that this methodology is sufficient for our study. As we discussed above, all the variables are corrected multiplicatively (Equation 6.3), except temperature that will be corrected additionally (Equation 6.4). The controlled period based on which we apply the bias correction is 35 years [1981-2015]. The correction is based on the following equations:

$$X'(t) = \frac{\overline{O(t)}}{\overline{X(t)}} X(t) \tag{6.3}$$

$$X'(t) = \left(\overline{O(t)} - \overline{X(t)}\right) + X(t) \tag{6.4}$$

where $X'(t)$ is one of the twelve unbiased UKCP18 ensembles, $X(t)$ is one of the twelve biased UKCP18 ensembles, $\overline{X(t)}$ is the mean of the ensembles for a specific month $t$ and $\overline{O(t)}$ is the mean for a specific month of CHESS dataset.

Below, we present the results of bias correction on UKCP18 dataset [1981-2080] based on the CHESS dataset [1981-2015]. We apply equations 6.3 and 6.4 to individually correct biases for each ensemble/variable combination; however, in order to evaluate the bias correction approach, we use a common Percentage bias (PBias) metric:

$$PBias = \frac{\sum_{i=1}^{i=12}(\overline{UKCP_i} - CHESS_i)}{\sum_{i=1}^{i=12} CHESS_i} \tag{6.5}$$

where $n$ is the month (from January to December) and $\overline{UKCP}$ is the average ensemble value for UKCP. Positive (negative) PBias means that the UKCP18 overestimates (underestimates) a particular variable. To understand the magnitude of bias, we plot the UKCP18 dataset against the CHESS dataset for the overlapping period between 1981 and 2015 for all the variables. We want to analyse the seasonality of the bias to see during which period the bias is higher.

Furthermore, we compare the mean monthly values before and after the bias correction to examine the accuracy of the bias correction method and if our approach can eliminate the bias.

**Temperature**

Figure 6.1 compares the monthly values of the domain averaged temperature for the period [1981-2015] produced by CHESS (black line) against UKCP18 ensembles (red lines). The first observation is that we cannot see a strong bias in that variable. CHESS could be one of the ensembles, as there is not a clear distinction between the two products. Looking Figure 6.1 (left), we could infer that UKCP18 ensembles are below CHESS during the spring months (March to May). Furthermore, in period between September and November, CHESS is in the upper percentile, that means that UKCP18 dataset tends to underestimate these months. Applying the Kolmogorov-Smirnoff test (K-S test) to the monthly datasets, we found that six out of twelve ensembles follow the same distribution with CHESS. Figure 6.1 (right) depicts the comparison of UKCP18 with CHESS after the bias correction. CHESS is lying in the centre of the UCKP18 ensembles. The PBias is -5.3% before the correction, but the bias correction method can reduce the PBias to zero. Furthermore, applying the K-S test in the bias-corrected projections, we found that nine out of twelve ensembles are draw from the same probability distribution compared to CHESS. Lower projected temperature for Match to May and from September to November could probably lead to less evaporation from soil, thus higher projected runoff, and wetter soil.

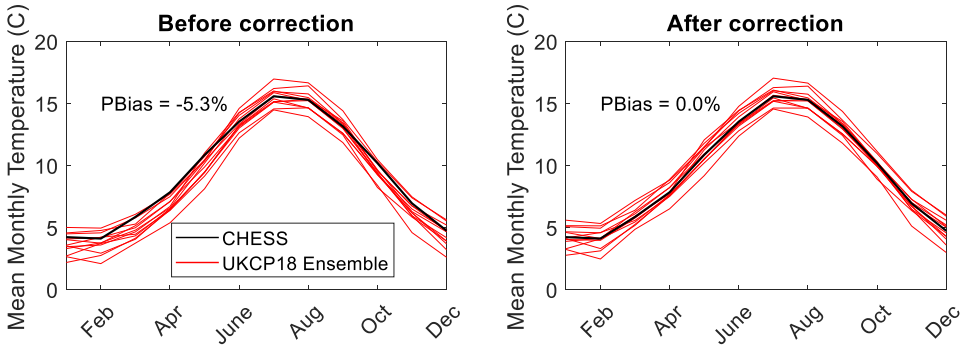


Figure 6.1: Comparison of monthly mean of domain average Temperature produced by CHESS (black) against UKCP18 ensembles (red) plotted before (left) and after (right) bias correction.

**Precipitation**

The mean annual precipitation from CHESS is found equal to 886 mm year$^{-1}$, whereas the corresponding values from UKCP18 dataset range between 820 and 1080 mm year$^{-1}$. This means that the annual rainfall from the UKCP18 product can be up to 22% higher compared to the CHESS dataset over the same period [1981-2015]. Figure 6.2 depicts the domain average of total monthly rainfall of UKCP18 ensembles against CHESS between 1981 and 2015. The first observation is that we cannot identify a very strong bias, except in Spring, looking the figure from a qualitative point of view. CHESS could be, again, one of the ensembles without significant distinction from the other members. However, UKCP18 shows a more seasonal behaviour compared to CHESS. Also, UKCP18 shows minimum on August, whereas CHESS has homogeneous low values from February to September. We can see that for the period between February and May, UKCP18 overestimates the rainfall, whereas we find the opposite behaviour for September and October. Therefore, not applying bias correction would distort the results of the water cycle during the year. Applying the K-S test to the monthly datasets, we found that 3 out of twelve ensembles follow the same distribution with CHESS. After the bias correction, seven out of twelve ensembles follow the same distribution with CHESS. The PBias is reduced to zero after the bias correction as compared with its initial value of ~9% before the correction. Interestingly, we correct not only bias, but the seasonality, as well.

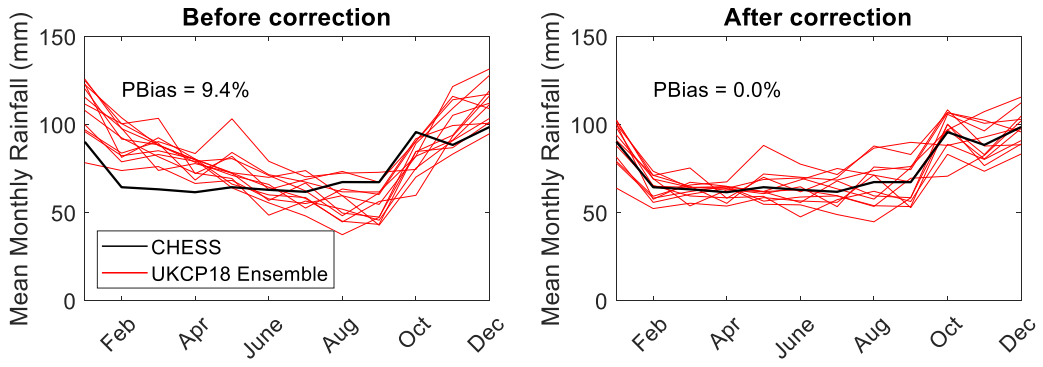


Figure 6.2: Comparison of monthly mean of domain average Precipitation produced by CHESS (black) against UKCP18 ensembles (red) plotted before (left) and after (right) bias correction.

**Additional meteorological forcing**

Figure 6.3a compares shortwave downwelling radiation from UKCP18 and CHESS datasets. This figure demonstrates maximum radiation value in June for CHESS and in July for UKCP18. It seems that there is one-month lag between the two products, but the magnitude of the two products is similar. Although the PBias is not high (-2.3%), the lag behaviour could affect the dynamics of evapotranspiration by changing its seasonality. In that case, bias correction is vital to simulate the dynamics and the seasonality correctly. Bias correction reduces PBias from -2.3% to -0.4%, while improving the seasonality. With regards to longwave downwelling radiation (Figure 6.3b), note the UKCP18 longwave radiation is calculated from the net longwave downwelling radiation, based on Equation 6.2. In that case, there is not a lag behaviour and the dynamics look similar, but it seems that there is an offset between the two products. Bias-correction achieves to remove this offset by reducing the PBias (9.2%) to zero.

Figure 6.3c presents the monthly mean of specific humidity. It is apparent the one-month lag of UKCP18 dataset compared to CHESS. Considering only when the maximum occurs, then it seems that for CHESS this happens on August, whereas for UKCP18 there are few ensembles in which the maximum occurs in September. However, the one-month lag is clear on the shape of the ensembles compared to CHESS. Again, bias correction is critical to simulate the dynamics and the seasonality accurately. Ignoring the bias of humidity (-3.4%) could possibly alert the seasonality of evapotranspiration. Regarding the surface pressure (Figure 6.3d), there is a distinctive offset between the two datasets, but the dynamics look similar, with the higher values for both datasets to occur in summer season and the lower between October and May. By applying the bias correction method, we remove the offset, and we create a more consistent dataset. Bias correction removes the PBias (2.5%) to zero. Finally, we analyse the wind speed (Figure 6.3e). In this case, there is not an obvious bias between the two products, whereas the seasonality is similar with higher values during winter and lower during summer. CHESS is inside the ensemble envelop and could be a member of the ensembles without distinction. However, looking all the ensembles, it seems that most of them underestimate the wind speed giving a total PBias equal to -7.2%. The bias correction eliminates the bias with a final PBias of 0.0%.

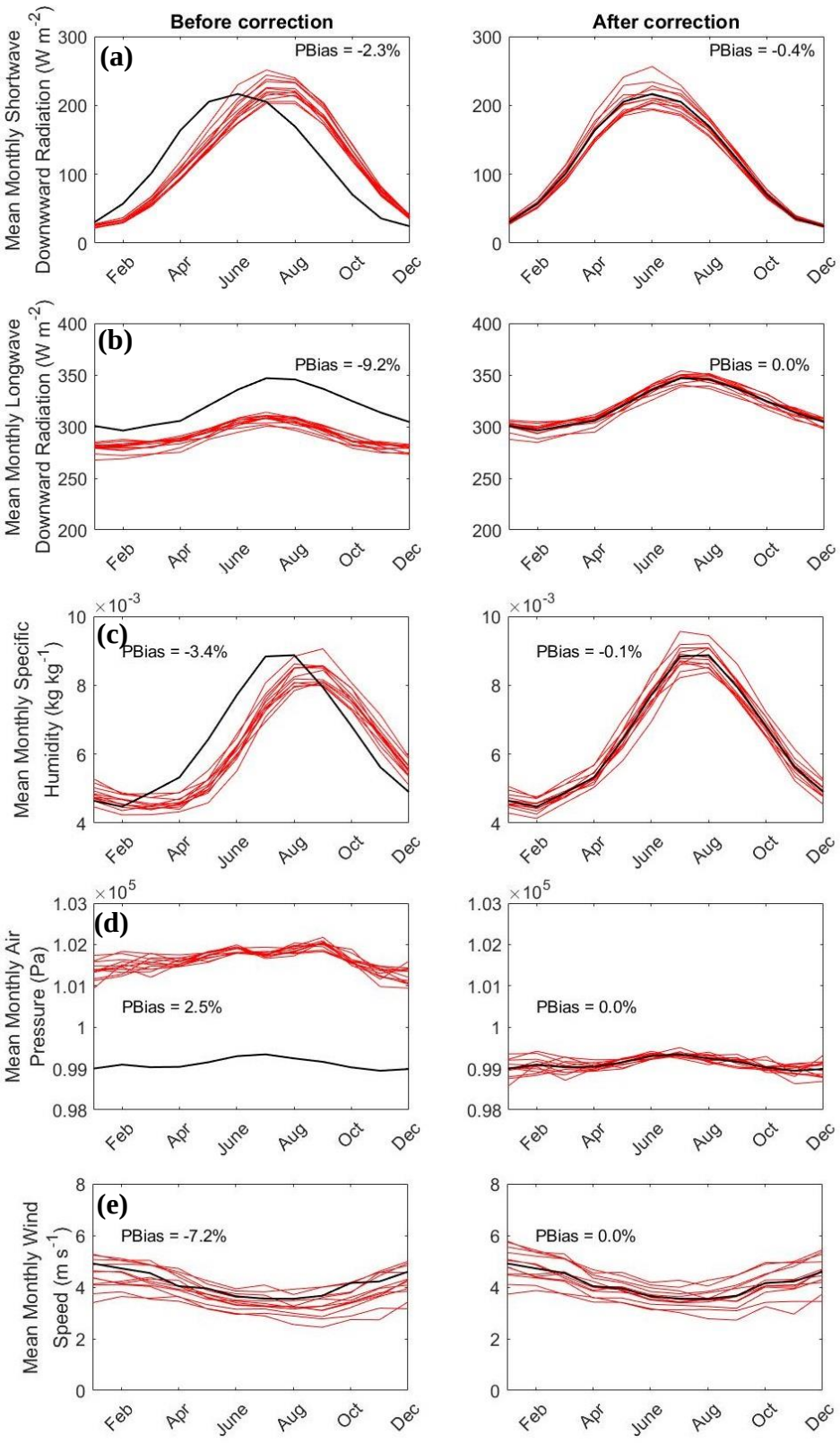


Figure 6.3: Comparison of monthly domain average of shortwave downward radiation (a), longwave downward radiation (b), specific humidity (c), air pressure (d) and wind speed (e) produced by CHESS (black) against UKCP18 ensembles (red) before (left) and after the bias correct (right).

### 6.2.4 Bias-corrected climate ensembles for Central Great Britain

This subsection presents the projections of UKCP18 for the study area. We focus only on temperature and precipitation because these are the typical hydrometeorological variables that characterize the climate of a region. However, we use all bias-corrected climate variables discussed in the previous subsection as forcing data for our model simulations. We look the annual changes and the seasonal changes during Winter (December-February) and Summer (June-August).

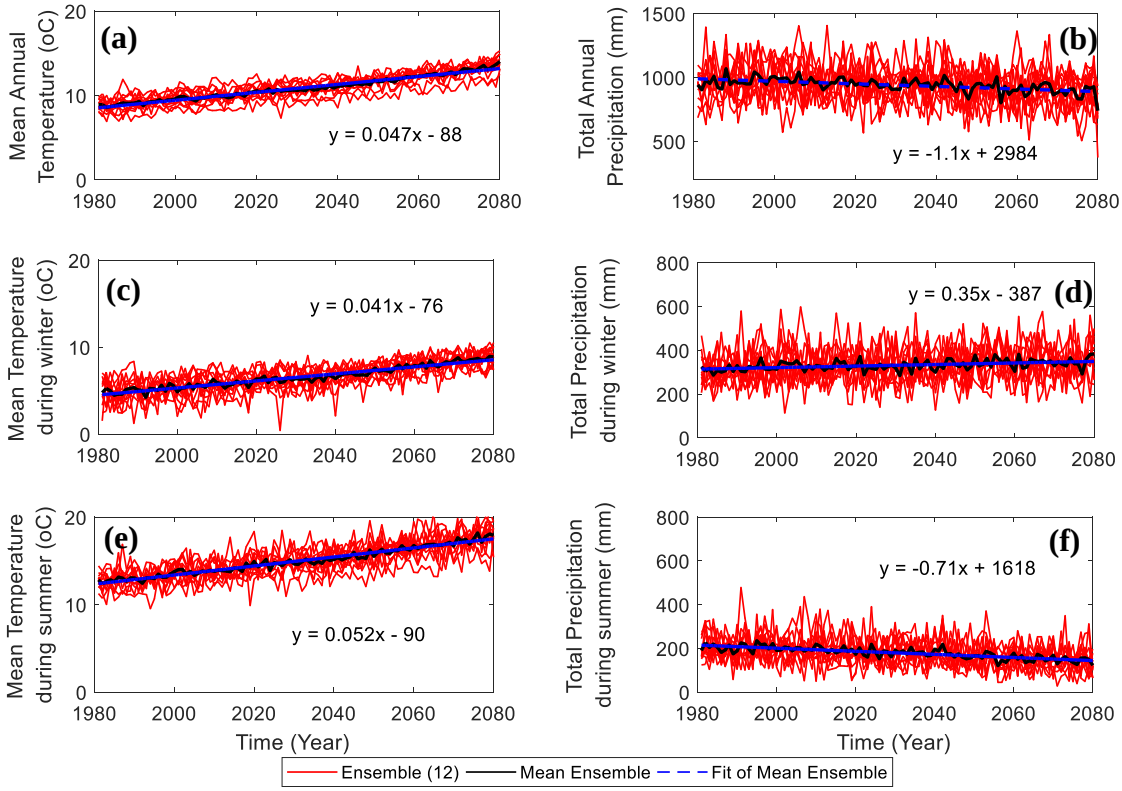


Figure 6.4: Mean projected temperature (left column) and total precipitation (right column) from UKCP ensembles (red) compared to the ensemble mean (black) for the whole year (first row), winter season (second row) and summer season (third row).

We want to identify how strongly the future climate change might affect the hydrology of the study area using the UKCP18 projections for the high emission scenario (RCP 8.5). In Figure 6.4a, the domain average for the mean annual temperature (left) and the total annual precipitation (right) after the bias correction for the twelve ensembles (red colour) and the mean of the ensemble (black colour) are plotted. The increase of the mean annual temperature, based on the mean of the ensemble, is 0.47 ºC per decade, with almost constant rate of increase. The

mean annual temperature of the ensembles in 1980 is about 8.5 ºC and it is expected to reach 13 ºC in 2080s. Regarding the corresponding value of annual precipitation, in that case it seems that the rainfall will decrease with a rate about 11 mm per decade (Figure 6.4b).

For both winter and summer seasons, UKCP18 predicts an increase in temperature. The steepest rise of temperature is expected during summer, with about 0.52 ºC per decade (Figure 6.4e). The corresponding value for winter is 0.41 ºC per decade (Figure 6.4c). The significance of slope for both variables was examined using the t-statistics test and was found significant for p-value = 0.05. It seems that there is a constant rate of increase for both seasons. The strongest change (i.e., higher slope) of precipitation occurs during the summer, with average decrease of rainfall about 7 mm per decade (Figure 6.4f). For the winter season, the rate of rainfall will be lower with average decrease of 3.5 mm per decade. Interestingly, the variability during the summer is higher than during the winter. Overall, the variability of the ensembles for rainfall is much higher from the variability of temperature, but constant throughout the 100-year period for both variables and for all the seasons.

