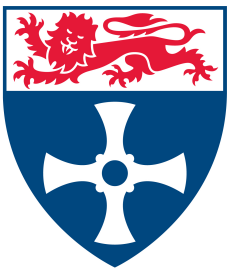
# A robust multi-purpose hydrological model for Great Britain



Elizabeth Lewis

*School of Civil Engineering and Geosciences*

*Newcastle University*

*A thesis presented for the degree of*

*Doctor of Philosophy*

April 2016

For Mum and Dad

# Abstract

Robust numerical models are an essential tool for informing flood and water management and policy around the world. Physically-based hydrological models have traditionally not been used for such applications due to prohibitively large data, time and computational resource requirements. Given recent advances in computing power and data availability, this study creates, for the first time, a robust, physically-based hydrological modelling system for Great Britain using the SHETRAN model and national datasets. Such a model has several advantages over less complex systems. Firstly, compared with conceptual models, a national physically-based model is more readily applicable to ungauged catchments, in which hydrological predictions are also required. Secondly, the results of a physically-based system may be more robust under changing conditions such as climate and land cover, as physical processes and relationships are explicitly accounted for. Finally, a fully integrated surface and subsurface model such as SHETRAN offers a wider range of applications compared with simpler schemes, such as assessments of groundwater resources, sediment transport and flooding from multiple sources.

In order to develop a national modelling system based on SHETRAN, a large array of data for the whole of Great Britain and the period 1960-2006 has been integrated into a framework that features a new, user-friendly graphical interface, which extracts and prepares the data required for a SHETRAN simulation of any catchment in Great Britain. This has vastly reduced the time it takes to set up and run a model from months to seconds. Structural changes have also been incorporated into SHETRAN to better represent lakes, handle pits in elevation data and accept gridded meteorological inputs. 306 catchments spanning Great Britain were then modelled using this system. The standard configuration of this system performs satisfactorily (NSE $> 0.5$) for 72% of catchments and well (NSE $> 0.7$) for 48%. Many of the remaining 28% of catchments that performed relatively poorly (NSE $< 0.5$) are located in the chalk in the south east of England. As such, the British Geological Survey 3D geology model for Great Britain (GB3D) has been incorporated for the first time in any hydrological model to pave the way for improvements to be made to simulations of catchments with important groundwater regimes. This coupling has involved development

of software to allow for easy incorporation of geological information into SHETRAN for any model setup. The addition of more realistic subsurface representation following this approach is shown to greatly improve model performance in areas dominated by groundwater processes.

The sensitivity of the modelling system to key inputs and parameters was tested, particularly with respect to the distribution and rates of rainfall and potential evapotranspiration. As part of this, a new national dataset of gridded hourly rainfall was created by disaggregating the 5km UK Climate Projections 2009 (UKCP09) gridded daily rainfall product with partially quality controlled hourly rain gauge data from over 1300 observation stations across the country. Of the sensitivity tests undertaken, the largest improvements in model performance were seen when this hourly gridded rainfall dataset was combined with potential evapotranspiration disaggregated to hourly intervals, with 61% of catchments showing an increase in NSE as a result of more realistic sub-daily meteorological forcing. Additional sensitivity analysis revealed that the slight over-estimation of runoff using the initial model configuration which has a median water balance bias of 5% was reduced in 62% of catchments by increasing daily potential evapotranspiration rates by 5%. Similarly, model performance was also found to improve by universally decreasing rainfall rates slightly, which together indicate the possibility of slight under-estimation of potential evapotranspiration derived from available data. In addition to extensive sensitivity testing, the national modelling system for Great Britain has also been coupled with the UKCP09 spatial weather generator to demonstrate the capability of the system to conduct climate change impact assessments. A set of 100 simulations for each of 20 representative catchments across the country were processed for a medium emissions scenario in the 2050s, in order to establish and demonstrate the methodology for conducting such an assessment. The results of these initial simulations suggest that higher potential evapotranspiration rates, combined with modest increases in rainfall under this climate change projection, lead to a general decrease in mean annual river flows. Changes in mean annual flow across the country vary between -26% to +8%, with the biggest reductions in flow found in the south of England and modest increases in runoff across Scotland.

This work represents a step-change in how the physically-based hydrological model SHETRAN can be used. Not only has this project made SHETRAN much easier to use on its own, but the model can now also be used in conjunction with external applications such as the UKCP09 spatial weather generator and GB3D. This means that the modelling system has great potential to be used as a resource at national, regional and local scales in an array of different applications, including climate change impact assessments, land cover change studies and integrated assessments of groundwater and surface water resources.

# Acknowledgements

Firstly I would like to thank my supervisors Chris Kilsby and Hayley Fowler for all of their excellent guidance, support and advice over the course of my research. Their inputs throughout the project were invaluable and are very much appreciated. I would also like to acknowledge Steve Birkinshaw for his help with SHETRAN, including teaching me how to use the model, troubleshooting my problems and modifying the code multiple times. I am additionally grateful to Greg O'Donnell for sharing his great technical knowledge and bits of code, while sometimes trying to send me off on a tangent. Thanks are also due to Geoff Parkin for his enthusiasm and support for my research, especially with respect to its hydrogeological aspects.

I would also like to thank NERC for funding this research, as well as those individuals and organisations who provided data specifically for this project. Stephen Blenkinsop kindly shared and quality-controlled the hourly rain gauge data used in Chapter 4, while Holger Kessler from the British Geological Survey supplied the geological data used in Chapter 7. Mark Whiteman at the Environment Agency also very helpfully shared the transmissivity data used in Chapter 7.

A huge thank you must go to my awesome brother, Frazer, for teaching me how to code and without whom I would have achieved nowhere near as much.

And finally thank you to David, for everything else.

# Contents

# List of Figures

# List of Tables

# Abbreviations

**ANN** Artificial Neural Networks
**AP** Affinity Propagation
**APM** Aquifer Properties Manual
**ARMA** Autoregressive Moving Average
**ASCII** American Standard Code for Information Interchange
**BGS** British Geological Survey
**BNG** British National Grid
**CEH** Centre for Ecology & Hydrology
**DDM** Data-Driven Modelling
**DEM** Digital Elevation Model
**DEFRA** Department for Environment, Food and Rural Affairs
**EA** Environment Agency
**ESDB** The European Soil Database
**EU** European Union
**G2G** Grid-to-Grid model
**GCM** Global Climate Model
**GPU** Graphical Processing Unit
**HadRM3** Regional Climate Model developed by the Met. Office
**HRU** Hydrological Response Unit
$\boldsymbol{K_s}$ Saturated Water Conductivity
**LCM** Land Cover Map
**LiDAR** Light Detection and Ranging data
**MORECS** Met Office Rainfall and Evaporation Calculation System
**NRFA** National River Flow Archive
**OS** Ordnance Survey
**PBSD** Physically-Based Spatially-Distributed
**PDM** Probability Distributed Model
**PE** Potential Evaporation

**PET** Potential Evapotranspiration

**PUB** Predictions in Ungauged Basins

**RCM** Regional Climate Model

**RMSE** Root Mean Square Error

**SAC-SMA** Sacremento Soil Moisture Accounting Model

**SHE** Systeme Hydrologique European model

**SWM** Stanford Watershed Model

$\boldsymbol{\theta_r}$ Residual Moisture Content

$\boldsymbol{\theta_s}$ Saturated Moisture Content

# Chapter 1

# Introduction

## 1.1 Water related challenges facing Great Britain

The scale of challenges in flood and water resource management facing the UK became abundantly clear during the winter of 2013/2014, when extensive and damaging flooding occurred in numerous parts of the country. A tidal surge on 5th December 2013 along the east coast of England was followed closely by a number of severe weather and fluvial and groundwater flooding events in subsequent months, which formed one of the wettest winters on record (Department for Communities and Local Government, 2014). The professional services firm PwC estimated that the insurance industry would face costs of up to £500M from the December 2013 and January 2014 weather events, with damage to the economy priced at £630M (PricewaterhouseCoopers, 2014). The UK government additionally pledged £560M to help the recovery of communities and businesses, while also repairing damaged infrastructure and flood defences (Department for Communities and Local Government, 2014). This example begins to highlight the magnitude of financial costs, risks to welfare and the extent of disruption associated with large flooding events, which may indeed become more frequent in the coming years as a result of climate change and continuing pressures from development on floodplains (Pachauri et al., 2014; Department for Communities and Local Government, 2009). The events of winter 2013/2014 also suggest the existence of vulnerabilities in current flood risk management practices, as well as the need for tools allowing assessment of broad-scale flood risk and options appraisal.

At the other extreme, droughts represent an additional challenge for water management in the UK, as well as large parts of the world. Rahiz and New (2013a) project large and widespread increases in drought characteristics for the 2050s and 2080s across the UK. Planning to cope with drought risks remains challenging, with unobserved extreme droughts difficult to account for using existing tools and limited national-scale integration in drought

plans. Water resources management strategies in the UK under more typical climatic conditions is also demanding, not least because of the requirement to balance water use with the needs of the environment under the European Union (EU) Water Framework Directive (Water Framework Directive, 2000). This legislation sets out stringent objectives for the ecological status of water bodies across the EU. Responsibility for implementation falls to the Environment Agency (EA) in the UK, who must seek to balance the quantitative and qualitative pressures on water bodies stemming from the requirements of water users with ecological and chemical environmental objectives. This is not an insignificant challenge, given existing pressures on resources in some regions of the UK, legacies from historically unsustainable practices and the confounding influence of climate change. Robust tools are needed to appropriately quantify these pressures and explore the implications of alternative management strategies.