### 6.2.5 Sampling of UKCP18 Projections

This subsection describes the methodology we follow to select samples for the climate projections. Considering the computational demand of JULES-GFB, the datasets and the literature, we decided to develop a new methodology to choose the sample simulations from all the ensembles. Choosing a few decades to represent a short-term and a long-term future is considered not very informative and subjective, because we cannot project the exact climate conditions of the future decades due to the inherent uncertainty from Global Climate models (Holman et al., 2009; Wootten et al., 2017). Thus, we use a new more objective approach that could give a better insight on the impact of climate change. We run 5-year simulations whose average precipitation and temperature are proportional to the climatology, as described by the CHESS dataset between 1980 and 2015. In that case, instead of trying to predict the impact of climate change for a specific decade in the future (e.g., 2060s), we will try to identify the effect of a specific change of annual temperature and precipitation (e.g., +2 °C of mean annual temperature and -20% of mean annual precipitation) to the hydrology of the study area.

Initially, we assume that a 5-year period is enough to represent the seasonality and the climatology of an area. The 35-year period of the CHESS dataset is used to describe the climatological values in the study area. The 5-year period [1991-1995] from CHESS is used as the control period, because it is similar to the climate average of temperature and precipitation

over the whole period [1981-2015] (Figure 6.5). For the 35-year period, the mean annual domain average temperature and precipitation are equal to 9.4 °C and 886 mm year$^{-1}$, respectively. The corresponding values for the 5-year control period are 9.3 °C and 876 mm year$^{-1}$, respectively. Therefore, this period could be representative for the whole period. However, we recognize that this approach has a few limitations. Firstly, the low variability in terms of rainfall for the 5-year period cannot depict the high variability of the entire period. Also, one simulation cannot provide with the uncertainty boundaries of the present climate conditions as if we used more simulations. Finally, if we run a 10-year simulation instead of 5-year we may get more robust results, as a bigger period may reduce the noise of the data.

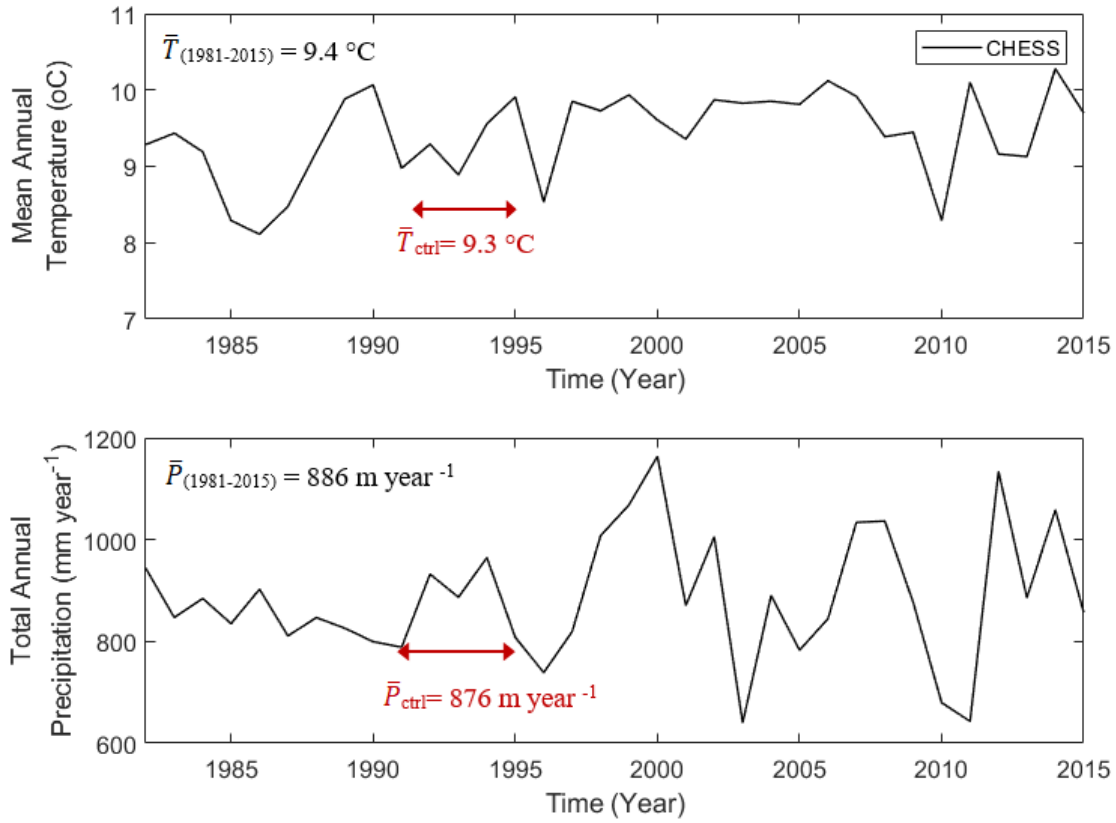


Figure 6.5: Comparison of mean annual Temperature (top) and total annual Precipitation (bottom) from CHESS for the domain average against a 5-year control period (dark red) from CHESS that represents the climatology of the study area.

In the next step, the 5-year moving average for the 12 ensembles is estimated for precipitation and temperature. For each ensemble, the 100-year time series gives 96 values of 5-year average precipitation and temperature. In total, we end up with 1152 sets of 5-year

periods for the 12 ensembles. There values are used against the control 5-year representative values for temperature ($\bar{T}$_ctrl) and precipitation ($\bar{P}$_ctrl) from CHESS to compute the deviations. For precipitation ($P$), it is estimated the percentage deviation ($\Delta P$) (Equation 6.6); whereas, for the temperature ($T$), it is estimated the difference of the two values ($\Delta T$) (Equation 6.7).

$$\Delta P=[(P(i,j)- \bar{P}\_\text{ctrl})/ \bar{P}\_\text{ctrl}]100 \quad (6.6)$$

where $P$ is the 5-year mean annual precipitation for every one of the 96 values ($i$) from each ensemble ($j$).

$$\Delta T=(T(i,j)- \bar{T}\_\text{ctrl}) \quad (6.7)$$

Figure 6.6 depicts the difference ($\Delta T$) of each 5-year mean temperature $T(i,j)$ with the control 5-year representative mean ($\bar{T}$_ctrl) against the percentage deviation ($\Delta P$) of each 5-year mean precipitation $P(i,j)$ with the control precipitation ($\bar{P}$_ctrl). The range of x and y axis are divided into 8 and 7 bins, respectively. The range of bins in x axis is 1 °C and in y axis is 10%. In that way, both axes have almost the same number of bins. In each bin, the point that is closer to the centre (red X in Figure 6.6) is selected to represent the centroid point of the bin. Every bin should have at least 6 points, that corresponds to 0.5% of the total number of points, in order to participate in the point selection. In total, 25 points/bins are selected which comprise 5-year daily data length. The samples represent the percentage change of annual precipitation from -35 to +25% with equal to 10% increments and additive change of mean annual temperature from -2 °C to +5 °C with +1 °C increment. With this methodology, we can include simulations that represent significant changes of rainfall and temperature and project the consequences while ensuring a compromise with regards to high computational demand from actually running our regional model simulations. In addition, for each run, we apply a 10-year spin-up by repeating the same 5-year period twice, in accordance to the same approach applied in Chapter 5.

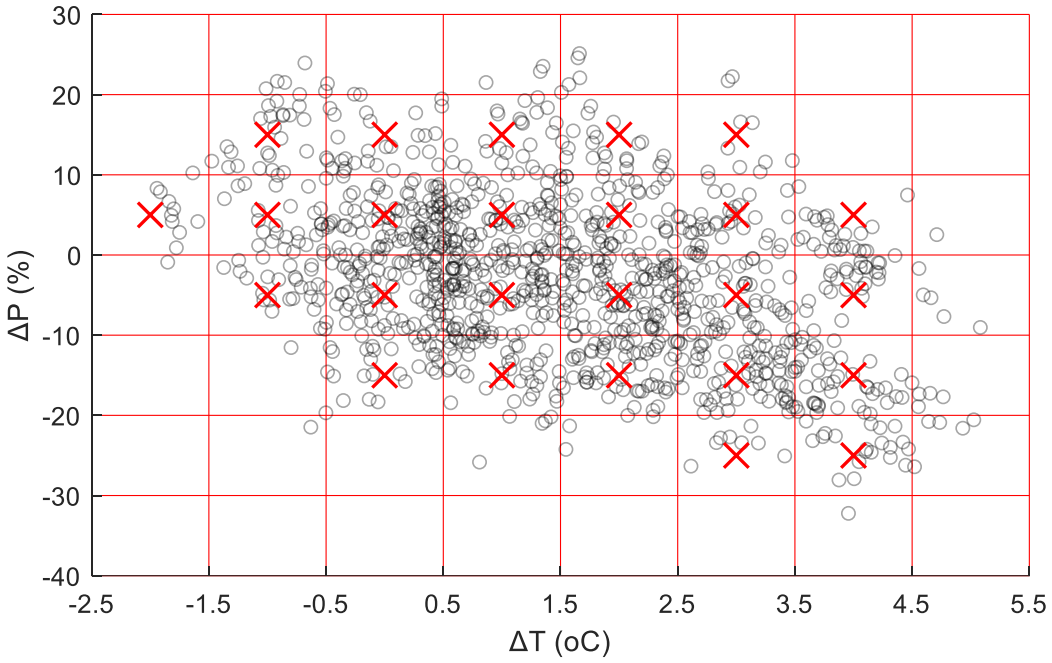


Figure 6.6: Different possible combinations of normalized differences for temperature Δ*T* (°C) and precipitation Δ*P* (%) between the 1152 sets of 5-year periods from the 12 ensembles of UKCP18 against the 5-year Control Period from CHESS. Red X represents the centroid of each bin. Note grids with fewer than six ensemble simulations available are excluded from our analysis.

## 6.3 Results and Discussion

### 6.3.1 Domain average

Figure 6.7 shows the domain average of the annual percentage change for the main hydrological variables (runoff, soil moisture, evaporation, and recharge) against the control case. The percentage change is calculated for each of the selected simulation based on the combined hydroclimatic conditions already presented in Figure 6.6. We can see that the magnitude of changes in runoff and recharge (i.e., different coloured grids) is changing more pronouncedly in the vertical direction (rainfall gradient) when compared to the horizontal direction (temperature gradient). We can infer that precipitation is the dominant factor and higher precipitation is associated with higher runoff and recharge. In general, in water-limited regions (i.e., water availability limits the values of actual evapotranspiration), runoff is very sensitive to the change of precipitation (Liang and Liu, 2014). However, UK is an energy-limited region (i.e., the rate of energy supply limits the values of actual evapotranspiration), so

we expect a lower sensitivity of the runoff to the change of precipitation (McVicar et al., 2012). This is the reason why we still see a change in runoff and recharge in the horizontal axis (temperature gradient).

On the other hand, it seems that for soil moisture the dominant factor is the temperature for up to -5% decrease of annual precipitation. In drier conditions ($\Delta P$ < -5%), there is a different behaviour with more significant percentage change in the annual soil moisture throughout the soil column (3 m) that does not change significantly when the mean annual temperature increases. With regards to evaporation, both precipitation and temperature affect the annual percentage change as we find changes in the diagonal. We can deduce that the higher precipitation in combination with higher temperature lead to more evaporation.

As we expected, the magnitude of change for each component is not the same (Figure 6.7). Runoff is expected to increase up to 23% if the annual precipitation increases 5% and the temperature decreases -2 °C in the same time. If the annual precipitation increases 15 %, then for almost all the changes in the temperature, the mean annual runoff will increase 15%. This shows that when the climate becomes very wet, the temperature does not affect the results significantly. It probably affects the evaporation, but we do not expect the 15% increase of evaporation to affect the runoff considerably, because the contribution of evaporation to the water balance is much lower compared to runoff. Finally, runoff is expected to decrease up to 22% if the temperature increases +4 °C and precipitation decreases 25%. Considering soil moisture there is similar percentage change for $\Delta P$ > -5%. In particular, the percentage change of soil moisture goes from 1.2% (for $\Delta T$ = +1 °C) to -2.5% (for $\Delta T$ = +4 °C). However, for $\Delta P$ ≤ -15%, the percentage change of soil moisture goes from -3.2% (for $\Delta T$ = +1 °C) to -5.5 % (for $\Delta T$ = +4 °C). In the latter case, the range is smaller. Evaporation shows a patchier behaviour. Simulations with negative temperature change and precipitation with change ±5% lead to evaporation decrease up to 15%. On the other hand, evaporation can increase up to 18% if temperature increases +3°C and precipitation +15%. Finally, in terms of recharge, there are two distinct behaviours if the rainfall increases (decreases) that lead to more (less) recharge. The increase of recharge can be from 70% for temperature +1 °C and precipitation +15% to -90% for temperature +4 °C and precipitation -25%. From all the water components, recharge shows the highest sensitivity to the climate change scenarios.

In the next step, we want firstly to identify which combinations lead to positive, neutral, or negative changes in runoff. Furthermore, we want to examine if the changes in the magnitude of runoff lead to changes in the dynamics and in the timing of runoff. The last question is related

to the soil moisture. We want to find if there is a tipping point for $\Delta P \geq$ -5% that lead to different behaviour in the soil moisture. As it is mentioned before, we find similar behaviour when $\Delta P \leq$ -15%. We want to know if there is a particular reason for that. Our hypothesis is that under these extreme dry conditions, the study domain may be converted from energy-limited to water-limited region. To address these questions, we choose different simulations for each question and we compare them either against the control case that represents climatology in the area or between other simulations from UKCP18 projections.

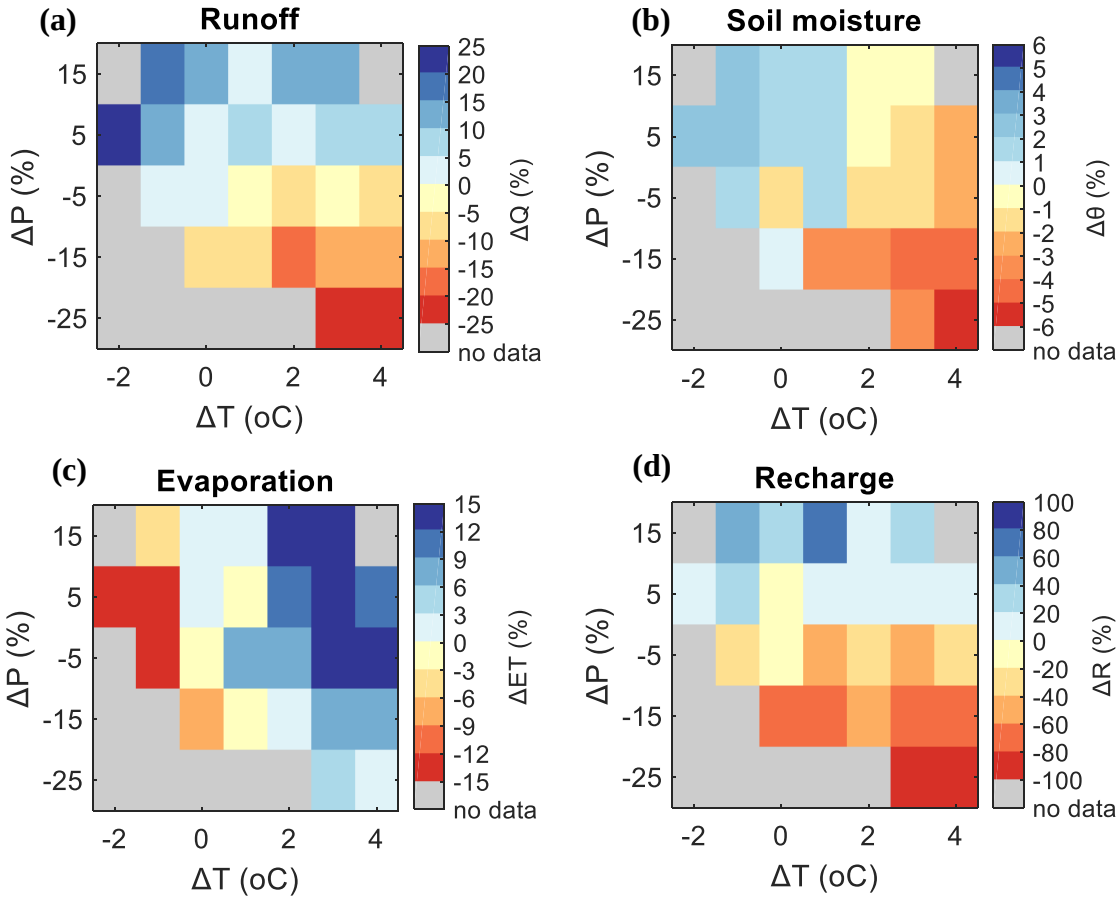


Figure 6.7: Percentage change of the domain average for Runoff (a), Soil moisture (b), Evaporation (c) and Recharge (d) against the control case that represents climatology for the 25 different simulations of UKCP18 projections.

**Which climate scenarios lead to positive, neutral, or negative change of runoff?**

In Figure 6.7, we showed that runoff is more strongly controlled by rainfall, albeit temperature has non negligible impact as UK is an energy-limited area. Thus, higher temperature (more energy) will increase the evapotranspiration flux. Small percentage change

of annual rainfall ($\Delta P$ ±5%) is usually linked with neutral impact on runoff ($\Delta Q$ ±5%). Exemptions are for extreme dry or cold conditions. Percentage increase (decrease) of rainfall about 15% (-15%) is linked with positive (negative) change of runoff. Interestingly, the highest percentage change of runoff (24.7%) does not happen for the wettest scenarios ($\Delta P$ +15%), but for $\Delta P$ set to +5% under the coldest scenario ($\Delta T$ -2 °C). It seems that the combination of extreme scenarios for both temperature and precipitation is linked with more significant change of runoff. This is reasonable, because under these extreme conditions, evaporation and runoff generation mechanisms work additively to increase or decrease runoff. In other words, under extreme wet and cold conditions, runoff generation mechanisms (i.e., infiltration excess and groundwater contribution) produce more water due to the higher precipitation. More water will be recharged to the aquifer and after a lag response, it will reach the surface through discharge and baseflow contribution. At the same time, soil evaporates less water because there is less available energy (lower temperature) and the moisture stress factor $\beta$ tends to zero, as soil moisture concentration decreases, that ultimately leads to lower net photosynthetic uptake. Therefore, the runoff increases because there is more water available for runoff and baseflow contribution, while less water coming from rainfall contributes to evapotranspiration. Thus, under these conditions, runoff is projected to increase significantly. Similarly, under extreme dry and warm conditions, the change of precipitation and temperature work additively towards the decrease of runoff. The significant lower precipitation causes the decrease of runoff due to infiltration excess mechanism and due to the lower contribution to the groundwater. The higher temperature increases substantially the evapotranspiration flux. The combination of all these physical processes lead to considerable reduction of runoff and recharge.

Intuitively, we should expect that changes of runoff to be somewhat proportional to rainfall changes, as runoff is more strongly controlled by rainfall. For example, we find that for $\Delta P$ -15% and $\Delta T$ +4 °C, the mean annual runoff decreases about $\Delta Q$ -12%. The increase in temperature will only exacerbate the issue because some of the water coming from the rainfall, that could have contributed to runoff, also gets evaporated. To analyse further how an "average year" looks like, we plot the domain daily average of the 5-year simulations of rainfall, runoff, temperature, and evaporation (Figure 6.8). In addition, seven-day moving average has been applied for smoothing purposes. Interestingly, we do not find significant difference during the year in the contribution of evaporation between the two simulations (Figure 6.8 bottom). As we expected, temperature is higher throughout the year for the climate simulation [JULES (P-15 T+4)] with peak to occur in late summer. The difference of temperature between JULES-Ref

and the UKCP18 simulation is more pronounced in later summer and early autumn. Thus, we should expect higher difference in the daily evaporation during this period. However, the evaporation timeseries does not follow the same pattern, as the peak occurs in early summer. The reason probably is that under the dry conditions, evaporation is lower because there is less available water in the soil domain for evaporation. Another explanation could be that the radiation is lower, thus there is less available energy.

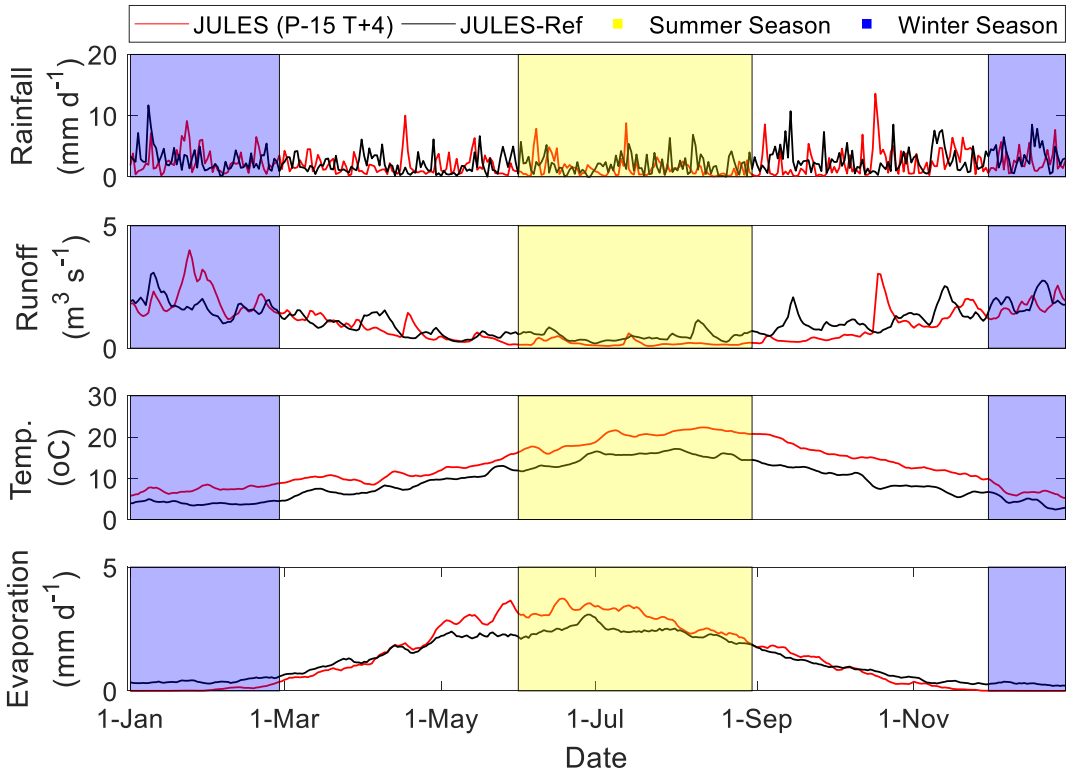


Figure 6.8: Domain daily average of the five-year simulation for rainfall (first figure), runoff (second figure), temperature (third figure) and evaporation (forth figure) for [JULES (P-15 T+4)] (red) against JULES-Ref (black). For illustration purposes, seven-day moving average has been applied in runoff, temperature, and evaporation.

For analysing the radiation, we calculate the outgoing longwave and the net radiation using equations 6.8 and 6.9, respectively.

$$LW\uparrow = \sigma \varepsilon_s T_s^4 \tag{6.8}$$

$$R_n = (1\text{-}\alpha)\ \mathrm{SW} - \sigma \varepsilon_s T_s^4 + \sigma \varepsilon_a T_a^4 \tag{6.9}$$

where $LW\uparrow$ is the outgoing longwave radiation (W m$^{-2}$), $R_n$ is the net radiation (W m$^{-2}$), $\varepsilon_a$ is the air emissivity assumed equal to 0.8 (unitless) and $T_s$ is the soil surface temperature (K). The outgoing longwave, the net radiation and the average relative saturation throughout the soil domain (3 m) for each simulation are presented in Figure 6.9.

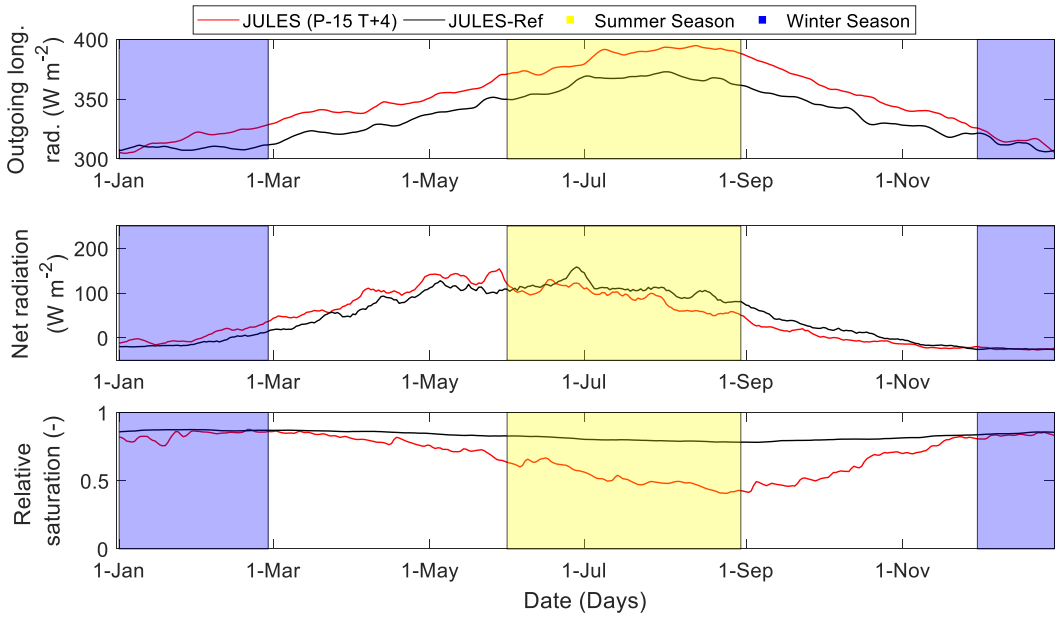


Figure 6.9: Domain daily average of the five-year simulation for net outgoing shortwave radiation (first figure), net radiation (second figure) and relative saturation (third figure) for [JULES (P-15 T+4)] (red) against JULES-Ref (black). For illustration purposes, seven-day moving average has been applied.

The outgoing longwave radiation is much higher for the climate change simulation and especially in summer and early autumn. Looking the net radiations, the results from the two simulations are similar with slightly lower values for [JULES (P-15 T+4)] during late summer and higher values during spring. This can partially explain the behaviour of evaporation that does not follow the pattern of air temperature during late summer and early autumn. However, it seems that the much drier soil moisture in the soil domain for the climate simulation has stronger impact on evaporation during late summer. In JULES-Ref, the average relative saturation throughout the 3 m soil domain ranges between 0.9 in wet season and 0.8 in dry season, while the corresponding values for [JULES (P-15 T+4)] are 0.85 and 0.4. In JULES, evapotranspiration includes bare soil evaporation, transpiration, and canopy evaporation. Soil evaporation and transpiration are restricted by the soil moisture state (Best et al., 2011). The minimum values for the relative saturation under the extreme dry conditions occur in late

summer to early autumn when we find similar evapotranspiration values for the two simulations. Hence, we could deduce that the reason that evapotranspiration does not exacerbate the decrease of the runoff under this climate scenario is the low availability of the water in the soil domain with the lower net radiation to have a smaller impact, especially during spring.

**Have the changes in magnitude of runoff also imply changes in its intensity and dynamics?**

Rainfall in [JULES (P-15 T+4)] is generally lower (Figure 6.8), especially during summer when only a few minor rainfall events occur with low intensity (less than 5 mm $d^{-1}$). However, there are a few significant rainfall events (higher than 20 mm $d^{-1}$) that occur in autumn, spring, or early summer for the UKP18 projection. Interestingly, although the JULES-Ref simulation gives more annual precipitation, it has lower intense rainfall events. This is consistent with the reports from Met Office (Met Office, 2018a, b) that project more intense thunderstorms in the future. The lower rainfall throughout the year and especially during the summer affects runoff with almost zero values for the UKCP18 simulation. The runoff during the spring and autumn has the same magnitude between the climate projection [(JULES (P-15 T+4)] and the reference model, whereas [JULES (P-15 T+4)] is wetter during the winter. Interestingly, rainfall events with same magnitude during summer season lead to lower runoff.

Figure 6.10 proves this statement by plotting the monthly rainfall values for the domain average against average monthly runoff rate during winter (blue) and summer (red) for the two simulations. The fit of [JULES (P-15 T+4)] simulation for summer season is more horizontal compared to JULES-Ref. It seems that the drier soil during summer does not generate significant runoff regardless the magnitude of precipitation. The explanation of this behaviour is the infiltration excess mechanism. In the model, the rainfall splits into runoff and infiltration. As soil is drier in [JULES (P-15 T+4)], more water enters the soil. Hence, there is less available water for direct runoff. Furthermore, the low available water in the soil domain is used for evapotranspiration and does not reach the aquifer. As a result, the contribution of groundwater is lower. The steeper linear fit during the winter compared to summer is consistent with our hypothesis that the domain flashes more water during the cold season for the same amount of rainfall. For instance, for monthly rainfall of 100 mm, the average runoff rate during the summer is less than 1 $m^3 s^{-1}$, whereas during the winter is 2 $m^3 s^{-1}$. The model tends to produce up to 3 times more runoff during the winter for the same amount of rainfall in [JULES (P-15 T+4)] simulation. On the other hand, the linear fit for winter is slightly steeper for [JULES (P-15

T+4)] and this can be attributed to the lower evaporation during this season (Figure 6.8 bottom). Similar conclusions can be drawn for the daily time step, as well, even though the data are more scattered.

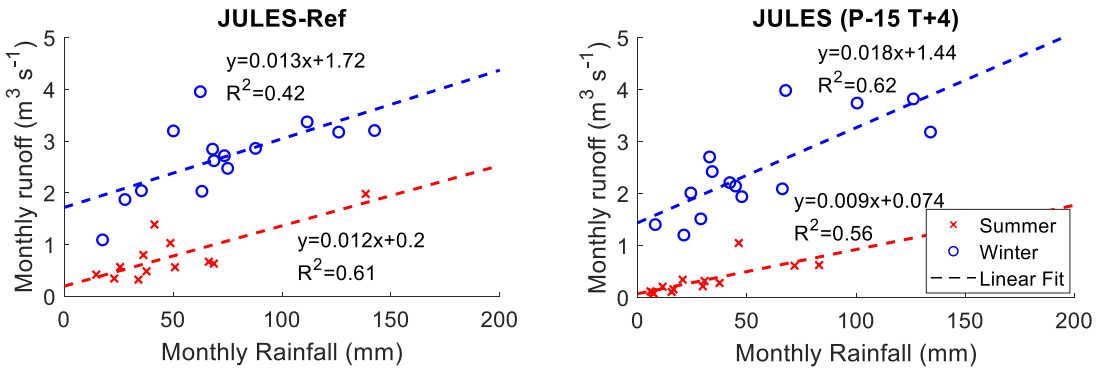


Figure 6.10: Scatter plot of monthly rainfall against monthly runoff (domain average values) for JULES (P-15 T+5) (right) against JULES-Ref (left) for summer (red) and winter (blue) seasons.

Previously, we showed that the response of the region to rainfall events is different between winter and summer seasons as during summer it is generated less runoff for the same amount of rainfall. However, so far, we have not discussed about the timing and the intensity of runoff. Figure 6.11 tries to reveal these characteristics of runoff. First, looking the month with the highest average runoff (Figure 6.11, left), we find similar behaviour for dry and warm climate scenarios (bottom right) with January to be the wettest month. On the contrary, December is the wettest month under warm and wet climate scenarios (upper right), while there is not a clear answer for cold and wet climate scenarios. In terms of the month with the lowest monthly runoff (Figure 6.11, medium), there is a distinct behaviour for simulations without change of temperature and for wet and warm climate scenarios. Under these cases, August is the driest month. There is another cluster of simulations for $\Delta T$ +3 °C, with July to be the driest month, while under $\Delta T$ +4 °C, again July or August is the driest month. We find this shift under extreme dry conditions and the reason is the lower wetness of the soil column. For wet and cold climate conditions, the driest month is the same with the month that gives the highest evaporation flux, usually June or July. We recognize that the bias correction method may have changed the timing of the variables, as the correction is based on the climatological conditions and dynamics between 1981 to 2015.

Finally, regarding the intensity of maximum daily runoff, it seems that the highest values occur for wet and cold conditions (Figure 6.11, right) and for extreme warm conditions (ΔT +4 °C). In the first case, the daily maximum runoff can be up to 20 $m^3$ $s^{-1}$ and in the second case up to 13 $m^3$ $s^{-1}$. The climate simulations between the two clusters give maximum daily runoff about 6 $m^3$ $s^{-1}$. The high maximum runoff values for extreme warm conditions is consistent with the report from Met Office (Met Office, 2018a, b) that predict more intense thunderstorms under these conditions.

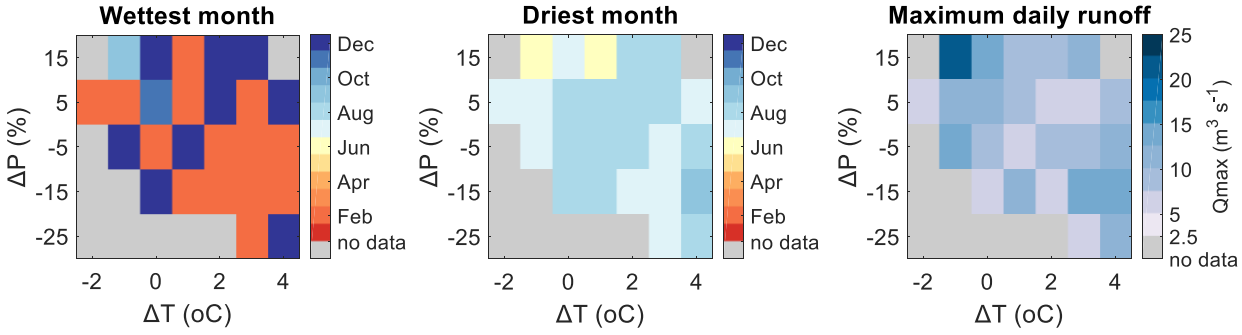


Figure 6.11: Month of the year with the maximum (left) and minimum monthly runoff (medium). The maximum daily runoff is presented on the right for the 25 different simulations of UKCP18 projections.

**Is there a tipping point that affect the soil moisture projections?**

In Figure 6.7c for soil moisture, we found two distinct behaviours. In particular, the results for percentage deviation $\Delta P \geq$ -5% give similar values across the y axis with temperature to be the dominant factor. However, for drier conditions, the soil moisture does not change more than 1% from simulation to simulation. We want to find if there is a tipping point that can explain this behaviour. We selected two simulations from climate change scenarios to identify the different dominant processes that dictate these different behaviours, namely [JULES (P-15 T+3)] and [JULES (P-5 T+3)]. In these simulations, the change in the temperature is the same, and the only difference is the precipitation change.

The timeseries of the four variables are presented in Figure 6.12. The first interesting finding is that although summer rainfall has a few significant events with intensity higher than 10 mm $day^{-1}$, summer runoff is almost zero. We find significant flood events during autumn for [JULES (P-15 T+3)] that lead to floods. As we discussed above, extreme climate conditions are associated with intense thunderstorms. However, when these thunderstorms occur during summer, they do not generate significant runoff. Figure 6.10 showed the different response of

the domain between summer and winter under the same amount of precipitation. The linear fit for drier conditions during summer becomes even flatter with the magnitude of precipitation to have low impact on the magnitude of the flood. As we expected, even though the temperature is similar for all the seasons between the two simulations, there are significant discrepancies in the evaporation, especially during summer and autumn. The explanation is the drier soil domain that cannot provide with enough water for evaporation. We find the opposite behaviour for the spring season when evaporation is higher for JULES [(P-15 T+3)].