In the context of these issues, the focus of this research is on the development and evaluation of a robust and practically useful national hydrological modelling system for Great Britain using the physically-based model SHETRAN (Ewen et al., 2000), which can ultimately be used to assist in flood and water management. Such a modelling system opens up new possibilities for quantifying multi-source flood risk, examining the potential effects of climate change on flood and drought frequency, and assessing water resource management options amongst other applications. In addition to practical significance for assisting in decision-making, constructing a national physically-based modelling system also provides the opportunity to explore prominent issues regarding the strengths and weaknesses of hydrological models on the basis of a large sample of catchments (e.g. Beven, 2001, 2012; Paniconi and Putti, 2015). This analysis can then feed back into evaluation of whether a national model based on SHETRAN is likely to give robust simulations of runoff in ungauged catchments and under changing conditions such as climate and land use change, which are questions of tremendous scientific and practical interest.

A brief discussion of drivers for hydrological modelling and an introduction to different modelling approaches are given below, in order to contextualise the aims and objectives of the study that are outlined subsequently. The chapter then concludes with an overview of the structure of the remainder of the thesis.

## 1.2 Why model?

Models are simplified descriptions or representations of systems or processes that may form a basis for evaluating understanding, testing hypotheses and making calculations or predictions (Refsgaard, 1996; Beven, 2012). These models may be perceptual (also referred to as

conceptual), reflecting qualitative understanding of how phenomena are perceived to work by scientists or engineers for example (Beven, 2012). There are also mathematical and numerical models, which are formed by translating perceptual understanding of system operation into equations intended to emulate observed empirical behaviour. Implicit or explicit perceptual models are of course integral to all areas of hydrology, but mathematical and numerical models have also found many applications. For example, these models are regularly used in flood management for purposes such as flood frequency estimation, prediction of hydrograph shape and peak flow rates for hydraulic structure design, and flood forecasting (e.g. Cameron et al., 1999; Boughton and Droop, 2003; Kjeldsen, 2007; Bell and Moore, 1998). In addition, models are central to contemporary water resources planning and management, in which water demand and supply considerations need to be carefully balanced against the requirements of the environment (e.g. Arnold et al., 1998; Loucks et al., 2005; Shepley et al., 2012). The potential impacts of climate change on important features of the hydrological cycle are also commonly assessed using numerical modelling, which forms a useful tool to help support mitigation and adaptation planning (e.g. Arnell, 1999; Vörösmarty et al., 2000; Christensen et al., 2004; Prudhomme et al., 2013).

The applications listed above are but a few examples of how hydrological models are used in research and practice. The extensive use of these models in various contexts stems from the simple fact that it is not possible to measure all quantities of interest for supporting important decision-making processes or indeed advancing scientific understanding in hydrology (Beven, 2012). As Beven notes, measurement limitations mean that key decisions in fields such as flood protection and water resources planning require means of extrapolating in space and time from available observations. This is particularly critical where observations of hydrological systems are very limited or absent, such as in ungauged catchments where historical data for characterising flow regimes do not exist. The absence of such data therefore presents a particularly significant challenge to effective management of water resources and hazards in the large areas of the world that are poorly instrumented with respect to key hydrological state variables and fluxes (Sivapalan, 2003). The importance of this challenge is compounded by the numerous pressures and factors affecting water resources and hazards in many regions (e.g. Vörösmarty et al., 2000). In many catchments, both gauged and ungauged, it is of great practical significance to assess possible future hydrological impacts arising from changes in land use or climate, for which no data are available but consequences for societies and environments around the world may be profound (e.g. Bates et al., 2008; Jimenez Cisneros et al., 2014).

In addition to using modelling as a tool for scientific and practical applications in individual catchments, there are also a number of reasons for developing hydrological models

that can be applied at regional or national scales. In terms of advancing science, an important driver for modelling large areas containing multiple catchments is the usefulness of comparative approaches in hydrology, which has been recognised during the recent decade of Prediction in Ungauged Basins (Hrachowitz et al., 2013). Simultaneously investigating multiple catchments forms a means of learning about hydrological processes and the reasons for variation in catchment behaviour. Modelling samples of catchments forms a useful part of this, as well as highlighting possible deficiencies in data, model structure, parameters or indeed process understanding (Henriksen et al., 2003). In addition, national-scale modelling can aid practical flood and water management in various ways. Near-term flood forecasting is one example of this, but a robust model can also assist in planning and decision-making over the longer-term. This can be through the ability to assess various issues, such as the potential effects of scenarios of climatic and land use change on water resources and flood risk. A national-scale model also represents a consistent way to quantify water resource availability, as well as test and evaluate the implications of different water resource management strategies with respect to both demand/supply management and environmental sustainability. Large-scale flood risk assessments are also possible with national-scale models (Hall et al., 2003), and flooding from various sources (such as fluvial and groundwater sources) can be assessed at the same time if an appropriate model structure is applied (see discussion below).

## 1.3 Which model?

The broad spectrum of hydrological models that can be used in various applications can generally be categorised into three overarching classes on the basis of model structure and process representation. The first class of models, empirical models, are based on mathematical relationships between hydrological system inputs and outputs derived solely on the basis of measurements, rather than physical catchment processes (Zhang, 2007; Pechlivanidis et al., 2011). Examples of this class include ARMA (autoregressive moving average) models (Box and Jenkins, 1976), as well as data-driven modelling (DDM) approaches such as artificial neural networks (ANN) (see Dawson and Wilby, 2001). In contrast, the second class of hydrological models consider more explicitly important catchment processes, albeit in simplified ways. Models of this nature, often known as conceptual models, are both widely used and varied, including codes such as the Stanford Watershed Model (Crawford and Linsley, 1966), HBV (Bergström et al., 1995) and TOPMODEL (Beven et al., 1995). Conceptual models typically attempt to represent dominant catchment processes through use of connected reservoirs, utilising mathematical functions to relate storage and flux terms (Beven, 2012). Yet the final class of models, physically-based models, attempt to adopt a higher level of realism

through their basis in equations that are considered to best describe the operation of physical catchment processes. These models typically attempt to solve differential equations based on conservation of mass, energy and momentum governing processes such as channel, overland and subsurface flow, often following from the blueprint put forward by Freeze and Harlan (1969).

This categorisation of model types of course simplifies the complexity and variation found across the spectrum of hydrological models, as discussed further in Chapter 2. The important point to note here is rather that different types of models are likely to have contrasting domains of applicability, which is highly relevant to selection of a model for investigating issues pertaining to ungauged catchments and non-stationary conditions, as explored in this research. For example, conceptual models are often sufficiently flexible that good model performance in relation to observed data can be obtained through calibration of parameters (Beven, 2012). This potential for accurate simulation of runoff has led to their application in many practical contexts, as exemplified by the conceptual Grid-To-Grid (G2G) model developed in the UK (Bell et al., 2007a). This model is used for operational flood forecasting in the UK (Price et al., 2012) and has also been applied in climate change impact studies using Regional Climate Model (RCM) projections (Bell et al., 2007a,b). However, the application of stores and functions parameterised to replicate observed behaviour in conceptual models may not represent physical catchment processes, potentially limiting the reliability of their predictions outside of the limits used in calibration, given the high dependency on data to determine model parameters (Beven, 2012). Physically-based models should theoretically be more reliable in applications such as predicting runoff in ungauged basins or under climate change, although potential limitations to our ability to prescribe equations for physical processes at the model element scale and estimating suitable parameters exist (Beven, 2001). There are ongoing debates surrounding issues such as these (e.g. Refsgaard et al., 2010; Ewen et al., 2012; Refsgaard et al., 2012), while high input data requirements, computational expense and often lower absolute accuracy make physically-based models more challenging to apply in some practical respects.

As the ability to predict flows in ungauged basins and under climatic and land use change are important aims for this research, the physically-based hydrological model SHETRAN is used as the basis for the new national modelling system for Great Britain reported here. The issues surrounding this type of model compared with other options are explored more fully in the literature review (Chapter 2), but in summary it is argued that using SHETRAN fits well with the aims of the research outlined in more detail below, while additionally complementing previous work. Conceptual models for the UK have already been developed (such as G2G mentioned above (Bell et al., 2007a)), but application of fully integrated physically-based

models to large numbers of catchments remains a relatively unexplored area, particularly for the UK. This approach therefore opens up possibilities to investigate a number of interesting questions using a reasonable sample size and comparative approach, for example regarding the degree of calibration required for physically-based models and the transferability of parameters in space and time. These issues impinge on both theoretical debates regarding the physicality of physically-based models, as well as more practical considerations such as the reliability of this category of models for making useful predictions for water management. Indeed, utilising an integrated surface-subsurface model such as SHETRAN paves the way for applications that are unachievable using existing national models for the UK, such as national-scale multi-source flood risk assessment.

## 1.4 Aims and objectives

The aim of this research therefore is to set up a national physically based hydrological model using SHETRAN (SHETRAN for GB) and to test its capacity for providing robust results, that can be used to model flood and drought frequencies, in gauged and ungauged catchments, both now and under future climate change.