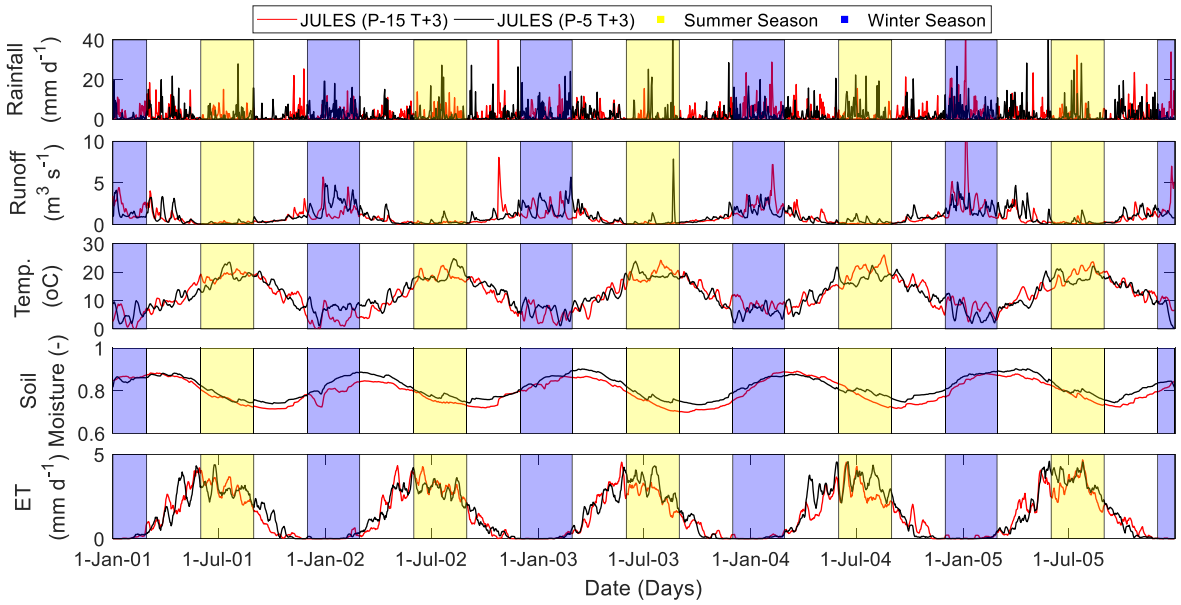


Figure 6.12: Domain average rainfall (first figure), runoff (second figure), temperature (third figure), soil moisture (fourth figure) and evaporation (fifth figure) for [JULES (P-15 T+3)] (red) against [JULES (P-5 T+3)] (black). For smoothing purposes, seven-day moving average is applied in temperature and evaporation.

In order to test our assumption that there is a tipping point, we plot the daily soil moisture against temperature values, while we use evaporation to colour the samples (Figure 6.13). While generally evaporation is increasing in the same direction as temperature increases, there is a threshold of the relative saturation around 0.74 under which there is shortage of water. The domain is unable to provide with enough water, so the high temperature is not sufficient to generate evaporation. We can deduce that there is a tipping point for extreme dry conditions under which the temperature is not the dominant factor of soil moisture and evaporation. This tipping point is related to the soil characteristics and it is not expected to be the same

everywhere. In our case this point is for annual precipitation decrease greater than or equal to 15%. Beyond this point we expect extreme droughts. Under these conditions, the study domain may move towards to becoming an energy-limited region.

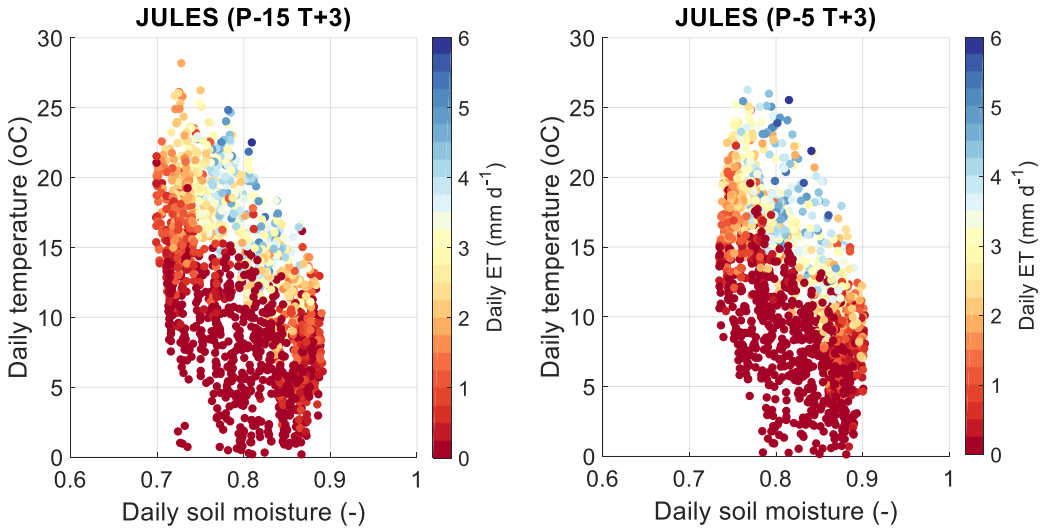


Figure 6.13: Daily soil moisture in the soil domain against daily temperature for [JULES (P-15 T+3)] (left) against [JULES (P-5 T+3)] (right). Daily evaporation values are used to colour the samples from red to represent zero evaporation and blue to represent the maximum value (6 mm).

### 6.3.2 Catchment analysis

In the domain analysis, we analysed the impact of climate change scenarios on water components for the whole domain. Here, we analyse the components separately for each catchment. In general, we identify similar patterns from catchment to catchment. It is something we should expect considering the small size of the area that lead to similar forcing. Furthermore, the characteristics of the catchments (land cover, soil type) are similar with all of them to be groundwater dominated catchments with BFI from 0.51 to 0.65. An interesting observation during this analysis is the change of runoff ratio (i.e., ratio of annual runoff over annual rainfall) for different climate change scenarios. Figure 6.14 presents the runoff ratio for four catchments and for the 25 different climate change scenarios. We find a change diagonally with high runoff ratio under drier and colder climate conditions and lower runoff ratio under wetter and warmer conditions. The range between the highest and the lowest value is about 0.15 for the smaller catchments (Lugg at Byton and Rea Brook at Hookagate) and 0.1 for the rest of them.

In general, runoff ratio is affected by precipitation, evapotranspiration, slope, soil characteristics and land cover type (Ratzlaff, 1994). As the domain is not large and the forcing is similar, we want to analyse why there are these differences among the catchments. We compare the daily timeseries of the main water components for the two catchments with the strongest differences (Lugg at Byton and Rea Brook at Hookagate). These two catchments have almost the same area (203 and 178 km$^2$). We want to find why the Rea Brook catchment gives lower runoff ratio values that implies that it generates less runoff.

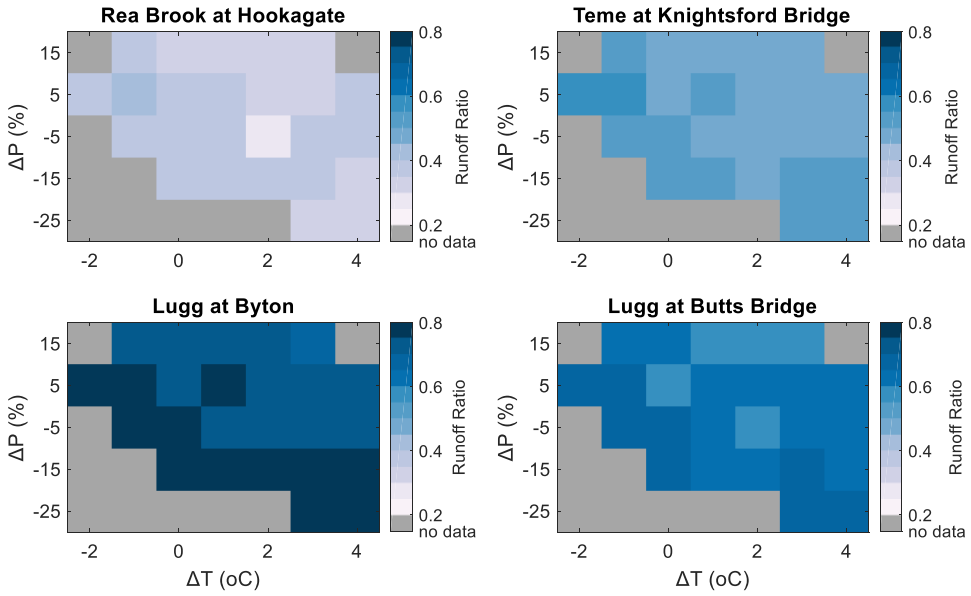


Figure 6.14: Runoff ratio for four different catchments in the domain under different climate conditions.

The daily rainfall, runoff, average relative saturation throughout the soil column (3 m) for the two catchments under the (ΔP = +5%, ΔT = -1 °C) climate scenario are shown in Figure 6.15. Positive values of recharge mean that water from the soil domain (3 m) moves into the aquifer, whereas negative values mean that water moves out of the saturated aquifer to the land surface (discharge). Under this climate scenario, Rea Brook catchment receives less water (830 mm year-1) compared to the Lugg at Byton (950 mm year-1) and it produces less runoff. The dominant land cover type is C3 grasses for the Rea Brook catchments and Shrubs for the Lugg catchments. Shrubs and C3 grasses have the same characteristics in JULES (e.g., infiltration enhancement factor, albedo) (Best et al., 2011), so we do not expect differences coming from the land cover type. The dominant soil type is clay loam with higher hydraulic saturated conductivity values ranging from 0.009 to 0.05 mm s$^{-1}$ for the Rea Brook catchment, whereas Lugg has clay loam and silty clay loam with lower conductivity values about 0.006 mm s$^{-1}$. The

higher saturated conductivity values for the Rea Brook catchment imply higher infiltration rate. Hence, less water is used for direct runoff. Furthermore, slope is the key element for this analysis. The Rea Brook catchment is relatively flat with a wide valley upstream of the outlet. On the other hand, Lugg is substantially steeper with a ravine upstream of the outlet. Thus, the groundwater contribution is more significant, as can be seen from Figure 6.15d, especially during late summer and early autumn. The steeper slope of the Lugg catchment in combination with the lower infiltration rate compared to the flat terrain of the Rea Brook catchment with the significant infiltration rate can explain the higher runoff ratio of the first catchment. Evapotranspiration flux is almost identical for the two catchments (data not shown here), so it does not affect the different runoff ratio.

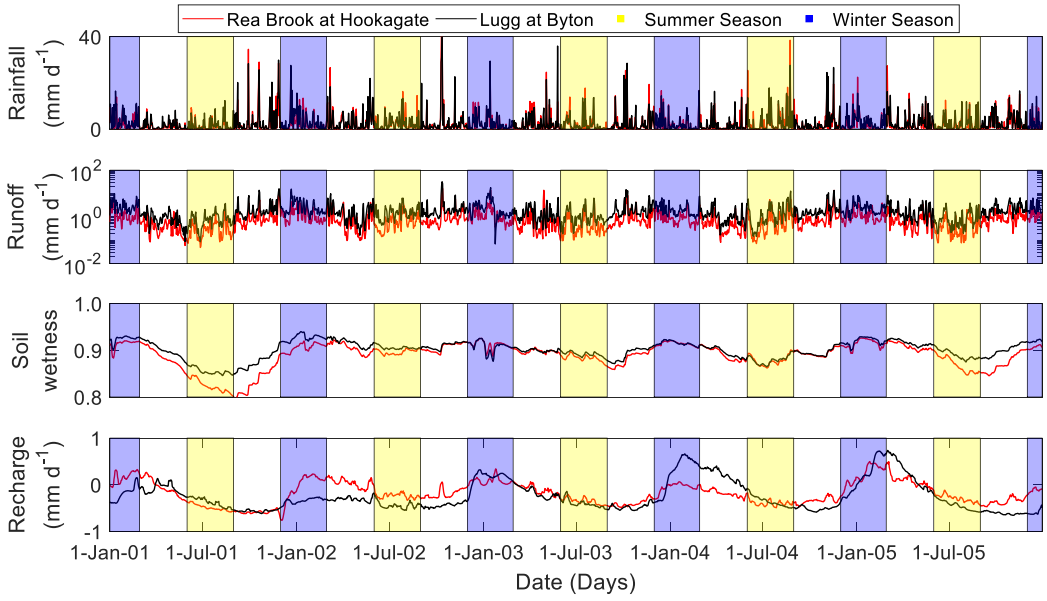


Figure 6.15: Daily rainfall (first figure), runoff (second figure), relative saturation (third figure) and recharge for the Rea Brook catchment (red) against Lugg at Byton catchment (P-5 T+3) (black) under the ($\Delta P$ = +5%, $\Delta T$ = -1 °C) climate scenario. For smoothing purposes, seven-day moving average has been applied in relative saturation and recharge. Positive (negative) recharge implies recharge (discharge).

In the last step we try to find which water component is responsible for the behaviour of runoff ratio (Figure 6.14) that is changing for each climate scenario. We analyse the Tame at the Knightsford Bridge catchment because it is the biggest one. To calculate the dynamics between the water components against runoff ratio, we use linear correlation. We calculate

cross correlation between the mean annual values of each water component for each climate scenario against runoff ratio (Figure 6.16). We have also included precipitation and temperature as the key hydrometeorological variables that define the climatology of an area. According to Schober et al. (2018) values of correlation above 0.7 should be considered as strong, between 0.4 and 0.7 as moderate and between 0.1 and 0.4 as weak. We recognize the limitation of this approach that assumes that one variable is caused by another only, but we know there are multiple interactions. However, we can gain insight at least for the direct influence of one variable on another. One more drawback of this approach is that linear correlation measures only linear relationship between two variables and ignores the non-linear relationship.

Evaporation has the largest correlation in absolute values against runoff ratio (-0.74). Recharge has a weak impact with -0.34 linear correlation. Precipitation and temperature have weak and moderate negative correlations equal to -0.38 and -0.43, respectively. The strong negative correlation between evaporation and runoff ratio suggests that when the evaporation increases, then runoff ratio decreases. The reason is the biggest percentage of rainfall that is converted into evaporation instead of runoff. So, as the numerator decreases the ratio decreases, as well. The weak negative correlation between recharge and runoff ratio shows that when recharge increases, water that was supposed to run overland go into the aquifer. Thus, again the numerator decreases and so does the runoff ratio. Finally, it seems that the runoff component and relative saturation have no impact on the behaviour of runoff ratio.

Figure 6.16 does not only help us to explain the change of runoff ratio under different climate change scenarios, but also summarizes the findings from this analysis. The strong positive correlation between evaporation and temperature (0.82) and the moderate negative correlation between evaporation and relative saturation (-0.47) prove that the temperature is the dominant factor that dictates the magnitude of evapotranspiration. However, the presence of water in the soil domain has moderate impact, as well. Interestingly, recharge is more strongly correlated with precipitation (0.94) compared to relative saturation (0.77). Furthermore, the similar correlation of recharge and precipitation (about 0.77) with relative saturation show that the latter does not depend only precipitation. Now, that there is a coupled groundwater model below JULES domain, the soil moisture in the soil domain depend on the groundwater, as well. We hypothesize that in the free drainage case, when the free drainage flux increases, the relative saturation in the soil moisture decreases because the soil domain is drying. Under these circumstances, we should expect negative correlation between recharge (free drainage) and relative saturation.

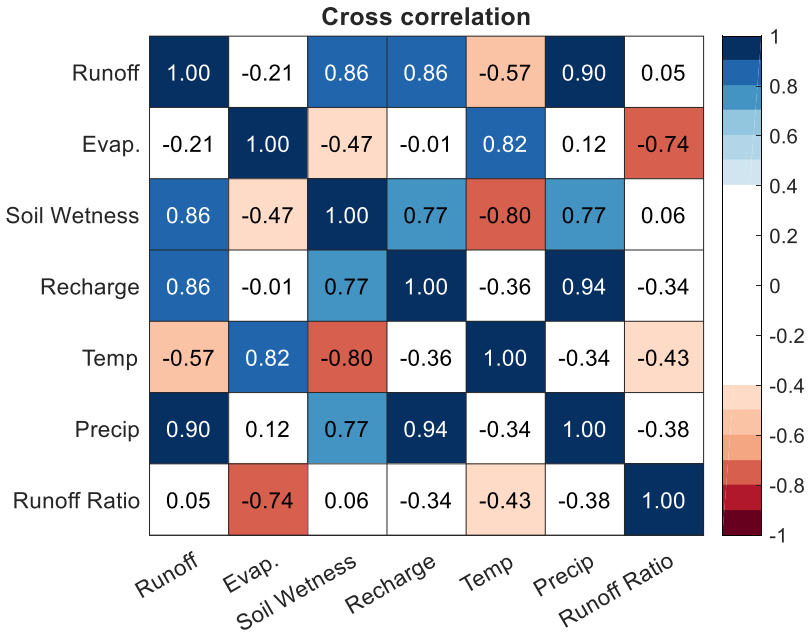


Figure 6.16: Cross correlation of the annual values of main water components and forcing for different climate scenarios at Tame at the Knightsford Bridge catchment. Colours have been used only for moderate and strong correlation values.

## 6.4 Conclusions

In this Chapter, we investigate the impact of climate change in Central England using JULES-GFB and the UKCP18 climate projections. Simulations are performed for 25 different climate scenarios. The aim is to examine the impact of the main water components under different climate conditions, from dry (25% decrease of annual precipitation) to wet (15% increase of annual precipitation) and from cold (-2 °C) to warm conditions (+4 °C).

The results from the domain analysis demonstrate that precipitation and not temperature has the greatest impact on surface and subsurface hydrology. Runoff and recharge are significantly dominated by the percentage change of precipitation, as we find no differences across different changes in the temperature. The percentage changes in these two components become substantial for higher than 15% decrease of annual precipitation. Runoff can decrease up to 23% and recharge up to 90% under extreme dry (-25%) and warm conditions (+4 °C). Under colder conditions evaporation decreases up to 15% because there is less energy for evaporation. Under warmer and wetter conditions evaporation can increase up to 18%. The range in soil moisture is much lower with +2% under wet and cold conditions to -5.5% under extreme dry and warm conditions.

Furthermore, different response of the region to rainfall events is found, as during winter the domain generates up to three times more runoff compared to summer. The reason is the drier soil during summer that implies higher proportion of rainfall to be infiltrated in soil and higher evapotranspiration rate. In terms of the timing of the peak and trough in runoff, January is the wettest month under dry and warm climate scenarios and December under warm and wet climate scenarios. August is usually the driest month. Regarding the intensity of maximum daily runoff, the highest values occur for wet and cold conditions (up to 20 $m^3$ $s^{-1}$) and for extreme warm conditions ($\Delta T$ +4 °C) (up to 13 $m^3$ $s^{-1}$).

For annual decrease of precipitation greater than or equal to 15%, it seems that the impact of temperature on soil moisture and evaporation is decreasing because there is not enough water for evaporation. This occurs when the average relative saturation over the model domain is less than 0.74. In the catchment analysis, we found runoff ratio to change diagonally when temperature and precipitation are increasing or decreasing simultaneously with lower values under wet and warmer conditions. Evaporation found to have the strongest correlation in absolute terms against runoff ratio. This negative correlation suggests that when the evaporation increases, then runoff ratio decreases because biggest percentage of rainfall is converted into evaporation instead of runoff.

# 7 SUMMARY AND CONCLUSIONS

Land Surface Models are now more widely used for hydrological studies, along with studies to assess the impact of climate change and for operational purposes. However, some of them still rely on the free drainage assumption, which describes the bottom boundary condition for soil moisture dynamics, while ignoring the groundwater contribution. This affects the realism as LSMs with no groundwater parameterization cannot capture the interactions between saturated and unsaturated zone, possibly leading to inaccurate water partitioning. In this thesis, the impact of explicitly representing groundwater in an LSM is explored in order to explain and quantify the effects on land surface water balance components, such as runoff, evapotranspiration, and soil moisture. The Joint UK Land Environment Simulator (JULES) model is used in this analysis. The regional forcing data were obtained from the Climate, Hydrology and Ecology Research Support System (CHESS) dataset (Robinson et al., 2016). The elevation map was derived from HydroSHEDS (Hydrological data and maps based on Shuttle Elevation Derivatives at multiple Scales) (Lehner et al., 2008). The Land Cover Map 2007 (LCM2007) (Morton et al., 2011) and the Land Information System provided by the Cranfield University were used to obtain land cover type information and the soil parameters, respectively. Daily streamflow and evaporation values from the National River Flow Archive (NRFA) and the Global Land Evaporation Amsterdam Model (GLEAM) (Miralles et al., 2011; Martens et al., 2018), respectively, were used to validate the model. Finally, the UK Climate Projection 2018 dataset (UCKP18) was applied to simulate climate change scenarios.

In the first step, the performance of the default version of JULES was assessed in terms of accurate daily streamflow predictions in flashy (overland-dominated) and groundwater-dominated catchments in Great Britain (Chapter 4). Three distinct behaviours were found in the experiment associated respectively with flashy, groundwater-dominated, and chalk catchments.

The model performed satisfactorily (i.e., Kling and Gupta Efficiency (KGE) above 0.5) in flashy catchments, which are rainfall-dominated and have low baseflow contribution to the outlet (i.e., low values of Base Flow Index; BFI). In these catchments, the infiltration excess and the saturation excess mechanisms are sufficient to represent the runoff generation mechanisms. The missing groundwater representation in JULES deteriorated its performance to predict runoff in groundwater-dominated catchments with BFI more than 0.5. In these cases, the KGE was very low and in many cases, negative suggesting the failure of the model in groundwater-dominated regions. However, in both overland and groundwater flow dominated catchments, the model fails to simulate low flows. This could affect the response of the model under different climate change scenarios by alerting the impact of a drought.

In the second step, a groundwater parameterization was coupled to JULES (JULES-GFB) to replace the free drainage assumption (Chapter 5). The model was first tested in a more controlled manner with two synthetic experiments, and subsequently using a real-case regional experiment in Great Britain. The results suggested improvement in the realism of the model as the new model had superior performance compared with the version of the model where free drainage occurs (JULES-FD), especially for shallow water regions and coarse soil types. JULES-GFB improved soil moisture dynamics, while successfully representing lateral water flow in the saturated zone, and consequently better estimating the contribution of groundwater to river discharge. Model evaluation of JULES-GFB in a groundwater-dominated region in Great Britain showed an increase of model performance relative to other hydrological fluxes, in particular to catchment-integrated streamflow and evapotranspiration. From this test case, it can also be concluded that representing groundwater hydrology explicitly can supersede the advantage of implementing a calibrated saturation excess runoff generation scheme in JULES.

In the last step, the impact of climate change was evaluated using the newly developed JULES-GFB and the climate change scenarios obtained from the UK Climate Projections 2018 (UKCP18) dataset (Chapter 6). The model demonstrated that runoff and recharge were significantly dominated by precipitation, as expected in humid environments such as those observed on the UK. Runoff could decrease up to 23% and recharge up to 90% under extreme dry (-25% of mean annual precipitation) and warm conditions (+4 °C of mean annual temperature). Different response of the region to rainfall events was found during winter as the domain generated up to three times more runoff compared to summer. The reason was the drier soil during summer that implied higher proportion of rainfall to be infiltrated in the soil domain and higher evapotranspiration. Under these extreme conditions, the impact of temperature on

soil moisture and evaporation was decreasing because there was not enough water for evaporation. Finally, in the catchment analysis, we found runoff ratio to change for different climate change scenario with evaporation to have the strongest correlation in absolute terms against runoff ratio.

In conclusion, we examined JULES performance to simulate runoff in 47 catchments in UK with different characteristics. The aim was to find the limitations of the free drainage assumption that JULES uses and how these affect the performance of the model. The geology and the baseflow contribution to the runoff were the dominant factors that influenced the model performance. We substantially improved the hydrological representation of JULES by introducing a simple groundwater parameterization (JULES-GFB). We extensively tested the new approach against two synthetic well-controlled experiments and subsequently applying the new model to the real-world in a regional domain in the UK. The objective of the first synthetic experiment was to assess how the wetting front interacts with the aquifer in a column ignoring the lateral flow. In the second experiment, we evaluate the contribution of the lateral flow to the simulated discharge in a tilted-V catchment. The real-world application helped us test the new model against actual data in conditions that are not homogeneous and ideal, like these in the synthetic experiments. The ultimate goal of our development aimed to provide a modelling framework capable of simulating the hydrometeorological processes for large-scale applications. The new model improved the realism and the water partition of the model and addressed the problem of the missing groundwater representation in JULES. As the newly developed JULES-GFB proved to be more accurate and robust from JULES-FD, we used it to assess the impact of climate change in Central England and Wales. We used the UKCP18 dataset to simulate climate change scenarios. The new model answered questions related to the impact of climate change on water components (e.g., recharge, runoff, soil moisture) in groundwater dominated catchments. The fact that the new model can capture the interactions between aquifer and unsaturated zone gives credibility to the model, especially in climate change studies. Finally, it is worth highlighting that the new groundwater parameterization proposed here is relatively simple and does not require substantial computational time.

**Recommendations for future work**

Here are a few specific points that highlight a few recommendations that may contribute to this work in the future:

- The column-scale infiltration experiment provided us with results on the impact of the free drainage scheme in a soil column using a specific forcing and ignoring the impact of climate. Future work should generalize the column-scale experiment globally by including different forcing based on climate classification, like that of Knoben et al. (2018). This analysis could provide us with a more accurate picture about the error we get using LSMs with the free drainage scheme without the need to run the model globally.
- The application of JULES-GFB was done in a region in Great Britain with a somewhat narrow range in terms of hydrogeological and climatic variabilities. This region is characterized by a humid climate (i.e., energy limited), hence the impact of the groundwater model on evapotranspiration predictions will be small when compared against the default model. The reason is that the soil domain is always wet, so the free drainage assumption cannot dry it. Thus, as the differences of the wetness of the soil domain between JULES-FD and JULES-GFB are not large, the differences of evapotranspiration are not large as well. However, in an arid environment, the domain would be much drier in the model with the free drainage scheme and consequently the difference of simulated evapotranspiration between JULES-FD and JULES-GFB would be higher. The application of the model in areas with different climate, especially in arid areas, could highlight the impact of representing explicitly groundwater and the difference between the two configurations.
- JULES Default (Best et al., 2011) and our approach does not take into consideration the capillary flux. Capillary flux is defined as the upward flux from shallow groundwater to the overlying rooting zone of the soil (Vergnes et al., 2014). However, we did not expect much difference in our results, because capillary flux affects mainly water-limited semiarid regions (Niu et al., 2007). Future works should focus on developing a new model parameterization to address the omission of the capillary rise in the model. LSMs usually ignore the capillary flux (Niu et al., 2007), while only in a few model developments with a new groundwater parameterization capillary flux is explicitly represented (e.g., Niu et al., 2007; Koirala et al., 2014; Vergnes et al., 2014). Introducing the representation of capillary flux to JULES-GFB could show wetter soil domain, especially in arid areas, with higher evapotranspiration values.
- The explicit solver of JULES requires improvement because it is generally fit for very small timesteps. In the explicit equations, variables are calculated directly from the

previous time step without iterations for convergence of the solution. To achieve stability in the solution, it is crucial to keep the timestep small, otherwise fluctuations will occur, and they will be amplified, as the model continues running. Furthermore, the model needs fine vertical discretization to achieve robustness and avoid oscillations. Probably, the usage of an implicit solver could make the model more robust to numerical oscillations and decrease the computational time by increasing the time step and the vertical discretization.

- The lateral subsurface flow in the unsaturated zone (within JULES soil domain) is not considered in JULES. The only instance when this happens is when the water table reaches the soil domain. However, previous studies (e.g., Miguez-Macho et al., 2007; Maxwell and Kollet, 2008) have shown the importance of vertical and lateral redistribution of soil moisture in the unsaturated zone (Hazenberg et al., 2015). The implementation of a three-dimensional flow based on Richards equation in the soil domain could be subject of future research. The new parameterization will require significant computer power, but it will improve the realism of the model because it will account for subgrid variation of infiltration capacity and topography (Hazenberg et al., 2015). The use of synthetic experiments could reveal the impact of the new parameterization, especially the cases (e.g., geology, initial conditions, climate) in which the 3-D representation has significant impact.
- The river routing scheme (RFM) is based on globally constant parameters ignoring topography and the land cover type. In our preliminary analysis, we replaced the river network scheme with the Manning equation. We found minor differences between the RFM and the Manning scheme suggesting low impact of river routing scheme on runoff predictions. However, we applied this change only in two small catchments with similar climatological conditions. Future works should select a bigger sample size to give a better picture of the impact of a new river routing approach. A comprehensive sensitivity and uncertainty analysis should assess the impact of each parameter.
- Finally, there are a few other recommendations related to the hydrological parameterization in JULES that need further investigation: (1) the effect of layer thickness on water partition and telescoping; (2) the effect of different soil water stress (beta) parameterisations; (3) the effect of different pedotransfer functions for calculation of hydraulic parameters (JULES uses the Cosby equations which are not suitable for the Van Genuchten model, personal communication with Anne Verhoef)(; (4) The effect of seasonally varying plant parameters (LAI, albedo).

# BIBLIOGRAPHY

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'connell, P.E. and Rasmussen, J., 1986. An introduction to the European Hydrological System—Systeme Hydrologique Europeen,"SHE", 2: Structure of a physically-based, distributed modelling system. Journal of hydrology, 87(1-2), pp.61-77.

Alley, W.M., Healy, R.W., LaBaugh, J.W. and Reilly, T.E., 2002. Flow and storage in groundwater systems. science, 296(5575), pp.1985-1990.

Aquanty, Inc., 2015. HydroGeoSphere User Manual, 435 pp., Waterloo, Ont.

Arnell, N.W., 2004. Climate-change impacts on river flows in Britain: the UKCIPO2 scenarios. Water and Environment Journal, 18(2), pp.112-117.

Arnell, N., 2011. Uncertainty in the relationship between climate forcing and hydrological response in UK catchments. Hydrology and Earth System Sciences, 15(3), pp.897-912.

Ashby, S.F. and Falgout, R.D., 1996. A parallel multigrid preconditioned conjugate gradient algorithm for groundwater flow simulations. Nuclear Science and Engineering, 124(1), pp.145-159.

Bakopoulou, C., 2015. Critical assessment of structure and parameterization of JULES land surface model at different spatial scales in a UK Chalk catchment.

Ball, J.T., Woodrow, I.E. and Berry, J.A., 1987. A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. In Progress in photosynthesis research (pp. 221-224). Springer, Dordrecht.

Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M. and Betts, A.K., 2009. A revised hydrology for the ECMWF model: Verification from field site to

terrestrial water storage and impact in the Integrated Forecast System. Journal of hydrometeorology, 10(3), pp.623-643.

Beck, H.E., van Dijk, A.I., de Roo, A., Dutra, E., Fink, G., Orth, R. and Schellekens, J., 2017. Global evaluation of runoff from ten state-of-the-art hydrological models. Hydrology and Earth System Sciences, 21(6), pp.2881-2903.

Bell, V.A., Kay, A.L., Cole, S.J., Jones, R.G., Moore, R.J. and Reynard, N.S., 2012. How might climate change affect river flows across the Thames Basin? An area-wide analysis using the UKCP09 Regional Climate Model ensemble. Journal of Hydrology, 442, pp.89-104.

Bell, V.A., Kay, A.L., Jones, R.G. and Moore, R.J., 2007. Development of a high-resolution grid-based river flow model for use with regional climate model output. Hydrology and Earth System Sciences Discussions, 11(1), pp.532-549.

Best, M.J., Pryor, M., Clark, D.B., Rooney, G.G., Essery, R., Ménard, C.B., Edwards, J.M., Hendry, M.A., Porson, A., Gedney, N. and Mercado, L.M., 2011. The Joint UK Land Environment Simulator (JULES), model description–Part 1: energy and water fluxes. Geoscientific Model Development, 4(3), pp.677-699.

Betts, R.A., Alfieri, L., Bradshaw, C., Caesar, J., Feyen, L., Friedlingstein, P., Gohar, L., Koutroulis, A., Lewis, K., Morfopoulos, C. and Papadimitriou, L., 2018. Changes in climate extremes, fresh water availability and vulnerability to food insecurity projected at 1.5 C and 2 C global warming with a higher-resolution global climate model. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376(2119), p.20160452.

Beven, K.J., 2001. Rainfall-runoff modelling: The Primer, John Wiley and Sons Press, Department of Geography Royal Holloway. University of London Egham, Surrey.

Beven, K., 2004. Robert E. Horton's perceptual model of infiltration processes. Hydrological processes, 18(17), pp.3447-3460.

Beven, K. and Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. Hydrological processes, 6(3), pp.279-298.

Beven, K.J. and Freer, J., 2001a. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems, Journal of Hydrology, 249, 11–29.

Beven, K.J. and Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrological Sciences Journal, 24(1), pp.43-69.

Bierkens, M.F., Bell, V.A., Burek, P., Chaney, N., Condon, L.E., David, C.H., de Roo, A., Döll, P., Drost, N., Famiglietti, J.S. and Flörke, M., 2015. Hyper-resolution global hydrological modelling: what is next? "Everywhere and locally relevant". Hydrological processes, 29(2), pp.310-320.

Bixio, A., Gambolati, G., Paniconi, C., Putti, M., Shestopalov, V., Bublias, V., Bohuslavsky, A., Kasteltseva, N. and Rudenko, Y., 2002. Modeling groundwater-surface water interactions including effects of morphogenetic depressions in the Chernobyl exclusion zone. Environmental Geology, 42(2-3), pp.162-177.

Black, T.L., 1994. The new NMC mesoscale Eta model: Description and forecast examples. Weather and forecasting, 9(2), pp.265-278.

Boorman, D.B., Hollis, J.M. and Lilly, A., 1995. Hydrology of soil types: a hydrologically-based classification of the soils of United Kingdom. Institute of Hydrology.

Brooks, R.H. and Corey, A.T., 1964. Hydraulic properties of porous media and their relation to drainage design. Transactions of the ASAE, 7(1), pp.26-0028.

Butts, M.B., Payne, J.T., Kristensen, M. and Madsen, H., 2004. An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. Journal of hydrology, 298(1-4), pp.242-266.

Camporese, M., Paniconi, C., Putti, M. and Orlandini, S., 2010. Surface-subsurface flow modeling with path-based runoff routing, boundary condition-based coupling, and assimilation of multisource observation data. Water Resources Research, 46(2).

Cao, F., Dan, L., Ma, Z. and Gao, T., 2020. Assessing the regional climate impact on terrestrial ecosystem over East Asia using coupled models with land use and land cover forcing during 1980–2010. Scientific Reports, 10(1), pp.1-15.

CCIRG (Climate Change Impacts Review Group), 1991. The Potential Effects of Climate Change in the United Kingdom. HMSO, London, UK.

Charlton, M.B. and Arnell, N.W., 2014. Assessing the impacts of climate change on river flows in England using the UKCP09 climate change projections. Journal of Hydrology, 519, pp.1723-1738.

Chow, V.T., Maidment, D.R. and Larry, W., 1988. Mays. Applied Hydrology. International edition, MacGraw-Hill, Inc, 149.

Christensen, J.H., Boberg, F., Christensen, O.B. and Lucas-Picher, P., 2008. On the need for bias correction of regional climate change projections of temperature and precipitation. Geophysical Research Letters, 35(20).

Christierson, B.V., Vidal, J.P. and Wade, S.D., 2012. Using UKCP09 probabilistic climate information for UK water resource planning. Journal of Hydrology, 424, pp.48-67.

Clark, D.B. and Gedney, N., 2008. Representing the effects of subgrid variability of soil moisture on runoff generation in a land surface model. Journal of Geophysical Research: Atmospheres, 113(D10).

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. and Boucher, O., 2011. The Joint UK Land Environment Simulator (JULES), model description–Part 2: carbon fluxes and vegetation dynamics. Geoscientific Model Development, 4(3), pp.701-722.

Clark, M.P., Fan, Y., Lawrence, D.M., Adam, J.C., Bolster, D., Gochis, D.J., Hooper, R.P., Kumar, M., Leung, L.R., Mackay, D.S. and Maxwell, R.M., 2015. Improving the representation of hydrologic processes in Earth System Models. Water Resources Research, 51(8), pp.5929-5956.