The main objectives are to:

- Collate and process national, freely available datasets required for use in SHETRAN
- Develop software to automate the set up of robustly parameterised catchment models
- Develop a gridded hourly rainfall data product for use with the national modelling system to provide higher resolution information
- Assess the uncalibrated performance of the national modelling system
- Assess the robustness and sensitivity of the system to meteorological inputs
- Relate model performance to catchment characteristics
- Make structural improvements to both SHETRAN and the national datasets where required
- Couple the national SHETRAN system to the UKCP09 spatial weather generator
- Develop a methodology for a national climate change impact assessment using SHETRAN and the weather generator.

## 1.5 Thesis outline

Chapter 2 reviews literature relating to types of hydrological models, their uses and the argument for physically based modelling. This leads to an outline of the SHETRAN hydrological model and its applications. A discussion of other national modelling systems follows and then this work is put into the context of predictions in ungauged basins.

Chapter 3 describes the data collated and processing methods for the SHETRAN for GB system. The software that has been developed as part of this work is presented and the underlying algorithms are explained. The general climatology and physiography of the country are also briefly described.

Chapter 4 outlines the creation of a gridded hourly rainfall product. A description of the sub-daily rainfall data and quality control procedures are presented, including a comparison of the hourly records to daily rainfall data. The disaggregation method is described and limitations are discussed. A brief analysis of the quality of the resulting dataset is presented.

Chapter 5 presents the results of the initial simulations from the modelling system. A number of structural changes are made based on these results and a standard version of the system is reviewed. A series of sensitivity tests to universal model parameters such as the roughness coefficient, rainfall and evaporation rates, temporal rainfall distributions and the inclusion of the hourly rainfall dataset described in Chapter 3. The model performance is then related to catchment characteristics.

Chapter 6 details a methodology for linking SHETRAN for GB to the UKCP09 spatial weather generator and conducting a climate change impact assessment. Metrics for analysis and some initial results are presented.

Chapter 7 addresses the representation of geology in the national modelling system. The BGS national 3D geological model and other datasets are incorporated into SHETRAN and the importance of accurate geological information is discussed.

Finally, the main conclusions are presented, limitations of the study are discussed and possible future work is suggested.

# Chapter 2

# Literature Review

## 2.1 Structure of literature review

From the discussion above it begins to become apparent that making reliable, robust and realistic hydrological predictions in ungauged catchments and/or under non-stationary conditions is a central focus and challenge in hydrology (e.g. Klemes, 1986; Sivapalan, 2003; Gupta et al., 2014). This review aims to explore this challenge through considering a number of its dimensions. Firstly, a brief overview of different types of hydrological models and their relative advantages and disadvantages is given. The purpose of this is to help identify the properties of hydrological models that may be required to make predictions in the cases of interest. On the basis of these considerations, the hydrological model selected for this study, SHETRAN, is then discussed. Other national modelling frameworks are described and the broader topic of prediction in ungauged basins is discussed. It should be noted that literature related to climate change impact assessment is presented in Chapter 6, as this in an exemplar application of the modelling system rather than a key focus of this study.

## 2.2 Hydrological model classification

Several classifications have been proposed for the substantial and varied range of hydrological models in existence (e.g. O'Connell, 1991; Wheater et al., 1993; Singh, 1995; Refsgaard, 1996). These classifications typically differentiate between models on the basis of at least three key attributes: whether or not the model is deterministic or stochastic; how space is discretised; and how hydrological processes are represented or

described (Refsgaard, 1996; Beven, 2012). These attributes and common relationships between them form the basis of discussion in this section. However, it should be noted that it is also possible to classify hydrological models based on other features, such as their treatment of time (e.g. event-based or continuous simulation).

### 2.2.1 Deterministic and stochastic models

As noted above, hydrological models can be differentiated based on whether they are deterministic or stochastic. A model may be considered deterministic if it consistently produces a single, identical set of output variables for a given set of inputs, parameter values and boundary conditions (Beven, 2012). This means that a given input will always result in the same output if other aspects of the model are held constant; model outputs are therefore entirely determined by specified and constant relationships between inputs and states. In contrast, stochastic models may produce multiple outputs for a given set of inputs by incorporating a random element into some aspect of the model or its inputs. This approach is often used in recognition of uncertainties in inputs, parameter values or boundary conditions, with the result that a distribution of outputs may be obtained for a single set of inputs (Beven, 2012). However, the distinction between deterministic models and stochastic models is not always simple. For example, Beven (2012) notes that there are examples where stochastic error models are coupled with outputs from deterministic hydrological models. This point is illustrated by the method described in Montanari and Koutsoyiannis (2012), in which stochastic perturbation of the inputs, parameters and outputs from a deterministic model is conducted to obtain a probability distribution and so assess uncertainty in model results. It may also be noted that there are models performing deterministic predictions but using probability density functions of state variables (e.g. PDM Moore, 2007). This leads Beven (2012) to suggest that if model output variables have some variance (at a given time step in the case of continuous simulation) then they could be classified as stochastic, whereas if outputs adopt a single value then the model may be considered deterministic.

### 2.2.2 Spatial discretisation

In addition to the distinction between deterministic and stochastic approaches, hydrological models may be further classified on the basis of the mode of spatial discretisation utilised. Two key approaches to the treatment of space are often identified. Lumped

models treat the entire catchment as a homogeneous entity, with no explicit representation of spatial variation in catchment structure, properties, processes, boundary conditions or inputs (Singh, 1995). State variables such as subsurface storage thus effectively represent averages over the catchment area (Beven, 2012). Conversely, distributed models divide the whole catchment into a finite number of elements, which may be assigned different properties and so exhibit contrasting hydrological responses to forcing inputs. In this case equations for state variables are solved for each element, with the resulting states representing local averages. The implication of this is that distributed models are able to account more explicitly for spatial variation in catchment hydrology and its implications for key processes. This feature is conceptually attractive; it seems intuitively more realistic than representing visibly heterogeneous catchments as spatially homogeneous.

However, there are several complications here, which are related to both theoretical and practical issues. For example, a number of questions have been posed concerning scale issues in distributed models, such as how to identify appropriate parameter values at the element scale or reconcile differences between the scale of variation of processes and the element scale required for computational feasibility (Beven, 2012). Furthermore, using common model performance metrics for evaluating the degree of fit between observed and modelled variables for a historical simulation, it is not necessarily the case that distributed models produce more accurate results than lumped models. Issues such as these are considered in more detail shortly.

It should also be noted that another form of spatial discretisation may be identified in between the two extremes of lumped and distributed models. Semi-distributed models are based on modelling multiple discrete sub-catchments, sometimes referred to as hydrological response units (HRUs) (Beven, 2012). These sub-catchments are typically represented as homogenous units based on some metric of hydrological similarity, though clearly they will differ from each other and so provide some level of spatial heterogeneity compared with lumped models. This approach has been suggested to represent a compromise between the computational efficiency of lumped models and the explicit representation of important spatial variability of models employing a fully distributed spatial discretisation (Orellana et al., 2008).

### 2.2.3 Model structure and process representation

#### Empirical models

One group of hydrological models that may be identified on the basis of model structure and process representation are empirical models. These models are essentially based on mathematical relationships between hydrological system inputs and outputs, which are derived solely on the basis of measurements and typically without detailed consideration of physical catchment processes (Zhang, 2007; Pechlivanidis et al., 2011). This means that empirical models generally rely on the availability of observation data to characterise the behaviour of a hydrological system (Wheater et al., 1993). This class of models could be considered to include a broad range of methods, such as transfer functions, which form the basis of the unit hydrograph and its variants (Sherman, 1932; Beven, 2012) or the data-based mechanistic method applied by e.g. Young and Beven (1994). This latter approach attempts to limit prior assumptions about the model structure, which is then identified from the available system input/output data and considered valid only if it is mechanistically plausible. Time series models such as autoregressive moving average (ARMA) methods (Box and Jenkins, 1976) could also be considered as part of the class of empirical models, as well as other data-driven modelling (DDM) approaches such as artificial neural networks (ANN) (see Dawson and Wilby, 2001). In ANN, a learning set of data is used to train weighting functions in a network of connected nodes relating system inputs and outputs.

Empirical models are therefore essentially based on induction from available data and may provide useful tools for the hydrologist in various applications. Yet in the cases of particular interest in this project (predictions in ungauged catchments and under changing conditions), empirical models may not be a reliable tool. Firstly, if observation data are not available then estimating the parameters and/or form of empirical models may be difficult. Secondly, as empirical models depend on historical data, future alterations to hydrological processes arising from changes in climate or land use may result in deviations from historical behaviour that are not easily accounted for in models conditioned so heavily on past events. However, in some cases it has been possible to use this class of models in certain applications despite the absence of site-specific data. For example, in some regions, statistical relationships have been derived between catchment descriptors and unit hydrograph parameters. These relationships allow for estimation of flood hydrographs in ungauged catchments through regionalisation, such as in the rainfall-runoff models in the UK's Flood Studies Report (FSR; NERC, 1975) and Flood Estimation Handbook (FEH; Institute of Hydrology, 1999). The potential

for applying ANN to predict flood statistics or index floods in ungauged catchments with comparable performance to FEH methods has also been demonstrated by Dawson et al. (2006) for example, although limitations of the large dependency on data availability for training have been recognised. The use of regionalisation in methods such as these will be discussed further shortly, but it should be noted at this stage that in general empirical methods are intrinsically limited by their limited physical underpinning, which is problematic when attempting to account for changing hydrological processes and estimate the range of quantities of interest desired for practical water management under current and future conditions.