Cloke, H.L., Jeffers, C., Wetterhall, F., Byrne, T., Lowe, J. and Pappenberger, F., 2010. Climate impacts on river flow: projections for the Medway catchment, UK, with UKCP09 and CATCHMOD. Hydrological Processes, 24(24), pp.3476-3489.

Collier, N., Hoffman, F.M., Lawrence, D.M., Keppel-Aleks, G., Koven, C.D., Riley, W.J., Mu, M. and Randerson, J.T., 2018. The international land model benchmarking (ILAMB) system: design, theory, and implementation. Journal of Advances in Modeling Earth Systems, 10(11), pp.2731-2754.

Condon, L.E. and Maxwell, R.M., 2015. Evaluating the relationship between topography and groundwater using outputs from a continental-scale integrated hydrology model. Water Resources Research, 51(8), pp.6602-6621.

Coon, E.T., Moulton, J.D. and Painter, S.L., 2016. Managing complexity in simulations of land surface and near-surface processes. Environmental modelling & software, 78, pp.134-149.

Cooper, D.M., Wilkinson, W.B. and Arnell, N.W., 1995. The effects of climate changes on aquifer storage and river baseflow. Hydrological sciences journal, 40(5), pp.615-631.

Coulthard, T.J., Ramirez, J., Fowler, H.J. and Glenis, V., 2012. Using the UKCP09 probabilistic scenarios to model the amplified impact of climate change on drainage basin sediment yield. Hydrology and Earth System Sciences, 16(11), pp.4401-4416.

Cox, P.M., Betts, R.A., Bunton, C.B., Essery, R.L.H., Rowntree, P.R. and Smith, J., 1999. The impact of new land surface physics on the GCM simulation of climate and climate sensitivity. Climate Dynamics, 15(3), pp.183-203.

Cuthbert, M.O., Gleeson, T., Moosdorf, N., Befus, K.M., Schneider, A., Hartmann, J. and Lehner, B., 2019. Global patterns and dynamics of climate–groundwater interactions. Nature Climate Change, 9(2), pp.137-141.

Dadson, S.J. and Bell, V.A., 2010. Comparison of Grid-2-Grid and TRIP runoff routing schemes.

Dadson, S.J., Bell, V.A. and Jones, R.G., 2011. Evaluation of a grid-based river flow model configured for use in a regional climate model. Journal of hydrology, 411(3), pp.238-250.

Dadson, S., Blyth, E., Clark, D., Hughes, A., Hannaford, J., Lawrence, B., Polcher, J. and Reynard, N., 2019, January. Hydro-JULES: Next Generation Land-surface and Hydrological Predictions. In Geophysical Research Abstracts (Vol. 21).

Dandar, E., Saaltink, M.W., Carrera, J. and Nemer, B., 2017. A surface model for water and energy balance in cold regions accounting for vapor diffusion. Hydrology and Earth System Sciences Discussions, pp.1-28.

Danierhan, S., Abudu, S. and Donghai, G., 2012. Coupled GSI-SVAT model with groundwater-surface water interaction in the riparian zone of Tarim River. Journal of Hydrologic Engineering, 18(10), pp.1211-1218.

Dankers, R., Arnell, N.W., Clark, D.B., Falloon, P.D., Fekete, B.M., Gosling, S.N., Heinke, J., Kim, H., Masaki, Y., Satoh, Y. and Stacke, T., 2014. First look at changes in flood hazard in the Inter-Sectoral Impact Model Intercomparison Project ensemble. Proceedings of the National Academy of Sciences, 111(9), pp.3257-3261.

Davie, J.C.S., Falloon, P.D., Kahana, R., Dankers, R., Betts, R., Portmann, F.T., Wisser, D., Clark, D.B., Ito, A., Masaki, Y. and Nishina, K., 2013. Comparing projections of future

changes in runoff from hydrological and biome models in ISI-MIP. Earth System Dynamics, 4, pp.359-374.

de Goncalves, L.G.G., Restrepo-Coupe, N., da Rocha, H., Saleska, S. and Stockli, R., 2008. The Large Scale Biosphere-Atmosphere Experiment in Amazônia, Model Intercomparison Project (LBA-MIP) protocol. LBA-MIP website: http://www. climatemodeling. org/lba-mip/March, 19.

de Graaf, I.D., Sutanudjaja, E.H., Van Beek, L.P.H. and Bierkens, M.F.P., 2015. A high-resolution global-scale groundwater model. Hydrology and Earth System Sciences, 19(2), pp.823-837.

de Graaf, I.E., van Beek, R.L., Gleeson, T., Moosdorf, N., Schmitz, O., Sutanudjaja, E.H. and Bierkens, M.F., 2017. A global-scale two-layer transient groundwater model: Development and application to groundwater depletion. Advances in water Resources, 102, pp.53-67.

Deardorff, J.W., 1978. Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation. Journal of Geophysical Research: Oceans, 83(C4), pp.1889-1903.

Dickinson, R.E., 1986. Biosphere/atmosphere transfer scheme (BATS) for the NCAR community climate model. Technical report, NCAR.

Dickinson, R.E., 1984. Modeling evapotranspiration for three-dimensional global climate models. Climate processes and climate sensitivity, 29, pp.58-72.

Dingman, S.L., 2015. Physical hydrology. Waveland press.

Dooge, J.C., 2001. Integrated management of water resources. In Understanding the Earth System (pp. 115-123). Springer, Berlin, Heidelberg.

Ducharne, A., 2016. The hydrol module of ORCHIDEE: Scientific documentation.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K. and Liebert, J., 2012. HESS Opinions" Should we apply bias correction to global and regional climate model data?". Hydrology and Earth System Sciences, 16(9), pp.3391-3404.

Endrizzi, S., Gruber, S., Dall'Amico, M. and Rigon, R., 2014. GEOtop 2.0: simulating the combined energy and water balance at and below the land surface accounting for soil freezing, snow cover and terrain effects. Geoscientific Model Development, 7(6), pp.2831-2857.

Environment Agency, 2005. CATCHMOD: Conceptual Rainfall Runoff Model. Technical User Guide and Software Manual for CATCHMOD v4.03 . Environment Agency: UK.

Environment Agency, 2006. Underground, Under Threat: The State of Groundwater in England and Wales.

Environment Agency, 2014. The stage of groundwater in England and Wales.

Famiglietti, J.S., Lo, M., Ho, S.L., Bethune, J., Anderson, K.J., Syed, T.H., Swenson, S.C., De Linage, C.R. and Rodell, M., 2011. Satellites measure recent rates of groundwater depletion in California's Central Valley. Geophysical Research Letters, 38(3).

Fan, Y., Li, H. and Miguez-Macho, G., 2013. Global patterns of groundwater table depth. Science, 339(6122), pp.940-943.

Fan, Y. and Miguez-Macho, G., 2011. A simple hydrologic framework for simulating wetlands in climate and earth system models. Climate dynamics, 37(1-2), pp.253-278.

Ferguson, I.M. and Maxwell, R.M., 2010. Role of groundwater in watershed response and land surface feedbacks under climate change. Water Resources Research, 46(10).

Forzieri, G., Alkama, R., Miralles, D.G. and Cescatti, A., 2017. Satellites reveal contrasting responses of regional climate to the widespread greening of Earth. Science, 356(6343), pp.1180-1184.

Fowler, H.J. and Kilsby, C.G., 2007. Using regional climate model data to simulate historical and future river flows in northwest England. Climatic Change, 80(3-4), pp.337-367.

Freeze, R. A. and Cherry, J.A., 1979. Groundwater, 604 pp., Prentice-Hall, Englewood Cliffs, N. J

Fung, F. and Gawith, M., 2018. UKCP18 for UKCP09 Users, UKCP18 Guidance. Met Office

Fung F, Lowe J, Mitchell JFB, Murphy J, Bernie D, Gohar L, Harris G, Howard T, Kendon E, Maisey P, Palmer M and Sexton D (2018). UKCP18 Guidance: How to use the UKCP18 Land Projections. Met Office Hadley Centre, Exeter.

Gangrade, S., Kao, S. and McManamay, R.A., 2020. Multi-model Hydroclimate Projections for the Alabama-Coosa-Tallapoosa River Basin in the Southeastern United States. Sci Rep 10, 2870.

Ganji, A. and Sushama, L., 2018. Improved representation of surface-groundwater interaction in the Canadian land surface scheme. International Journal of Climatology.

Gedney, N. and Cox, P.M., 2003. The sensitivity of global climate model simulations to the representation of soil moisture heterogeneity. Journal of Hydrometeorology, 4(6), pp.1265-1275.

Giuntoli, I., Vidal, J.P., Prudhomme, C. and Hannah, D.M., 2015a. Future hydrological extremes: the uncertainty from multiple global climate and global hydrological models.

Giuntoli, I., Villarini, G., Prudhomme, C., Mallakpour, I. and Hannah, D.M., 2015b. Evaluation of global impact models' ability to reproduce runoff characteristics over the central United States. Journal of Geophysical Research: Atmospheres, 120(18), pp.9138-9159.

Givati, A., Gochis, D., Rummler, T. and Kunstmann, H., 2016. Comparing one-way and two-way coupled hydrometeorological forecasting systems for flood forecasting in the Mediterranean region. Hydrology, 3(2), p.19.

Gleeson, T., Wagener, T., Doell, P., Bierkens, M., Wada, Y., Lo, M.H., Taylor, R., Rahman, S., Rosolem, R., Hill, M. and West, C., 2019. Groundwater representation in continental to global hydrologic models: a call for open and holistic evaluation, conceptualization and classification.

Gleick, P.H., 1993. Water in crisis. Pacific Institute for Studies in Dev., Environment & Security. Stockholm Env. Institute, Oxford Univ. Press. 473p, 9

Gochis, D.J., W. Yu, D.N. Yates, 2015: The WRF-Hydro model technical description and user's guide, version 3.0. NCAR Technical Document. 120 pages. Available online at: http://www.ral.ucar.edu/projects/wrf_hydro/.

Green, W.H. and Ampt, G.A., 1911. Studies on Soil Physics. The Journal of Agricultural Science, 4(1), pp.1-24.

Griffin, A., Young, A. and Stewart, L., 2019. Revising the BFIHOST catchment descriptor to improve UK flood frequency estimates. Hydrology Research, 50(6), pp.1508-1519.

Gudmundsson, L., Tallaksen, L.M., Stahl, K., Clark, D.B., Dumont, E., Hagemann, S., Bertrand, N., Gerten, D., Heinke, J., Hanasaki, N. and Voss, F., 2012. Comparing large-scale hydrological model simulations to observed runoff percentiles in Europe. Journal of Hydrometeorology, 13(2), pp.604-620.

Gudmundsson, L., Wagener, T., Tallaksen, L.M. and Engeland, K., 2012. Evaluation of nine large-scale hydrological models with respect to the seasonal runoff climatology in Europe. Water Resources Research, 48(11).

Guillod, B.P., Jones, R.G., Dadson, S.J., Coxon, G., Bussi, G., Freer, J., Kay, A.L., Massey, N.R., Sparrow, S.N., Wallom, D.C. and Allen, M.R., 2018. A large set of potential past, present and future hydro-meteorological time series for the UK. Hydrology and Earth System Sciences, 22(1), pp.611-634.

Gulden, L.E., Rosero, E., Yang, Z.L., Rodell, M., Jackson, C.S., Niu, G.Y., Yeh, P.J.F. and Famiglietti, J., 2007. Improving land-surface model hydrology: Is an explicit aquifer model better than a deeper soil profile?. Geophysical research letters, 34(9).

Guppy, L., Uyttendaele, P., Villholth, K.G. and Smakhtin, V., 2018. Groundwater and Sustainable Development Goals: Analysis of Interlinkages.

Gupta, H.V., Kling, H., Yilmaz, K.K. and Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. Journal of hydrology, 377(1-2), pp.80-91.

Gutenson, J.L., Tavakoly, A.A., Wahl, M.D. and Follum, M.L., 2019. Comparison of Generalized Non-Data-Driven Reservoir Routing Models for Global-Scale Hydrologic Modeling. Hydrology and Earth System Sciences Discussions, pp.1-41.

Gutowski, W.J., Vörösmarty, C.J., Person, M., Ötles, Z., Fekete, B. and York, J., 2002. A Coupled Land-Atmosphere Simulation Program (CLASP): Calibration and validation. Journal of Geophysical Research: Atmospheres, 107(D16).

Guzman, J.A., Moriasi, D.N., Gowda, P.H., Steiner, J.L., Starks, P.J., Arnold, J.G. and Srinivasan, R., 2015. A model integration framework for linking SWAT and MODFLOW. Environmental Modelling & Software, 73, pp.103-116.

Habets, F., Noilhan, J., Golaz, C., Goutorbe, J.P., Lacarrere, P., Leblois, E., Ledoux, E., Martin, E., Ottlé, C. and Vidal-Madjar, D., 1999. The ISBA surface scheme in a macroscale hydrological model applied to the Hapex-Mobilhy area: Part I: Model and database. Journal of Hydrology, 217(1-2), pp.75-96.

Haerter, J., Hagemann, S., Moseley, C. and Piani, C., 2011. Climate model bias correction and the role of timescales. Hydrology and Earth System Sciences, 15, pp.1065-1073.

Hagemann, S., Chen, C., Haerter, J.O., Heinke, J., Gerten, D. and Piani, C., 2011. Impact of a statistical bias correction on the projected hydrological changes obtained from three GCMs and two hydrology models. Journal of Hydrometeorology, 12(4), pp.556-578.

Hagemann, S., Chen, C., Clark, D., Folwell, S., Gosling, S.N., Haddeland, I., Hannasaki, N., Heinke, J., Ludwig, F., Voss, F. and Wiltshire, A., 2013. Climate change impact on available water resources obtained using multiple global climate and hydrology models. Earth System Dynamics, 4, pp.129-144.

Hallett, S.H., Sakrabani, R., Keay, C.A. and Hannam, J.A., 2017. Developments in land information systems: examples demonstrating land resource management capabilities and options. Soil use and management, 33(4), pp.514-529.

Hanasaki, N., Kanae, S., Oki, T., Masuda, K., Motoya, K., Shirakawa, N., Shen, Y. and Tanaka, K., 2008. An integrated model for the assessment of global water resources–Part 1: Model description and input meteorological forcing. Hydrology and Earth System Sciences, 12(4), pp.1007-1025.

Hari, P., Andreae, M.O., Kabat, P. and Kulmala, M., 2009. A comprehensive network of measuring stations to monitor climate change. Boreal environment research, 14, pp.442-446.

Harman, J., Gawith, M. and Colley, M., 2005. Progress on assessing climate impacts through the UK Climate Impacts Programme. Weather, 60(9), pp.258-262.

Harris, I.P.D.J., Jones, P.D., Osborn, T.J. and Lister, D.H., 2014. Updated high-resolution grids of monthly climatic observations–the CRU TS3. 10 Dataset. International Journal of Climatology, 34(3), pp.623-642.

Hay, L.E., Wilby, R.L. and Leavesley, G.H., 2000. A comparison of delta change and downscaled GCM scenarios for three mountainous basins in the United States 1. JAWRA Journal of the American Water Resources Association, 36(2), pp.387-397.

Hazenberg, P., Fang, Y., Broxton, P., Gochis, D., Niu, G.Y., Pelletier, J.D., Troch, P.A. and Zeng, X., 2015. A hybrid-3D hillslope hydrological model for use in E arth system models. Water Resources Research, 51(10), pp.8218-8239.

Hedger, M.M., Connell, R. and Bramwell, P., 2006. Bridging the gap: empowering decision-making for adaptation through the UK Climate Impacts Programme. Climate Policy, 6(2), pp.201-215.

Hemlin Söderberg, M., 2015. Measuring soil infiltrationrates in cultivated land: A case study of Ifakara, Tanzania.

Hempel, S., Frieler, K., Warszawski, L., Schewe, J. and Piontek, F., 2013. A trend-preserving bias correction–the ISI-MIP approach. Earth System Dynamics, 4(2), pp.219-236.

Herrera-Pantoja, M. and Hiscock, K.M., 2008. The effects of climate change on potential groundwater recharge in Great Britain. Hydrological Processes: An International Journal, 22(1), pp.73-86.

Holman, I.P., Tascone, D. and Hess, T.M., 2009. A comparison of stochastic and deterministic downscaling methods for modelling potential groundwater recharge under climate change in East Anglia, UK: implications for groundwater resource management. Hydrogeology Journal, 17(7), pp.1629-1641.

Hough, M.N. and Jones, R.J.A., 1997. The United Kingdom Meteorological Office rainfall and evaporation calculation system: MORECS version 2.0-an overview. Hydrology and Earth System Sciences Discussions, 1(2), pp.227-239.

Houldcroft, C.J., Grey, W.M., Barnsley, M., Taylor, C.M., Los, S.O. and North, P.R., 2009. New vegetation albedo parameters and global fields of soil background albedo derived from MODIS for use in a climate model. Journal of Hydrometeorology, 10(1), pp.183-198.

Huang, Z., Tang, Q., Lo, M.H., Liu, X., Lu, H., Zhang, X. and Leng, G., 2019. The influence of groundwater representation on hydrological simulation and its assessment using satellite-based water storage variation. Hydrological Processes, 33(8), pp.1218-1230.

Hulme, M., Turnpenny, J. and Jenkins, G.J., 2002. Climate change scenarios for the United Kingdom: the UKCIP02 briefing report (p. 15). Norwich, UK: Tyndall Centre for Climate Change Research, School of Environmental Sciences, University of East Anglia.

Hutchins, M.G., Abesser, C., Prudhomme, C., Elliott, J.A., Bloomfield, J.P., Mansour, M.M. and Hitt, O.E., 2018. Combined impacts of future land-use and climate stressors on water resources and quality in groundwater and surface waterbodies of the upper Thames river basin, UK. Science of the Total Environment, 631, pp.962-986.

Jackson, C.R., 2001. The development and validation of the object-oriented quasi three-dimensional regional groundwater model ZOOMQ3D.

Jackson, C.R., Meister, R. and Prudhomme, C., 2011. Modelling the effects of climate change and its uncertainty on UK Chalk groundwater resources from an ensemble of global climate model projections. Journal of Hydrology, 399(1-2), pp.12-28.

Jarvis, P.G., 1976. The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field. Philosophical Transactions of the Royal Society of London. B, Biological Sciences, 273(927), pp.593-610.