### Conceptual models

In contrast to purely data-driven empirical modelling, conceptual models typically attempt to represent dominant hydrological processes understood to be significant in modulating system input-output relationships at the catchment-scale (Wheater, 2002). Conceptual models are often considered to have the feature that model structure is specified a priori (Wheater et al., 1993), in contrast to some of the data-based mechanistic modelling described above for example. The specification of model structure is subjective, drawing on the perceptual model of a catchment(s) and so the experience of the hydrologist (Beven, 2012). Although there is a large range of structures used in different conceptual models, these are usually formed from a series of connected reservoirs or stores (Pechlivanidis et al., 2011; Beven, 2012). Mathematical functions are defined to relate the storage and flux terms associated with these reservoirs. Some models may use a small number of reservoirs or stores, whereas others may implement much more complex structures. However, the schematic representation of hydrological processes in this way means that the parameters of conceptual models do not all have physical meaning (Wheater et al., 1993). This means that at least some of the parameters are not directly measurable, such that they have to be estimated through the process of calibration against observed data (Pechlivanidis et al., 2011; Beven, 2012).

It is often the case that conceptual models adopt a lumped spatial discretisation. As discussed above, this means that spatial variation and heterogeneity is not explicitly accounted for. Rather, computed state variables are in effect catchment averages. Perhaps the earliest example of this type of lumped conceptual model is the Stanford Watershed Model (SWM, Crawford and Linsley, 1966). This model simulates a number of commonly important hydrological processes in a simplified manner. For example, effective rainfall is apportioned to several subsurface stores, with linear and nonlinear reservoirs

used to simulate interflow and active groundwater respectively. Since the development of SWM, many more lumped-parameter conceptual models have been developed, including the Sacremento Soil Moisture Accounting Model (SAC-SMA, Burnash et al., 1973), the TANK model (Sugawara et al., 1983), the APIC model (Sittner et al., 1969), the SSARR model (Rockwood et al., 1972) and the Probability Distributed Model (PDM, Moore, 2007) amongst others. The latter model, PDM, exemplifies a set of conceptual models that utilise a probability distribution function of absorption capacity (or soil moisture) to attempt to implicitly account for spatial variation in catchment soil moisture levels and its controlling effect on saturation excess runoff generation. This model has been used in a number of hydrological applications, such as for deriving flood frequency curves through conducting long model runs (e.g. Lamb, 1999).

However, there are also examples of conceptual models adopting semi-distributed or fully distributed modes of spatial discretisation. One example of this is the ARNO model developed by Todini (1996), which also uses a probability distribution function of soil moisture but additionally allows for division of a catchment into sub-catchments that contribute to the overall rainfall-runoff response. HBV is another example of a semi-distributed conceptual model that has been applied in a wide range of contexts (Bergström et al., 1995; Götzinger and Bárdossy, 2007; Seibert, 2003; Seibert and McDonnell, 2010). In addition, the PDM has also been applied in more distributed forms. As Beven (2012) notes, concepts underpinning PDM have been applied in the Grid Model described by Bell and Moore (1998) and more recently implemented in the distributed Grid-To-Grid (G2G) model (Bell et al., 2007a). The G2G model is used for operational flood forecasting in the UK using inputs from the Met Office Global and Regional Ensemble Prediction System Forecasts (MOGREPS, Beven, 2012). G2G has also been used in conjunction with radar rainfall for the purpose of real-time forecasting (Cole and Moore, 2008). Future climate projections for the UK from a 25 km Regional Climate Model (RCM) have also been assessed using the G2G model (Bell et al., 2007a,b). TOPMODEL (Beven and Kirkby, 1979) is another popular conceptual model that makes use of topographical data and a minimal parameter set, which can be used to evaluate spatial patterns of results, without being a fully distributed model.

One of the general advantages of conceptual models is that there is often sufficient flexibility arising from their mathematical structure and parameters to obtain good model performance in relation to observed data (Beven, 2012). Additionally, conceptual models may also be considered to have relatively modest data requirements, as well as the option to adapt the complexity of a conceptual model according to the level of data available (Wheater, 2002). This class of hydrological models are also often fairly simple

and rapid to set up and apply. This means that they currently have many applications, providing an important tool for decision-making in areas such as flood forecasting and management as mentioned above.

However, although conceptual models may reproduce historical catchment behaviour with some reference to key processes perceived to be important, this class of models may not exhibit a large degree of fidelity with respect to physical catchment processes. The application of stores and functions parameterised to replicate observed behaviour, as opposed to physical relations governing hydrological processes, means that conceptual models are subject to some of the same limitations as empirical models. For example, predictions outside of the limits used in calibration may be questionable, given the high dependency on data to determine model parameters (Beven, 2012). This is potentially an issue for investigations of possibly unobserved hydrological extremes, which may need to be taken into account in water resources and flood management. In addition, omission of physical relationships and incorporation of parameters with limited physical meaning result in difficulties applying conceptual models under conditions of changing climate or land use. This is because parameters calibrated based on historical observations may not be robust under future conditions, particularly when those parameters are not easily related to physical catchment properties. Indeed, identifying the most representative parameters is not a simple task even for historical periods in many cases, given the limited information content of available data, which may be a particular problem for conceptual models with high-dimensional parameter spaces (Beven, 2012). Finally, reliance on observed data again raises issues of how conceptual models could be used to predict runoff in ungauged catchments. Some possible solutions for this last problem have been attempted, which will be discussed shortly.

### Physically based models

Physically based hydrological models differ from empirical or conceptual approaches in that they adopt an ostensibly higher level of physical realism by explicitly solving equations describing the operation of catchment processes. This class of models typically follows from Freeze and Harlan (1969)'s seminal blueprint paper, in which equations characterising surface and subsurface processes are described and linked to form an integrated model of the hydrological cycle. These nonlinear partial differential equations describe processes of infiltration, overland flow, unsaturated and saturated subsurface flow on the basis of continuity of mass and momentum. The equations are typically solved using numerical methods such as finite difference or finite element ap-

proximations (Beven, 2012). These methods are often applied in conjunction with a fully distributed spatial discretisation in two or three dimensions, with the catchment discretised into a grid of elements for which hydrological processes are resolved. The physical nature of the equations employed is often suggested to mean that in theory their parameters are measurable, thereby allowing incorporation of realistic variability in catchment properties into model simulations in principle (Beven and O'Connell, 1982).

A number of different modelling programs have been developed since the elucidation of this blueprint for physically based spatially distributed (PBSD) hydrological models. One such early realisation of this approach is in the Systeme Hydrologique Europeen (SHE) modelling system developed by the British Institute of Hydrology, the Danish Hydraulic Institute and the French company SOGREAH (Abbott et al., 1986). The modules comprising SHE reflect the guiding principles outlined by Freeze and Harlan (1969) of formulating and solving equations based on conservation of mass, momentum and energy. For example, diffusion wave approximations to the St Venant equations are used to simulate overland and channel flow, with unsaturated and saturated zone flows modelled using the one-dimensional Richards equation and two-dimensional Boussinesq equation respectively (Abbott et al., 1986). The significant input data and parameter specification requirements implied by this formulation were recognised during the design of SHE, such that flexibility through a modular structure was incorporated so that simpler calculation methods could be employed in the absence of sufficient data to justify more sophisticated approaches. This is reflected in the options for calculating evapotranspiration, where three calculation modes are available in view of practical difficulties in constraining all terms in the Penman-Monteith equation. Similarly, the option to use a degree day approach to snowmelt modelling was included, given that sufficient data are not consistently available for use of the full energy balance approach.

The SHE project has branched into two main strands, namely MIKE SHE (Refsgaard and Storm, 1995) and SHETRAN (Ewen et al., 2012). These models represent advancements over the SHE formulation through the inclusion of fully three-dimensional subsurface components, as well as addition of sediment transport and water quality modules. A description and review of SHETRAN the model used in this research is given below but it is worth noting that several other physically based modelling codes have been developed in addition to SHE and its derivatives. These models include the Institute of Hydrology Distributed Model (IHDM) (Beven et al., 1987), in which discretised hillslope planes provide a flexible representation of subsurface flow. The rationale for this approach is that, when compared with using square grid elements of

the size typically employed in SHE, it may more accurately represent processes related to topographic convergence and divergence, such as expansion and contraction of areas of saturation and surface flow close to channels (Beven et al., 1987).

Other examples of physically based models include THALES (Grayson et al., 1992), GSSHA (Downer and Ogden, 2004), WASIM-ETH (Schulla and Jasper, 2015) and HydroGeoSphere (Therrien et al., 2010; Brunner and Simmons, 2012). The latter model is one of the most recently developed physically based hydrological models. It uses a finite element approach to simultaneously solve the Richards equation describing three-dimensional variably saturated subsurface flow and a diffusion wave approximation to the St Venant equation for two-dimensional surface flows. This approach could be considered as one of the most fully integrated ways for physically based catchment simulation, but notably it is still fundamentally rooted in the seminal blueprint described earlier (Freeze and Harlan, 1969; Brunner and Simmons, 2012). This is the case for the majority of physically based models, which tend to differ more in their approaches to spatial discretisation and methods of solution rather than their underlying principles (Beven, 2012).