Jenkins, G.J., Murphy, J.M., Sexton, D.M., Lowe, J.A., Jones, P. and Kilsby, C.G., 2010. UK Climate Projections briefing report.

Johannes Dolman, A. and Gregory, D., 1992. The parametrization of rainfall interception in GCMs. Quarterly Journal of the Royal Meteorological Society, 118(505), pp.455-467.

Kay, A.L. and Jones, R.G., 2012. Comparison of the use of alternative UKCP09 products for modelling the impacts of climate change on flood frequency. Climatic Change, 114(2), pp.211-230.

Keune, J., Gasper, F., Goergen, K., Hense, A., Shrestha, P., Sulis, M. and Kollet, S., 2016. Studying the influence of groundwater representations on land surface-atmosphere feedbacks during the European heat wave in 2003. Journal of Geophysical Research: Atmospheres, 121(22).

Kim, J., Warnock, A., Ivanov, V.Y. and Katopodes, N.D., 2012. Coupled modeling of hydrologic and hydrodynamic processes including overland and channel flow. Advances in Water Resources, 37, pp.104-126.

Knoben, W.J., Woods, R.A. and Freer, J.E., 2018. A Quantitative Hydrological Climate Classification Evaluated With Independent Streamflow Data. Water Resources Research, 54(7), pp.5088-5109.

Koirala, S., Kim, H., Hirabayashi, Y., Kanae, S. and Oki, T., 2018. Sensitivity of global hydrological simulations to groundwater capillary flux parameterizations. Water Resources Research.

Koirala, S., Yeh, P.J.F., Hirabayashi, Y., Kanae, S. and Oki, T., 2014. Global-scale land surface hydrologic modeling with the representation of water table dynamics. Journal of Geophysical Research: Atmospheres, 119(1), pp.75-89.

Kollet, S.J. and Maxwell, R.M., 2006. Integrated surface–groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model. Advances in Water Resources, 29(7), pp.945-958.

Kollet, S.J. and Maxwell, R.M., 2008. Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model. Water Resources Research, 44(2).

Kollet, S., Sulis, M., Maxwell, R.M., Paniconi, C., Putti, M., Bertoldi, G., Coon, E.T., Cordano, E., Endrizzi, S., Kikinzon, E. and Mouche, E., 2017. The integrated hydrologic model

intercomparison project, IH-MIP2: A second set of benchmark results to diagnose integrated hydrology and feedbacks. Water Resources Research, 53(1), pp.867-890.

Konikow, L.F., 2011. Contribution of global groundwater depletion since 1900 to sea-level rise. Geophysical Research Letters, 38(17).

Koutroulis, A.G., 2019. Dryland changes under different levels of global warming. Science of The Total Environment, 655, pp.482-511.

Koutroulis, A.G., Papadimitriou, L.V., Grillakis, M.G., Tsanis, I.K., Warren, R. and Betts, R.A., 2019. Global water availability under high-end climate change: A vulnerability based assessment. Global and planetary change, 175, pp.52-63.

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S. and Prentice, I.C., 2005. A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. Global Biogeochemical Cycles, 19(1).

Kumar, R., Shankar, V. and Jat, M.K., 2015. Evaluation of root water uptake models–a review. ISH Journal of Hydraulic Engineering, 21(2), pp.115-124.

Kurtz, W., He, G., Kollet, S.J., Maxwell, R.M., Vereecken, H. and Franssen, H.J.H., 2016. TerrSysMP-PDAF (version 1.0): a modular high-performance data assimilation framework for an integrated land surface-subsurface model. Geoscientific Model Development, 9(4).

Labat, D., 2006. Oscillations in land surface hydrological cycle. Earth and Planetary Science Letters, 242(1-2), pp.143-154.

Lafon, T., Dadson, S., Buys, G. and Prudhomme, C., 2013. Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods. International Journal of Climatology, 33(6), pp.1367-1381.

Lane, R.A., Coxon, G., Freer, J.E., Wagener, T., Johnes, P.J., Bloomfield, J.P., Greene, S., Macleod, C.J. and Reaney, S.M., 2019. Benchmarking the predictive capability of hydrological models for river flow and flood peak predictions across over 1000 catchments in Great Britain. Hydrology and Earth System Sciences, 23(10), pp.4011-4032.

Largeron, C., Cloke, H.L., Verhoef, A., Martinez-de la Torre, A. and Mueller, A., 2018. Impact of the representation of the infiltration on the river flow during intense rainfall events in JULES. European Centre for Medium-Range Weather Forecasts.

Le Moigne, P., Boone, A., Calvet, J.C., Decharme, B., Faroux, S., Gibelin, A.L., Lebeaupin, C., Mahfouf, J.F., Martin, E., Masson, V. and Mironov, D., 2009. SURFEX scientific documentation. Note de centre (CNRM/GMME), Météo-France, Toulouse, France.

Le Vine, N., Butler, A., McIntyre, N. and Jackson, C., 2016. Diagnosing hydrological limitations of a land surface model: application of JULES to a deep-groundwater chalk basin. Hydrology and Earth System Sciences, 20(1), pp.143-159.

Leander, R. and Buishand, T.A., 2007. Resampling of regional climate model output for the simulation of extreme river flows. Journal of Hydrology, 332(3-4), pp.487-496.

Ledoux, E., Girard, G., De Marsily, G., Villeneuve, J.P. and Deschenes, J., 1989. Spatially distributed modeling: conceptual approach, coupling surface water and groundwater. In Unsaturated Flow in Hydrologic Modeling (pp. 435-454). Springer, Dordrecht.

Lee, K.T. and Chang, C.H., 2005. Incorporating subsurface-flow mechanism into geomorphology-based IUH modeling. Journal of Hydrology, 311(1), pp.91-105.

Lehner, B., Döll, P., Alcamo, J., Henrichs, T. and Kaspar, F., 2006. Estimating the impact of global change on flood and drought risks in Europe: a continental, integrated analysis. Climatic Change, 75(3), pp.273-299.

Lehner, B., Verdin, K. and Jarvis, A., 2008. New global hydrography derived from spaceborne elevation data. Eos, Transactions American Geophysical Union, 89(10), pp.93-94.

Leng, G., Tang, Q. and Rayburg, S., 2015. Climate change impacts on meteorological, agricultural and hydrological droughts in China. Global and Planetary Change, 126, pp.23-34.

Lewis, M.A., Cheney, C.S. and O Dochartaigh, B.E., 2006. Guide to permeability indices.

Liang, L. and Liu, Q., 2014. Streamflow sensitivity analysis to climate change for a large water-limited basin. Hydrological Processes, 28(4), pp.1767-1774.

Liang, X., Lettenmaier, D.P., Wood, E.F. and Burges, S.J., 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. Journal of Geophysical Research: Atmospheres, 99(D7), pp.14415-14428.

Liang, X., Xie, Z. and Huang, M., 2003. A new parameterization for surface and groundwater interactions and its impact on water budgets with the variable infiltration capacity (VIC) land surface model. Journal of Geophysical Research: Atmospheres, 108(D16).

Lo, M.H. and Famiglietti, J.S., 2010. Effect of water table dynamics on land surface hydrologic memory. Journal of Geophysical Research: Atmospheres, 115(D22).

Luo, Y., OuYang, Z., Yuan, G., Tang, D. and Xie, X., 2003. Evaluation of macroscopic root water uptake models using lysimeter data. Transactions of the ASAE, 46(3), p.625.

MacKellar, N.C., Dadson, S.J., New, M. and Wolski, P., 2013. Evaluation of the JULES land surface model in simulating catchment hydrology in Southern Africa. Hydrology and Earth System Sciences Discussions, 10(8), pp.11093-11128.

Mahrt, L. and Pan, H., 1984. A two-layer model of soil hydrology. Boundary-Layer Meteorology, 29(1), pp.1-20.

Manabe, S., 1969. Climate and the ocean circulation: I. The atmospheric circulation and the hydrology of the earth's surface. Monthly Weather Review, 97(11), pp.739-774.

Manley, R.E., 1977. The soil moisture component of mathematical catchment simulation models. Journal of Hydrology, 35(3-4), pp.341-356.

Mansour, M.M. and Hughes, A.G., 2004. User's manual for the distributed recharge model ZOODRM.

Martens, B., Waegeman, W., Dorigo, W.A., Verhoest, N.E. and Miralles, D.G., 2018. Terrestrial evaporation response to modes of climate variability. npj Climate and Atmospheric Science, 1(1), p.43.

Martínez-de la Torre, A., Blyth, E.M. and Weedon, G.P., 2019. Using observed river flow data to improve the hydrological functioning of the JULES land surface model (vn4. 3) used for regional coupled modelling in Great Britain (UKC2). Geoscientific Model Development, 12(2), pp.765-784.

Mathias, S.A., Butler, A.P., McIntyre, N. and Wheater, H.S., 2005. The significance of flow in the matrix of the Chalk unsaturated zone. Journal of Hydrology, 310(1-4), pp.62-77.

Maxwell, R.M., 2013. A terrain-following grid transform and preconditioner for parallel, large-scale, integrated hydrologic modeling. Advances in Water Resources, 53, pp.109-117.

Maxwell, R.M. and Kollet, S.J., 2008. Interdependence of groundwater dynamics and land-energy feedbacks under climate change. Nature Geoscience, 1(10), pp.665-669.

Maxwell, R.M., Kollet, S.J., Smith, S.G., Woodward, C.S., Falgout, R.D., Ferguson, I.M., Baldwin, C., Bosl, W.J., Hornung, R. and Ashby, S., 2009. ParFlow user's manual. International Ground Water Modeling Center Report GWMI, 1(2009), p.129p.

Maxwell, R.M., Lundquist, J.K., Mirocha, J.D., Smith, S.G., Woodward, C.S. and Tompson, A.F., 2011. Development of a coupled groundwater–atmosphere model. Monthly Weather Review, 139(1), pp.96-116.

Maxwell, R.M. and Miller, N.L., 2005. Development of a coupled land surface and groundwater model. Journal of Hydrometeorology, 6(3), pp.233-247

Maxwell, R.M., Condon, L.E. and Kollet, S.J., 2015. A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3. Geoscientific model development, 8(3), p.923.

Maxwell, R.M., Putti, M., Meyerhoff, S., Delfs, J.O., Ferguson, I.M., Ivanov, V., Kim, J., Kolditz, O., Kollet, S.J., Kumar, M. and Lopez, S., 2014. Surface-subsurface model intercomparison: A first set of benchmark results to diagnose integrated hydrology and feedbacks. Water resources research, 50(2), pp.1531-1549.

McCall, G.J.H. and Marker, B., 1989. Earth science mapping for planning, development, and conservation. Graham & Trotman.

McVicar, T.R., Roderick, M.L., Donohue, R.J., Li, L.T., Van Niel, T.G., Thomas, A., Grieser, J., Jhajharia, D., Himri, Y., Mahowald, N.M. and Mescherskaya, A.V., 2012. Global review and synthesis of trends in observed terrestrial near-surface wind speeds: Implications for evaporation. Journal of Hydrology, 416, pp.182-205.

Meenal, M. and Eldho, T.I., 2011. Simulation of groundwater flow in unconfined aquifer using meshfree point collocation method. Engineering Analysis with Boundary Elements, 35(4), pp.700-707.

Meinshausen, M., Smith, S.J., Calvin, K., Daniel, J.S., Kainuma, M.L.T., Lamarque, J.F., Matsumoto, K., Montzka, S.A., Raper, S.C.B., Riahi, K. and Thomson, A.G.J.M.V., 2011. The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. Climatic change, 109(1-2), p.213.

Met Office, 2018a. UKCP18 National Climate Projections. (Available online: https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/research/ukcp/ukcp18-overview-slidepack.ff-compressed.pdf).

Met Office, 2018b. UKCP18 Headline Findings. (Available online: https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/research/ukcp/ukcp18-headline-findings.pdf).

Met Office, 2018c. UKCP18 Guidance: Bias correction (Available online: https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/research/ukcp/ukcp18-guidance---how-to-bias-correct.pdf).

Miguez-Macho, G., Fan, Y., Weaver, C.P., Walko, R. and Robock, A., 2007. Incorporating water table dynamics in climate modeling: 2. Formulation, validation, and soil moisture simulation. Journal of Geophysical Research: Atmospheres, 112(D13).

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H. C., Meesters, A. G. C. A., & Dolman, A. J., 2011. Global land-surface evaporation estimated from satellite-based observations.

Mitchell, V.G., Cleugh, H.A., Grimmond, C.S. and Xu, J., 2008. Linking urban water balance and energy balance models to analyse urban design options. Hydrological Processes: An International Journal, 22(16), pp.2891-2900.

Monteith, J.L., 1965. Evaporation and environment. In Symposia of the society for experimental biology (Vol. 19, pp. 205-234). Cambridge University Press (CUP) Cambridge.

Moore, R.J., 1985. The probability-distributed principle and runoff production at point and basin scales. Hydrological Sciences Journal, 30(2), pp.273-297.

Moore, R.J., 2007. The PDM rainfall-runoff model. Hydrology and Earth System Sciences Discussions, 11(1), pp.483-499.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D. and Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE, 50(3), pp.885-900.

Morton, D., Rowland, C., Wood, C., Meek, L., Marston, C., Smith, G., Wadsworth, R. and Simpson, I., 2011. Final Report for LCM2007-the new UK land cover map. Countryside Survey Technical Report No 11/07.

Murphy, J.M., Sexton, D.M.H., Jenkins, G.J., Boorman, P.M., Booth, B.B.B., Brown, C.C., Clark, R.T., Collins, M., Harris, G.R., Kendon, E.J. and Betts, R.A., 2009. UK Climate

Projections Science Report: Climate change projections–Met Office Hadley Centre. Exeter, UK.

Nakicenovic, N., Alcamo, J., Davis, G., De Vries, B., Fenhann, J., Gaffin, S., Gregory, K., Grubler, A., Jung, T.Y., Kram, T. and La Rovere, E.L., 2000. IPCC Special Report on Emissions Scenarios, Cambridge.

Nash, J.E. and Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I—A discussion of principles. Journal of hydrology, 10(3), pp.282-290.

Nie, W., Zaitchik, B.F., Rodell, M., Kumar, S.V., Anderson, M.C. and Hain, C., Groundwater Withdrawals Under Drought: Reconciling GRACE and Land Surface Models in the United States High Plains Aquifer. Water Resources Research.

Niu, G.Y., Yang, Z.L., Dickinson, R.E., Gulden, L.E. and Su, H., 2007. Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data. Journal of Geophysical Research: Atmospheres, 112(D7).

Niu, G.Y., Yang, Z.L., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E. and Tewari, M., 2011. The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. Journal of Geophysical Research: Atmospheres, 116(D12).

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Flanner, M. G., Kluzek, E., Lawrence, P. J., Levis, S., Swenson, S. C., Thornton, P. E., Aiguo Dai, M. D., Dickinson, R., Feddema, J., Heald, C. L., Hoffman, F., Lamarque, J.F., Mahowald, N., Niu, G.Y., Qian, T., Randerson, J., Running, S., Sakaguchi, K., Slater, A., Stöckli, R., Wang, A., Yang, Z.-L., Zeng, X. and Zeng, X. 2010. Technical description of version 4.0 of the Community Land Model (CLM). NCAR Technical Note

Oleson, K.W. et al., 2013. Technical description of version 4.5 of the Community Land Model (CLM).

Overgaard, J., Rosbjerg, D. and Butts, M.B., 2006. Land-surface modelling in hydrological perspective? a review.

Painter, S.L., Coon, E.T., Atchley, A.L., Berndt, M., Garimella, R., Moulton, J.D., Svyatskiy, D. and Wilson, C.J., 2016. Integrated surface/subsurface permafrost thermal hydrology: Model formulation and proof-of-concept simulations. Water Resources Research, 52(8), pp.6062-6077.

Pan, H.L. and Mahrt, L., 1987. Interaction between soil hydrology and boundary-layer development. Boundary-Layer Meteorology, 38(1-2), pp.185-202.

Paniconi, C., Aldama, A.A. and Wood, E.F., 1991. Numerical evaluation of iterative and noniterative methods for the solution of the nonlinear Richards equation. Water Resources Research, 27(6), pp.1147-1163.

Papadimitriou, L.V., Koutroulis, A.G., Grillakis, M.G. and Tsanis, I.K., 2016. High-end climate change impact on European runoff and low flows-exploring the effects of forcing biases. Hydrology and Earth System Sciences, 20(5), p.1785.

Pappenberger, F., Cloke, H.L., Balsamo, G., Ngo-Duc, T. and Oki, T., 2010. Global runoff routing with the hydrological component of the ECMWF NWP system. International Journal of Climatology, 30(14), pp.2155-2174.

Pechlivanidis, I.G., Jackson, B.M., Mcintyre, N.R. and Wheater, H.S., 2011. Catchment scale hydrological modelling: a review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. Global NEST journal, 13(3), pp.193-214.

Pilling, C. and Jones, J.A.A., 1999. High resolution climate change scenarios: implications for British runoff. Hydrological Processes, 13(17), pp.2877-2895.

Pinder, G.F. and Bredehoeft, J.D., 1968. Application of the digital computer for aquifer evaluation. Water Resources Research, 4(5), pp.1069-1093.

Poméon, T., Diekkrüger, B. and Kumar, R., 2018. Computationally Efficient Multivariate Calibration and Validation of a Grid-Based Hydrologic Model in Sparsely Gauged West African River Basins. Water, 10(10), p.1418.

Potočki, K., 2018. Hydrology-theory and general concepts; Freshwater and pollution-sources and pathways of diffuse water pollutants

Price, A., Downing, R. A., and Edmunds, W. M., 1993. The chalk as an aquifer, in: The hydrogeology of the chalk of north-west Europe, edited by: Downing, R. A., Price, M., and Jones, G. P., Claredon Press, Oxford, 35–58.

Prickett, T.A. and Lonnquist, C.G., 1971. Selected digital computer techniques for groundwater resource evaluation. Bulletin (Illinois State Water Survey) no. 55.

Prudhomme, C., Dadson, S.J., Morris, D., Williamson, J., Goodsell, G., Crooks, S.M., Boelee, L., Davies, H.N., Buys, G., Lafon, T. and Watts, G., 2012. Future Flows Climate: an ensemble of 1-km climate change projections for hydrological application in Great Britain.