Physically based spatially distributed models have been applied in a range of studies of hydrological processes, climate change and land use change, as well as in some practical water management applications. For example, Goderniaux et al. (2009) utilised HydroGeoSphere for fully integrated modelling of surface and subsurface processes in the Geer basin, a catchment underlain by a Chalk aquifer in eastern Belgium. Satisfactory simulation of both stream flows and groundwater heads was obtained, with the authors finding that the integrated modelling approach and spatially explicit representation of evapotranspiration contribute to sound physical realism in the model. The calibrated model was then used for assessment of potential climate change impacts using downscaled and bias-corrected regional climate model (RCM) outputs from a medium-high emissions scenario. These projections indicated potential reductions in groundwater heads of up to 8m and decreases in stream flow of between 9 and 33% in this catchment by 2080. In terms of applications in water management, Refsgaard et al. (2010) note how MIKE SHE has been utilised in studies of impacts of different operational strategies for the Gabcikovo hydropower scheme on the Danube, including interactions between stream flows, sediment processes, aquifer chemistry and contamination. Refsgaard et al. (2010) also describe how MIKE SHE has been utilised to support management of flood risk, water resources and environmental sustainability in the South Florida Water Management District. Other examples of physically based models being applied in national-scale modelling are discussed below.

The combination of a basis in physical process descriptions and an ability to account for catchment heterogeneity therefore makes physically based models conceptually attractive for numerous applications, such as modelling ungauged catchments or basins under non-stationary climatic or land use changes (Beven and O'Connell, 1982). However, a number of critiques of physically based spatially distributed models have emerged over time (most recently Paniconi and Putti (2015)), particularly following from Beven (1989). These critiques are neatly characterised by Beven (1999a), who identifies problems of nonlinearity, scale, equifinality, uniqueness and uncertainty in physically based distributed hydrological modelling. The first issue, nonlinearity, is related to the fact that the dynamics of nonlinear systems do not average in a simple manner and extremes exert a large influence on overall system response. Beven connects this issue with the problem of scale through the argument that certain equations typically used in physically based models particularly the nonlinear Richards equation are not necessarily applicable at feasible grid resolutions in spatially distributed catchment modelling. In conjunction with local heterogeneity meaning it is difficult to estimate parameters of the equations at the scale of a model element, the nonlinearity of the Richards equation means that applying it for a model element would not be consistent with integrating the Richards equation applied more locally. Beven additionally notes that sub-grid heterogeneities also arise from complicated factors such as preferential flow pathways, which are not the focus of the Richards equation and particularly difficult to parameterise at the model element scale.

What Beven terms problems of uniqueness of place and equifinality are connected to the issue of differentiating between multiple model structures and parameter sets that provide similar levels of performance in simulating catchment behaviour. Given the complexities of real catchments and the limitations of current theories for quantitatively describing all of their nuances, Beven argues that identifying a single optimal model structure for most catchments is very difficult. This is closely related to the fact that insufficient data are typically available to identify a unique optimal parameter set for any particular model, an observation stemming from earlier work by Duan et al. (1992) and Beven (1993) on simulating catchment discharge using multiple parameter sets. As Beven (2001) notes, this issue is particularly relevant for distributed models, which can in theory have very high numbers of parameters that can vary between model grid cells, and it is often the interaction between parameters that determines the performance of a simulation. Both of the problems of specifying optimal model structures and parameter set are further complicated by limitations in the availability of sufficient input and validation data, the errors in which are often difficult to fully quantify. All

of these issues in combination lead Beven (2001) to suggest that uncertainty analysis is an integral part of hydrological modelling, in order to display more faithfully the limitations of model results.

The arguments put forward by Beven highlight some of the potential issues with physically based spatially distributed models, as well as key problems and uncertainties in hydrological modelling more generally. It is of course important to note that, while issues of nonlinearity and scale may indeed be associated with models that have their origins in Freeze and Harlan (1969)'s blueprint, other categories of hydrological models are rooted in mathematics that also only approximates catchment functioning, albeit with typically less of a basis in understanding of physical processes. It is also certainly the case that unique model structures and parameter sets cannot necessarily be identified for conceptual or empirical hydrological models, such that these issues are not only relevant for physically based distributed models. In addition, there are many examples of successful use of physically based models in simulating catchment behaviour and investigating different practical problems, some of which have been detailed above and others pertaining to national modelling systems are given below. This does not diminish some of the issues raised by Beven of course, but it does show that physically based models in their current form are able to add value in water management problems, although their limitations need to be clearly acknowledged. For investigating particular problems, such as nonstationary climatic or land use conditions, physically based models are therefore considered to still provide the best option, as they are more likely to provide physically plausible behaviour under different forcings compared with conceptual or empirical models.

## 2.3 SHETRAN

The physically-based spatially-distributed hydrological modelling system employed in this research is SHETRAN, which has its origins in the Systeme Hydrologique Europeen (SHE) model developed by the British Institute of Hydrology, the Danish Hydraulic Institute and the French company SOGREAH (Abbott et al., 1986; Ewen et al., 2012). As discussed earlier, the foundations of SHE were strongly influenced by Freeze and Harlan (1969)'s blueprint, with these principles taken further and additional processes incorporated during the development of SHETRAN, along with some divergences in approach. This development process was partly funded by United Kingdom Nirex Limited for the practical purpose of assessing the safety of a proposed deep underground repos-

itory for radioactive wastes (Ewen et al., 2012). For this assessment a model capable of analysing transport of radionuclides in surface and shallow subsurface hydrological systems was required. This led to the current form of SHETRAN, in which finite difference approximations are used to solve equations describing fully three-dimensional coupled surface/subsurface water flow and transport of sediments and reactive solutes. Given the context of its development, it is unsurprising then that a significant strength of SHETRAN is its detailed treatment of the subsurface as a variably saturated porous medium and the direct coupling of surface and subsurface flow and transport.

The sediment and solute transport capabilities of SHETRAN are not explored in this project, but of course the modelling system developed here could provide a starting point for rapid setup of catchments for investigations of these processes. Focussing instead on the water flow component, however, SHETRAN includes processes of canopy interception of rainfall, evapotranspiration, infiltration, surface runoff (overland, channelised and overbank), snowpack formation and melting, storage and flow in the variably saturated subsurface, flows between surface and subsurface water bodies and groundwater seepage discharge (Ewen et al., 2012). Human influences on catchments are also accounted for, as SHETRAN can incorporate river and groundwater abstractions, flow augmentation schemes and irrigation. Subsurface flow is based on equations for variably saturated flow, which accounts for variations in hydraulic conductivity and storage properties as functions of saturation. Overland and channelised flow are simulated using a diffusion approximation to the Saint-Venant equations in two and one dimensions respectively. Evapotranspiration may be represented using the Penman-Monteith formulation or as a fraction of externally supplied potential evapotranspiration, while snowmelt can be simulated using either a temperature-index or energy balance approach.

Uses and application of SHETRAN beyond supporting the Nirex work described above are many and diverse. For example, Bovolo and Bathurst (2012) used SHETRAN to determine the magnitude of shallow landslides as a function of rainfall return period, while Parkin et al. (2007) investigated the impact of groundwater abstractions on stream flow and use SHETRAN to generate a suite of hypothetical groundwater abstraction impact realisations. In addition, temperature is demonstrated to be a useful tracer of flow pathways using a calibrated SHETRAN model in Birkinshaw and Webb (2010), whereas Birkinshaw et al. (2011) used SHETRAN to demonstrate the impact of deforestation on peak flows and sediment yield in central-southern Chile. SHETRAN was also found to be suitable for real-time flow forecasting in Mellor et al. (2000). SHETRAN has additionally been used to study nitrate transport (Koo and

O'Connell, 2006), groundwater contaminants (Ewen, 1990) and the impact of climate change on Atlantic salmon (Walsh and Kilsby, 2007) amongst other issues. However, it is of note that many of these studies are generally focussed on modelling individual catchments. Simultaneous simulation of large samples of catchments has not been extensively attempted for SHETRAN, although such investigation provides a very useful opportunity to evaluate model performance over a range of catchment types and explore the potential for application of the model to support broad-scale flood and water resources management.

### 2.3.1 National hydrological modelling

Following the discussion of different types of hydrological modelling above, examples of applications of models at the national-scale are now considered. This provides some context for the work undertaken in this project to configure a physically-based model for robust and relatively rapid simulations of large numbers of catchments across Great Britain.

The first example of national hydrological modelling is taken from Denmark, where the integrated groundwater-surface water model MIKE SHE mentioned earlier has been applied over an area of 43000km$^2$ (Henriksen et al., 2003). This project was driven by the pressures facing groundwater and related effects on stream flows and habitats in Denmark, where around 99% of the country's water supply comes from groundwater. The hydrological context and issues therefore dictated deployment of an integrated surface/subsurface model, for which input data were determined from national databases of geology, soil, topography, river systems, climate and hydrology. The modelling methodology was developed through several iterations over a period of 5 years by a team of government scientists, ultimately culminating in the creation of 11 regional sub-models for Denmark (referred to collectively as the DK model). Catchments within the regional sub-models are not predetermined in this system but are calculated by MIKE SHE, allowing for the possibility that the divide between surface water catchment and groundwater catchments may not coincide. The DK model includes similar processes and components to SHETRAN, although notably the unsaturated zone component in MIKE SHE based on the Richards equation was excluded to reduce run times. Three-dimensional geology including superficial deposits was incorporated, with the model set up at 1km resolution with universally applied parameters throughout based on the literature, pumping tests, fieldwork and previous modelling results.

The problem of over-parameterisation and excessive tuning was kept to a minimum

by using universal parameter values consistently across the model, 10 free parameters remained for calibration. This was conducted principally using manual trial-and-error. Outputs from the model were calibrated and evaluated using data from 4439 boreholes and 28 river gauging stations, with a system of model performance statistics developed to capture different aspects of how well the simulations matched observed data. This provides a useful multi-criteria approach to model evaluation, which has been shown to be of significant value in constraining parameters in other physically-based modelling studies (Anderton et al., 2002). Henriksen et al. (2003) concluded that the DK model produces reliable results, with effective simulation of both groundwater heads and river flows, leading the authors to suggest that the model is appropriate for operational applications and national impact studies. There was some degree of discrepancy in overall water balance for the simulations, however, for which possible explanations include overestimation of precipitation or underestimation of potential evapotranspiration. Discerning between these different explanations remains challenging, as does identifying the most appropriate calibration strategy and determining whether all of the many processes in MIKE SHE are correctly represented. Henriksen et al. (2003) also noted how the challenges in collating and processing large amounts of data are complemented by advances in understanding of the hydrological processes in Denmark.

In comparison with the MIKE SHE model for Denmark, several differences in approach were taken in the construction of the national model for France reported in Habets et al. (2008). The French national model results from coupling a meteorological analysis system (SAFRAN) with a land surface scheme (ISBA) and a groundwater model (MODCOU) that solves the diffusion equation. The representation of hydrological processes is therefore potentially less integrated compared with the MIKE SHE model for Denmark, particularly with respect to surface/subsurface coupling. The French national model is typically run at an 8km grid square resolution but can be refined down to 1km for particular applications. Calibration of the modelling system was not undertaken, with parameters taken from national databases. Evaluation with respect to gauged flows for the four largest basins in France revealed Nash-Sutcliffe Efficiency (NSE) values ranging from 0.68 to 0.88 and water balance biases between -10% to +6%. In addition, a 10-year simulation showed good results for 610 river gauges, with 66% of gauges showing NSE greater than 0.55 and 36% showing NSE greater than 0.65. The model was also compared to snow depth observations and piezometric head observations. The results were generally considered to be encouraging, but it was noted that there is notable variability in performance, with NSE ranging from less than 0 to almost 1 and water balance biases ranging from -60% to +60%. It was also found that dry

years tended to be more poorly simulated than wet years. Influences from hydropower and dams were not incorporated in the model. The modelling system has been used at MeteoFrance operationally since 2003 for real-time monitoring of the water budget estimation of soil moisture for informing flood risk and drought monitoring.

A further example of national hydrological modelling comes from Sweden, where a national model has been constructed based on the water quality and quantity model HYPE (Strömqvist et al., 2012; Lindstrom et al., 2010). HYPE is a dynamic, semi-distributed process-based integrated catchment model, which is built around a modular structure in which hydrological response units are the basic model elements. Unlike national models focusing on quantifying water availability or flood risk, Strömqvist et al. (2012) indicate that the main purpose of this system is to assess water quality, including levels of nitrogen and phosphorus in response to the European Water Framework Directive (Water Framework Directive, 2000). The HYPE model has many hundreds of parameters that could be adjusted during calibration, but when configuring the model, parameters were mainly estimated from previous modelling experience and from the literature apart from 15 parameters for each land use and soil type which are calibrated and a further 10 parameters calibrated for the model globally. These parameters were calibrated manually and in a stepwise manner from the most important parameters to the least sensitive ones. Importantly, the calibration of sub-groups of gauged sub-basins was performed simultaneously. The model was able to simulate large unregulated catchments better than small catchments and unregulated catchments better than regulated catchments.

The current standard national hydrological model for the UK is the 1km national Grid to Grid model (G2G) (Bell et al., 2009). Runoff production in G2G is parameterized on slope and the soil and geology present. It is based on the probability distributed model (PDM) (Moore, 2007) and models surface water flow, soil water volumes and drainage. Model performance was assessed on river flows from 42 gauging stations using a daily time step. The model was found to simulate upland catchments well but struggled to properly simulate flat, groundwater dominated catchments.

Overall the UK national model performs similarly well to the other national models discussed here, however, this model is severely limited in its capabilities. The Danish MIKE SHE model can simulate groundwater levels well and the Swedish HYPE model is able to simulate nutrient transport and water quality. There is therefore clearly a need for a more comprehensive UK national model that is able to be used for several purposes instead of limited to just looking at flood peaks. The G2G model is also

limited by its lack of representation of snow melt and lakes which can have a significant effect on hydrology in the UK, especially in the north of England and Scotland.

### 2.3.2 Predictions in ungauged catchments

Making reliable, robust and realistic hydrological predictions for ungauged catchments and/or non-stationary conditions - such as land use or climate change - is considered to be a central focus and challenge in hydrology (e.g. Klemes, 1986; Sivapalan, 2003; Gupta et al., 2014). This challenge has been the subject of a substantial body of research in recent years (Hrachowitz et al., 2013). As discussed earlier, the majority of the worlds basins are ungauged, with significant regional variation in the degree of monitoring of stream flows. This substantially complicates the application of hydrological models to support management of the array of water-related issues experienced in many catchments, as calibration and evaluation with respect to measured flows is essential in most model development processes, particularly for empirical and conceptual models. Alternative options able to account for the multi-scale spatio-temporal heterogeneity of catchments must therefore be explored if the benefits of models are to be realised in ungauged basins. One approach to this is relating the parameters of models to more easily measurable catchment characteristics, in order to transfer parameters from gauged catchments where calibration is possible to ungauged catchments (Vogel, 2005). Another possibility is to apply physically-based models, if sufficient prior confidence in the likely accuracy of their predictions can be attained.

The International Association of Hydrological Sciences (IAHS) led a Predictions in Ungauged Basins (PUB) initiative over 10 years in order to further this aspect of hydrology (Sivapalan, 2003; Hrachowitz et al., 2013). The aim of PUB was to move away from models that rely on calibration and move towards models that emphasise understanding of catchment processes. One of the questions they investigated was whether point-scale physics really represent large scale patterns and dynamics, i.e. is 'the whole greater than the sum of its parts' in catchment hydrology? If so, is this best accounted for by building models from a bottom-up approach (similar to physically-based models such as SHETRAN) or using a top-down method, in which processes are parameterised directly at larger scales? As part of this, various research groups initiated experiments to help develop process understanding. The conclusion following from this is that hydrological models needed to change in order to more flexibly incorporate competing conceptualisations and permit testing of alternative model formulations. Modelling frameworks such as FUSE (Clark et al., 2008) and SUPERFLEX (Fenicia et al., 2011)

were developed to allow for intercomparison of conceptual model components, recognising that no particular model is necessarily universally preferable for all applications and that model performance depends upon the setting (Rutter et al., 2009).

To use these models in ungauged catchments, parameters must somehow be transferred from gauged catchments. The process for doing this is discussed below.

### Catchment characteristics

Defining catchment characteristics is an essential first step in transferring model parameters by any method from a gauged to ungauged basin. This process is also considered to be a fundamental part of improving catchment science, capitalising on the advantages of comparative approaches in learning about hydrological systems (Hrachowitz et al., 2013; Vogel, 2005). Catchment characteristics used for various applications can be numerous, but generally relate to physical features of a basin, such as area, slope and metrics describing relevant properties of land cover and soils, as well as climate (for example annual average rainfall and temperature) and hydrology (for example baseflow index). The Flood Estimation Handbook (FEH) (Institute of Hydrology, 1999) defines 25 catchment descriptors in its industry-standard methodology for pooling UK catchments on the basis of their properties. These characteristics include catchment mean altitude, longest drainage path, extent of urban and suburban cover, baseflow index, mean annual rainfall and mean flood depth amongst other descriptors. While some commonly used catchment characteristics can be relatively easily and objectively calculated, such as metrics of catchment elevation or slope, definition of other characteristics can be more subjective. To explore this further, Singh et al. (2014) attempt to assess which catchment characteristics control the success of parameter transfer to ungauged catchments in hydrological modelling. Singh et al. carried out their analysis using example catchments from across the whole of the USA, finding that the controls on successful parameter transfer vary significantly with scale, region and objective function. In general, however, it was concluded that the performance of hydrological models was largely controlled by the level of similarity in climate, medium and low flow signatures and elevation between climates.