Prudhomme, C., Giuntoli, I., Robinson, E.L., Clark, D.B., Arnell, N.W., Dankers, R., Fekete, B.M., Franssen, W., Gerten, D., Gosling, S.N. and Hagemann, S., 2014. Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. Proceedings of the National Academy of Sciences, 111(9), pp.3262-3267.

Quante, M. and Matthias, V., 2006, December. Water in the Earth's atmosphere. In Journal de Physique IV (Proceedings) (Vol. 139, pp. 37-61). EDP sciences.

Rahman, M.,2015. Influence of Subsurface Hydrodynamics on the Lower Atmosphere at the Catchment Scale. PhD diss., Universitäts-und Landesbibliothek Bonn.

Rahman, M. and Rosolem, R., 2017. Towards a simple representation of chalk hydrology in land surface modelling. Hydrology and Earth System Sciences, 21(1), p.459.

Rahman, M., Rosolem, R., Kollet, S.J. and Wagener, T., 2019. Towards a computationally efficient free-surface groundwater flow boundary condition for large-scale hydrological modelling. Advances in Water Resources, 123, pp.225-233.

Rajib, M.A., Merwade, V. and Yu, Z., 2016. Multi-objective calibration of a hydrologic model using spatially distributed remotely sensed/in-situ soil moisture. Journal of hydrology, 536, pp.192-207.

Ratzlaff, J.R., 1994. Mean annual precipitation, runoff, and runoff ratio for Kansas, 1971-1990. Transactions of the Kansas Academy of Science (1903), pp.94-101.

Reinecke, R., Foglia, L., Mehl, S., Trautmann, T., Cáceres, D. and Döll, P., 2019. Challenges in developing a global gradient-based groundwater model (G 3 M v1. 0) for the integration into a global hydrological model. Geoscientific Model Development, 12(6), pp.2401-2418.

Rigon, R., Bertoldi, G. and Over, T.M., 2006. GEOtop: A distributed hydrological model with coupled water and energy budgets. Journal of Hydrometeorology, 7(3), pp.371-388.

Robinson, E. L., Blyth, E. M., Clark, D. B., Finch, J., and Rudd, A. C.: Trends in evaporative demand in Great Britain using high-resolution meteorological data, Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2015-520, in review, 2016.

Rodell, M., Houser, P.R., Jambor, U.E.A., Gottschalck, J., Mitchell, K., Meng, C.J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M. and Entin, J.K., 2004. The global land data assimilation system. Bulletin of the American Meteorological Society, 85(3), pp.381-394.

Saraiva Okello, A.M., Masih, I., Uhlenbrook, S., Jewitt, G.P. and Van der Zaag, P., 2018. Improved process representation in the simulation of the hydrology of a meso-scale semi-arid catchment. Water, 10(11), p.1549.

Schaap, M.G., Leij, F.J. and Van Genuchten, M.T., 2001. Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. Journal of hydrology, 251(3-4), pp.163-176.

Schaefli, B. and Gupta, H.V., 2007. Do Nash values have value?. Hydrological Processes: An International Journal, 21(15), pp.2075-2080.

Schellekens, J., Dutra, E., Martínez-de la Torre, A., Balsamo, G., van Dijk, A., Weiland, F.S., Minvielle, M., Calvet, J.C., Decharme, B., Eisner, S. and Fink, G., 2017. A global water resources ensemble of hydrological models: the eartH2Observe Tier-1 dataset. Earth System Science Data, 9(2), pp.389-413.

Schewe, J., Heinke, J., Gerten, D., Haddeland, I., Arnell, N.W., Clark, D.B., Dankers, R., Eisner, S., Fekete, B.M., Colón-González, F.J. and Gosling, S.N., 2014. Multimodel assessment of water scarcity under climate change. Proceedings of the National Academy of Sciences, 111(9), pp.3245-3250.

Schlemmer, L., Schär, C., Lüthi, D. and Strebel, L., 2018. A groundwater and runoff formulation for weather and climate models. Journal of Advances in Modeling Earth Systems, 10(8), pp.1809-1832.

Schober, P., Boer, C. and Schwarte, L.A., 2018. Correlation coefficients: appropriate use and interpretation. Anesthesia & Analgesia, 126(5), pp.1763-1768.

Scibek, J. and Allen, D.M., 2006. Modeled impacts of predicted climate change on recharge and groundwater levels. Water Resources Research, 42(11).

Sellers, P.J., Mintz, Y.C.S.Y., Sud, Y.E.A. and Dalcher, A., 1986. A simple biosphere model (SiB) for use within general circulation models. Journal of the Atmospheric Sciences, 43(6), pp.505-531.

Sene, K.J., Gash, J.H.C. and McNeil, D.D., 1991. Evaporation from a tropical lake: comparison of theory with direct measurements. Journal of Hydrology, 127(1-4), pp.193-217.

Shrestha, P., Sulis, M., Masbou, M., Kollet, S. and Simmer, C., 2014. A scale-consistent terrestrial systems modeling platform based on COSMO, CLM, and ParFlow. Monthly weather review, 142(9), pp.3466-3483.

Shuttleworth, W.J., 2012. Terrestrial hydrometeorology. John Wiley & Sons.

Siebert, S., Burke, J., Faures, J.M., Frenken, K., Hoogeveen, J., Döll, P. and Portmann, F.T., 2010. Groundwater use for irrigation–a global inventory. Hydrology and Earth System Sciences, 14(10), pp.1863-1880.

Solheim, A.L., Globevnik, L., Austnes, K., Kristensen, P., Moe, S.J., Persson, J., Phillips, G., Poikane, S., van de Bund, W. and Birk, S., 2019. A new broad typology for rivers and lakes in Europe: Development and application for large-scale environmental assessments. Science of the Total Environment, 697, p.134043.

Sorribas, M.V., Paiva, R.C., Melack, J.M., Bravo, J.M., Jones, C., Carvalho, L., Beighley, E., Forsberg, B. and Costa, M.H., 2016. Projections of climate change effects on discharge and inundation in the Amazon basin. Climatic change, 136(3-4), pp.555-570.

Sterling, S.M., Ducharne, A. and Polcher, J., 2013. The impact of global land-cover change on the terrestrial water cycle. Nature Climate Change, 3(4), pp.385-390.

Stoll, S., Hendricks Franssen, H.J., Butts, M. and Kinzelbach, W.K., 2011. Analysis of the impact of climate change on groundwater related hydrological fluxes: a multi-model approach including different downscaling methods. Hydrology and Earth System Sciences, 15(1), pp.21-38.

Sulis, M., Meyerhoff, S.B., Paniconi, C., Maxwell, R.M., Putti, M. and Kollet, S.J., 2010. A comparison of two physics-based numerical models for simulating surface water–groundwater interactions. Advances in Water Resources, 33(4), pp.456-467.

Sutanudjaja, E.H., Beek, R.V., Wanders, N., Wada, Y., Bosmans, J.H., Drost, N., Ent, R.J., De Graaf, I.E., Hoch, J.M., Jong, K.D. and Karssenberg, D., 2018. PCR-GLOBWB 2: a 5 arcmin global hydrological and water resources model. Geoscientific Model Development, 11(6), pp.2429-2453.

Takata, K., Emori, S. and Watanabe, T., 2003. Development of the minimal advanced treatments of surface interaction and runoff. Global and planetary Change, 38(1-2), pp.209-222.

Tan, T.S., Phoon, K.K. and Chong, P.C., 2004. Numerical study of finite element method based solutions for propagation of wetting fronts in unsaturated soil. Journal of geotechnical and geoenvironmental engineering, 130(3), pp.254-263.

Tavakoly, A.A., Snow, A.D., David, C.H., Follum, M.L., Maidment, D.R. and Yang, Z.L., 2017. Continental-scale river flow modeling of the Mississippi River Basin using high-resolution NHDPlus dataset. JAWRA Journal of the American Water Resources Association, 53(2), pp.258-279.

Teutschbein, C. and Seibert, J., 2012. Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. Journal of Hydrology, 456, pp.12-29.

Thomson, N., Barrie, I.A. and Ayles, M., 1981. The Meteorological Office Rainfall and Evaporation Calculation System MORECS. 45, UK Met.

Tian, J., Liu, J., Wang, Y., Wang, W., Li, C. and Hu, C., 2020. A coupled atmospheric-hydrologic modeling system with variable grid sizes for rainfall-runoff simulation in semi-humid and semi-arid watersheds: How does the coupling scale affects the results?. Hydrology and Earth System Sciences Discussions, pp.1-36.

Tian, W., Li, X., Wang, X.S. and Hu, B.X., 2012. Coupling a groundwater model with a land surface model to improve water and energy cycle simulation. Hydrology and Earth System Sciences Discussions, 9(1), pp.1163-1205.

Tijerina, D.T., 2018. Improving US national water modeling: an intercomparison of two high-resolution, continental scale models, ParFlow-CONUS and the National Water Model v1. 2 configuration of WRF-Hydro (Doctoral dissertation, Colorado School of Mines. Arthur Lakes Library).

Todini, E., 1996. The ARNO rainfall—runoff model. Journal of hydrology, 175(1-4), pp.339-382.

Trenberth, K.E., 2011. Changes in precipitation with climate change. Climate Research, 47(1-2), pp.123-138.

Van Genuchten, M.T., 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil science society of America journal, 44(5), pp.892-898.

Van Huijgevoort, M.H.J., Van Lanen, H.A.J., Teuling, A.J. and Uijlenhoet, R., 2014. Identification of changes in hydrological drought characteristics from a multi-GCM driven ensemble constrained by observed discharge. Journal of Hydrology, 512, pp.421-434.

Vergnes, J.P. and Decharme, B., 2012. A simple groundwater scheme in the TRIP river routing model: global off-line evaluation against GRACE terrestrial water storage estimates and observed river discharges. Hydrology and Earth System Sciences, 16(10), p.3889.

Vergnes, J.P., Decharme, B. and Habets, F., 2014. Introduction of groundwater capillary rises using subgrid spatial variability of topography into the ISBA land surface model. Journal of Geophysical Research: Atmospheres, 119(19), pp.11-065.

Verseghy, D.L., 1991. CLASS—A Canadian land surface scheme for GCMs. I. Soil model. International Journal of Climatology, 11(2), pp.111-133.

Vetter, T., Reinhardt, J., Flörke, M., van Griensven, A., Hattermann, F., Huang, S., Koch, H., Pechlivanidis, I.G., Plötner, S., Seidou, O. and Su, B., 2017. Evaluation of sources of uncertainty in projected hydrological changes under climate change in 12 large-scale river basins. Climatic Change, 141(3), pp.419-433.

Villarini, G., Mandapaka, P.V., Krajewski, W.F. and Moore, R.J., 2008. Rainfall and sampling uncertainties: A rain gauge perspective. Journal of Geophysical Research: Atmospheres, 113(D11).

Wang, H.F. and Anderson, M.P., 1982. Introduction to groundwater modelling. Finite difference and finite element methods. Academic Press Inc., 237 p.

Weedon, G.P., Gomes, S., Viterbo, P., Shuttleworth, W.J., Blyth, E., Österle, H., Adam, J.C., Bellouin, N., Boucher, O. and Best, M., 2011. Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century. Journal of Hydrometeorology, 12(5), pp.823-848.

Weedon, G.P., Prudhomme, C., Crooks, S., Ellis, R.J., Folwell, S.S. and Best, M.J., 2015. Evaluating the performance of hydrological models via cross-spectral analysis: case study of the Thames Basin, United Kingdom. Journal of Hydrometeorology, 16(1), pp.214-231.

Weill, S., Mouche, E. and Patin, J., 2009. A generalized Richards equation for surface/subsurface flow modelling. Journal of Hydrology, 366(1-4), pp.9-20.

Wheater, H.S., Peach, D. and Binley, A., 2007. Characterising groundwater-dominated lowland catchments: the UK Lowland Catchment Research Programme (LOCAR).

Wilby, R.L., 2006. When and where might climate change be detectable in UK river flows?. Geophysical Research Letters, 33(19).

Wild, M. and Liepert, B., 2010. The Earth radiation balance as driver of the global hydrological cycle. Environmental Research Letters, 5(2), p.025203.

Wilkinson, W.B. and Cooper, D.M., 1993. The response of idealized aquifer/river systems to climate change. Hydrological sciences journal, 38(5), pp.379-390.

Williams, K. and Clark, D., 2014. Disaggregation of daily data in JULES.

Wood, A.W., Leung, L.R., Sridhar, V. and Lettenmaier, D.P., 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Climatic change, 62(1-3), pp.189-216.

Wootten, A., Terando, A., Reich, B.J., Boyles, R.P. and Semazzi, F., 2017. Characterizing sources of uncertainty from global climate models and downscaling techniques. Journal of Applied Meteorology and Climatology, 56(12), pp.3245-3262.

Yang, W.Y., Li, D., Sun, T. and Ni, G.H., 2015. Saturation-excess and infiltration-excess runoff on green roofs. Ecological Engineering, 74, pp.327-336.

Yeh, P.J. and Eltahir, E.A., 2005. Representation of water table dynamics in a land surface scheme. Part I: Model development. Journal of climate, 18(12), pp.1861-1880.

Yeh, P.J. and Eltahir, E.A., 2005. Representation of water table dynamics in a land surface scheme. Part II: Subgrid variability. Journal of Climate, 18(12), pp.1881-1901.

York, J.P., Person, M., Gutowski, W.J. and Winter, T.C., 2002. Putting aquifers into atmospheric simulation models: An example from the Mill Creek Watershed, northeastern Kansas. Advances in Water Resources, 25(2), pp.221-238.

Young, A.R., 2006. Stream flow simulation within UK ungauged catchments using a daily rainfall-runoff model. Journal of Hydrology, 320(1-2), pp.155-172.

Zeng, Y., Xie, Z., Liu, S., Xie, J., Jia, B., Qin, P. and Gao, J., 2018. Global Land Surface Modeling Including Lateral Groundwater Flow. Journal of Advances in Modeling Earth Systems.

Zhang, K., Kimball, J.S. and Running, S.W., 2016. A review of remote sensing based actual evapotranspiration estimation. Wiley Interdisciplinary Reviews: Water, 3(6), pp.834-853.

Zulkafli, Z., Buytaert, W., Manz, B., Rosas, C.V., Willems, P., Lavado-Casimiro, W., Guyot, J.L. and Santini, W., 2016. Projected increases in the annual flood pulse of the Western Amazon. Environmental Research Letters, 11(1), p.014013.

Zulkafli, Z., Buytaert, W., Onof, C., Lavado, W. and Guyot, J.L., 2013. A critical assessment of the JULES land surface model hydrology for humid tropical environments. Hydrology and Earth System Sciences, 17(3), pp.1113-1132.

# APPENDIX: PHD RELATED ACTIVITIES

This appendix includes a synopsis of the academic activities performed during or before the PhD studies of the author.

## Refereed Publications

Batelis, S.C., Rahman, M., Kollet, S., Woods, R. and Rosolem, R., 2020. Towards the representation of groundwater in the Joint UK Land Environment Simulator. Hydrological Processes.

## Miscellaneous Publications

Markonis, Y., **Batelis, S.C.**, Dimakos, Y., Moschou, E. and Koutsoyiannis, D., 2017. Temporal and spatial variability of rainfall over Greece. Theoretical and Applied Climatology, 130(1-2), pp.217-232.

**Batelis, S.C.** and Nalbantis, I., 2014. Potential effects of forest fires on streamflow in the Enipeas river basin, Thessaly, Greece. Environmental Processes, 1(1), pp.73-85.

## Conference abstracts

**Batelis, S.C.**, Rahman, M., Rosolem, R., Woods, R. and Kollet, S., 2018, December. Assessing the impact of explicit representation of groundwater on the Joint UK Land Environment Simulator (JULES). In AGU General Assembly Conference Abstracts.

**Batelis, S.C.**, Rahman, M., Rosolem, R., Woods, R. and Han, D., 2018, April. Improving streamflow predictions in groundwater-dominated catchments using JULES. In EGU General Assembly Conference Abstracts (Vol. 20, p. 6251).

**Batelis, S.C.**, Rosolem, R., Han, D. and Rahman, M., 2017, April. Testing the Joint UK Land Environment Simulator (JULES) for flood forecasting. In EGU General Assembly Conference Abstracts (Vol. 19, p. 4118).

Rosolem, R., Wagener, T., Rahman, M., **Batelis, S.C.**, Kollet, S., Sarrazin, S., Hartmann, A. and Pianosi, F., 2018, December. Benefits of parsimonious representations of hydrological processes in Earth System Models. In AGU General Assembly Conference Abstracts.

Zhang, J, Rosolem, R., Rocha, H., **Batelis, S.C.** and Rahman, M., 2018. Exploration of drought propagation mechanisms and controls for the 2014-2017 Southeast Brazil drought. In AGU General Assembly Conference Abstracts.

## Miscellaneous conference abstracts

Dimakos, Y., Moschou, E., **Batelis, S.**, Markonis, Y. and Koutsoyiannis, D., 2014, May. Monthly rainfall trends in Greece (1950-2012). In EGU General Assembly Conference Abstracts (Vol. 16).

Moschou, E.C., **Batelis, S.C.**, Dimakos, Y., Fountoulakis, I., Markonis, Y., Papalexiou, S.M., Mamassis, N. and Koutsoyiannis, D., 2013. Spatial and temporal rainfall variability over Greece. In 5th EGU Leonardo conference–Hydrofractals.

## Other presentations

**S.C. Batelis** et al., Water resources, infrastructures and services: Privatization or municipalization? Athens. Presentation title “Privatization and remunicipalization of water services in Argentine”.

## Memberships

Student member of British Hydrological Society

## Teaching activities

[2018-2019] Fluids 3 (Year 3, Department of Mechanical Engineering)

[2016-2018] Water Resources Project (Year 3, Department of Civil Engineering)

[2016-2018] Engineering Hydrology and Open Channel Flow (Year 2, Department of Civil Engi-neering)

[2016] Fluid Mechanics (Year 2, Department of Civil Engineering)

## Awards

Fully funded PhD Scholarship from EPSRC (Engineering and Physical Sciences Research Council).