Following definition of catchment characteristics, one step commonly taken next is to quantify in some way catchment similarity, i.e. the extent to which properties (and therefore hopefully functional behaviour) match between the catchments of interest. This approach was taken by Ali et al. (2012), who conducted a study using 36 catchments in Scotland to examine similarity indices for catchment classification. Unlike

previous studies where parameters have been tied to individual catchment characteristics, the authors used the Affinity Propagation (AP) clustering algorithm (Frey and Dueck, 2007) to divide catchments into groups of similar characteristics, before identifying an exemplar catchment that was most representative of any given group. This methodology has the benefit of allowing the methodology to objectively determine patterns of similarity without significant prior imposition by the researcher, although of course the parameters required as input to the clustering methodology must be specified. However, one of the drawbacks of AP is that it does not allow for fuzzy clustering', which would better account for cases such as catchments situated on the boundary of two groups. In AP, membership of a cluster is categorical, with no possibility of partial degrees of membership, which would perhaps be more intuitive and potentially useful. Notably, Frey and Dueck (2007) also found that physical catchment characteristics do not always correlate well to functional characteristics describing catchment behaviour, such as runoff coefficients. In addition, the choice of similarity metrics may depend on the context or region under consideration. In this case it was also found that (geographically) neighbouring catchments were usually more hydrologically similar than more distant catchments.

The work by Frey and Dueck (2007) is useful in demonstrating some of the issues inherent in relating catchment characteristics to functional similarity in hydrological response, particularly in relation to the imperfect match usually found between physical descriptors and functional response between catchments. This issue of the relationship between physical and hydrological similarity in catchments was additionally explored by Oudin et al. (2010). Using aridity index, catchment area, mean slope, median altitude, river network density, fraction of forest cover and a BFI estimate,Oudin et al. (2010) compared 10 UK catchments with a large group of potential analogues in France. Using this methodology to test how well the hydrology of physically similar catchments can be predicted when geographical proximity is reduced, it was found that only 60% of catchments were both hydrologically and physically similar. This relatively low degree of correspondence between physical and functional properties was considered likely to be due to either difficulties in accounting for uniqueness and specificity of catchment behaviour or the omission of sufficiently informative descriptors of hydrogeological properties from the catchment characteristics. In contrast to findings underpinning the FEH methods, it was also concluded that the least useful catchment descriptors for determining hydrological similarity in this context were catchment area and BFI, although this result could be at least partly an artefact of attempting to convert HOST classifications into the French equivalent. The most useful catchment characteristics were found to be

slope index, aridity index, drainage density and forest cover. It is interesting to note that the FEH does not include an explicit indicator for geology.

A further step often undertaken using the results from such similarity and grouping analyses (known as pooling) is then to directly transfer parameters from gauged catchments that would be members of each group, based on their characteristics. This is based on the assumption that the same parameter set should be applicable for similar catchments, which in turn presumes an accurate characterisation of catchment similarity. There are a number of studies that apply this method, such as Ouarda et al. (2006)) who modified the approach to examine the effect of pooling catchments based on a seasonality regionalisation method. Cunderlik and Burn (2002) used rainfall seasonality to pool catchments, which allowed catchments with no flow record but a rainfall record to be modelled.

A slightly different approach to catchment classification was undertaken by He et al. (2011), who consider catchments to be similar if many sets of parameters lead to similar model performance and instead of transferring model parameters from a pool of catchments only the most similar catchments are used as parameter donors.

This methodology is limited by the fact that catchment descriptors often do not capture the complexity and heterogeneity of natural catchments (Beven, 1999b). It is also important to consider the impact that climate change will have on catchment descriptors, and it is unclear whether a pool of catchments necessarily be representative of each other under different conditions.

### Regionalisation

Regionalisation of model parameters attempts to extend the simple donor catchment methodology by modelling the relationship between catchment characteristics and model parameters. The assumptions made in regionalisation are 1) that similar parameter sets will result in similar rainfall- runoff behaviours (an assumption inherent in all models) and 2) that physical similarity between catchments will result in hydrological similarity (Oudin et al., 2010). Regionalisation can be carried out using relatively simple statistical methods, such as multiple regression (Oudin et al., 2010; Kokkonen et al., 2003), or more complex methods such as those described later in this section. A two-step procedure is typically applied in which models of gauged catchments are calibrated individually and then the relationship between calibrated parameters and the characteristics of each catchment are determined (Seibert, 1999; Merz and Blöschl, 2004).

However, this method can be limited by the fact that the regression relationship between parameters and catchment characteristics may often be weak, leading to mixed results in the success of predicting parameters of models for ungauged catchments ((Vogel, 2005)). This problem relates at least partly to the problem of equifinality, whereby multiple sets of parameters may produce equally good simulations (Beven and Freer, 2001). If only a single parameter set is determined for the (donor) gauged catchment, only a single relationship between parameters and catchment characteristics can be found. Hundecha et al. (2008) work around this by developing a methodology in which models are calibrated to both maximise performance and to give a well defined structure in physiographic-climatic space. The results show that models perform equally well for catchments used in training/calibration as those used for validation, but that model parameters outside of the training data can only be extrapolated slightly indicating that it is only a useful methodology for transferring parameters between very similar catchments. An alternative approach is demonstrated by Götzinger and Bárdossy (2007), who define transfer functions in which model parameters are a function of flow time, land use, soil properties, area and geology. Instead of calibrating the parameters of the hydrological model directly, the parameters of the transfer function were instead calibrated. This reduces the dimensionality of the parameter space and goes some way to reducing the problem of equifinality.

Several other studies also seek to develop this two-step procedure of calibration followed by relating model parameters to catchment characteristics. Vogel (2005) compares this basic method to one termed regional calibration, whereby model parameters for all catchments in a region are calibrated simultaneously. Based on catchments in the U.S, this latter method resulted in as good a level of calibration essentially equal to the goodness-of-fit obtained by calibrating each catchment individually, again reflecting the pervasive issue of equifinality. However, the regional calibration approach had the advantage of lower water balance bias and much stronger regression relationships between catchment descriptors and model parameters. Interestingly, it was also observed that, for the validation period, both methods gave similar results. This indicates that, despite more robust regression relationships obtained from regional calibration, predictive capacity for ungauged catchments did not necessarily improve relative to using parameters from individually calibrated catchments with more scatter in their relationship to catchment properties. In addition, Li et al. (2010) outline a method for defining catchment similarity and regionalisation of parameters using the 'Index Model' that takes into account the geographical proximity and hydrological similarity of catchments instead of just catchment descriptors. This regionalisation method was

shown to improve hydrological model performance compared with parameter transfer based on linear regression, particularly for catchments where regression relationships are poor.

Engeland et al. (2001) tried a slightly different approach again by parameterising a large-scale regional model with one set of global parameters, in order that the same parameters can be used in the large catchment and all of its sub catchments. The authors used a Bayesian method to ascertain parameter probability distributions conditioned on observations and then used those to select robust parameters for their model. The parameter space was predefined by the authors and then sampled using two techniques. The resulting parameter sets are shown to be narrow in parameter space and provide good simulations.

Regionalisation methods have been shown to perform well across a range of studies, which is to be expected with a good set of donor catchments available. Few studies are able to extend their parameter transfers to dissimilar catchments successfully which is a significant limitation. Physically based modelling then becomes necessary as no parameter transfer is required at all.

### Physically based approach

An alternative approach to working around the lack of data available for ungauged basins is to apply a physically-based model using parameters known a priori from measured datasets as far as possible. This solution is not always possible for various reasons. For example, sufficient input data may not exist, the model structure may not capture all relevant processes in a catchment, the model may be difficult to apply at the required scale and calibration may still be required to achieve an acceptable degree of accuracy in runoff predictions (Hundecha et al., 2008). This last point relates to a number of the common criticisms of physically-based modelling approaches, such as those regarding scale and parameter estimation put forward by Beven (2001). However, blind validation tests of physically based models have been conducted that show with good catchment information, physically based models can perform well without calibration (Bathurst et al., 2004).

Several recent studies have also explored how physically-based modelling approaches compare with methods based on empirical and conceptual models for simulating ungauged catchments. Booker and Woods (2014) compare model performance for 485 catchments across New Zealand using an uncalibrated physically-based model (TopNet) with predictions from a purely empirical machine-learning regression model (Random

Forests). Booker and Woods found that the empirical approach using Random Forests outperformed TopNet across all performance measures, which is not unsurprising, given that tuning of parameters in empirical or conceptual models regularly achieves very high goodness-of-fit. Booker and Woods (2014) note that TopNet has been demonstrated to provide more accurate simulations in other applications, but clearly some calibration would be required to achieve this. In this case, applying TopNet to an ungauged basin would require the transfer of parameters. It would be interesting to see what methods of parameter transfer would work best in this case, particularly with regard to whether direct transfer of parameters linked to physical properties could be undertaken or whether more complicated methods were required for accurate simulation. If simple, direct transfers were not possible, it would potentially raise significant questions about the application of physically-based models to ungauged basins, at least in the context of this model structure and geographical/hydrological context.

In contrast to Booker and Woods (2014), however, a number of studies applying physically-based models to ungauged catchments show much more promising results. For example, Fang et al. (2010) demonstrated that an uncalibrated physically-based hydrological model of an agricultural catchment - using parameters derived a priori and a LiDAR DEM - could perform as well as a calibrated model, mainly due to the high resolution DEM data available. This raises some interesting issues of the extent to which it is in fact inadequate data that sometimes compromises the application of physically-based models.

Fang et al. (2013) went on to show that physically-based models with the correct structure, flexibility and parameters estimated from measurements on-site, or in similar conditions, can provide robust estimation of snowpack, soil moisture and streamflow across multiple scales, by using an uncalibrated physically-based model that included a full-suite of snow and cold region processes for the Marmot Creek research basin in Canada.

The strong potential of physically-based models in this respect is also apparent in work by Dornes et al. (2008). Unlike the regionalisation methods described above, which tend to work best on neighbouring catchments, Dornes et al. (2008) showed that parameters with a physical meaning could be transferred to catchments at a great distance when using physically-based models. This was achieved through calibrating a physically based model for a catchment and then transferring those parameters, as their physical meaning was well understood, to another catchment 1350km away.

## 2.4 Conclusion

This study aims to create a national modelling system using a physically based model, SHETRAN, that can be used for modelling both gauged and ungauged catchments robustly. This literature review has therefore summarised the types of hydrological models available, the pros and cons of physically based modelling, SHETRAN and its uses, other national modelling systems and approaches taken for modelling ungauged basins. Physically based models and SHETRAN in particular, whilst unpopular, offer real value in their ability to provide robust estimations of many variables whilst also remaining perfectly clear in their processes allowing the modeller to have a better understanding of the system they are studying. There is therefore much to be gained by having such a model for Great Britain.

# Chapter 3

# Creating SHETRAN for GB

## 3.1 Modelling the whole of Great Britain

A central objective of this research is to create a system that enables the automatic setup of a robust, physically-based spatially-distributed model for any catchment, either gauged or ungauged, in Great Britain. This necessitates the collation, processing and storage of large amounts of data, as well as methods to retrieve and configure catchment models with minimal user-intervention. This chapter outlines the data sources, modifications made to SHETRAN and automated processes used to build this system.

## 3.2 Great Britain as a study area

Great Britain is a relatively small, self-contained and data-rich region with reasonable heterogeneity of hydrology, which makes it suitable for developing and testing a national hydrological modelling system in terms of practical tractability and availability of input and evaluation data. There is also a long history of research into catchment processes in Great Britain, as well as many previous hydrological modelling studies, which means that catchment behaviour and processes are relatively well researched. This is an important benefit for assessing the modelling system, as conceptual understanding of catchment processes and their variation across Great Britain can be used to evaluate whether a model is simulating runoff generation mechanisms in a plausible way. Furthermore, there is an interesting range of catchment types across Great Britain, with significant variation in topography, geology, land use and climate amongst other factors. This heterogeneity presents an opportunity to explore the performance

of the SHETRAN modelling system across large numbers of catchments with varying properties, which could help to identify strengths and limitations of the system and its inputs.

The SHETRAN model is configured to run separately for individual catchments, rather than for multiple catchments in a given execution of the program. The catchments simulated in this work were determined on the basis of the records for all of the gauged catchments in the UK held by the National River Flow Archive (NRFA) (National River Flow Archive, 2014b). There are 1537 gauged catchments in the UK, of which 306 have freely available historic flow data, lie within Great Britain and are of an appropriate catchment area to be modelled with a 1km SHETRAN model (i.e. not so small that the grid size is too coarse to represent the catchment and not so large that processing times become unreasonable). These 306 catchments were therefore selected as the sample for consideration in this study, in order to evaluate the performance of the modelling system against historic daily flow data.

The suitability of these 306 catchments for investigating performance of the modelling system across a range of catchment types is supported by their good geographical distribution across Great Britain, as shown in Figure 3.1. From this figure, it can be seen that the 306 catchments cover the whole of Great Britain, ranging from impermeable upland catchments in Scotland to flat catchments underlain by major chalk aquifers in East Anglia. The associated variation in catchment properties is significant, as discussed further in Chapter 4 in relation to model performance. A full list of the catchments used in this study can be found in Appendix A. Figure 3.1 also highlights that a substantial range of catchment sizes are also covered by this sample. Some of the smallest and very largest catchments in Great Britain were omitted due to the spatial resolution used in this study and computational limits of the parallel processing infrastructure used in this work (discussed further below).

For each catchment, the NRFA provides descriptions of the setup of the gauging station and any associated flow measurement issues, a characterisation of the physical features of the catchment and flow statistics amongst other information. Importantly, human factors affecting runoff are also listed by categorising the artificial influences on flow as follows:

- Natural catchment - variation due to abstractions and discharges is so limited that the gauged flow is considered to be within 10% of the natural flow when it is greater than or equal to the Q95 flow

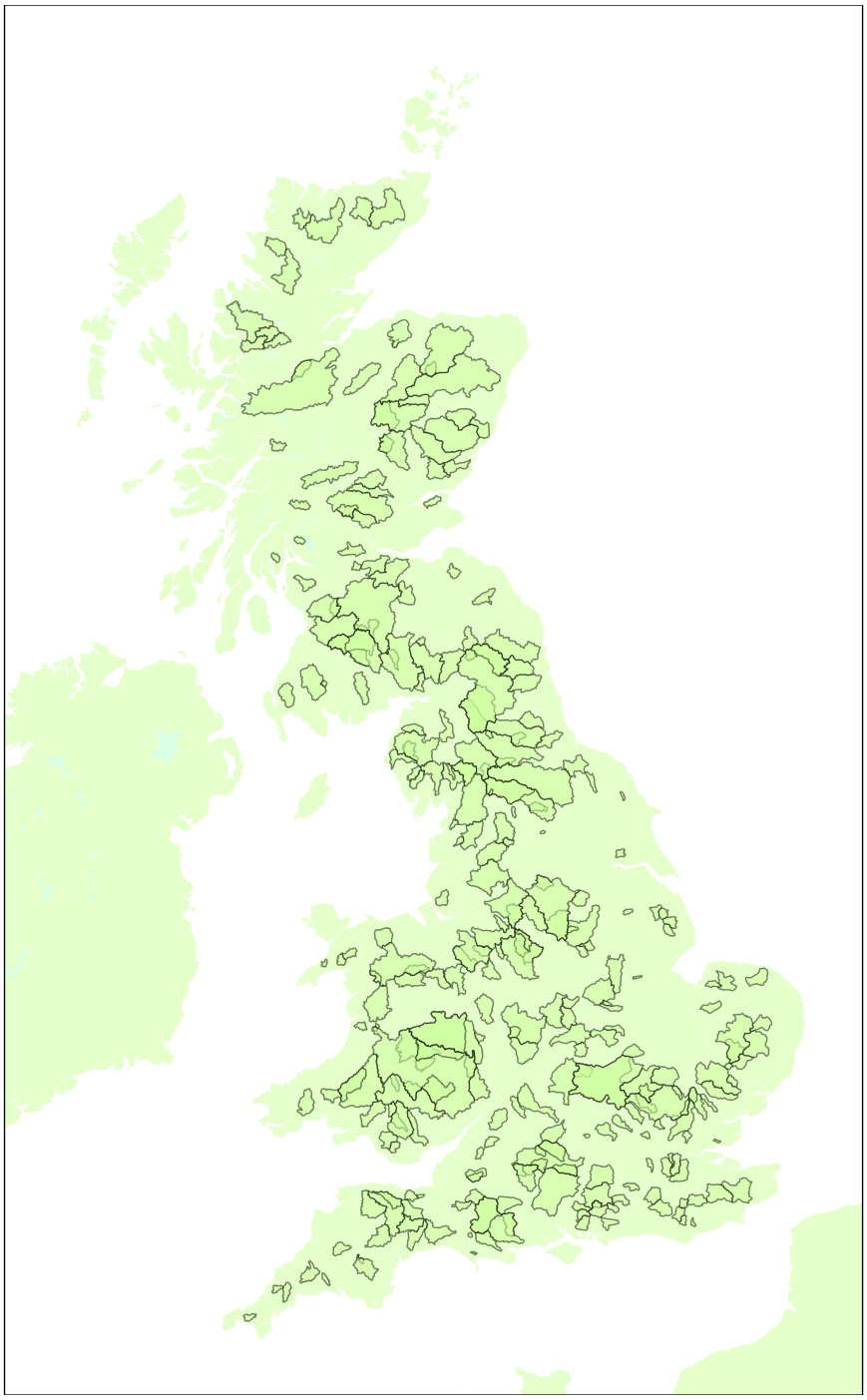


*Figure 3.1: The 306 catchments used in this study.*

- Storage or impounding reservoir - natural river flows are affected by water stored in a reservoir situated in, and supplied from, the catchment
- Regulated river - under certain flow conditions the river will be augmented from surface water and/or groundwater storage
- Public water supplies - natural runoff is reduced by the quantity abstracted from a reservoir or by a river intake if the water is conveyed outside the gauging station's catchment area
- Groundwater abstraction - natural river flow may be reduced or augmented by groundwater abstraction or recharge, including mine-water influences
- Effluent return - outflows from sewage treatment works will augment the river flow if the effluent originates from outside the catchment
- Industrial and agricultural abstractions - direct industrial and agricultural abstractions from surface water and from groundwater may reduce the natural river flow
- Hydro-electric power - river flow is regulated to suit the demands for power generation; catchment to catchment diversions may also significantly affect average runoff.

(Descriptions taken from the NRFA website (National River Flow Archive, 2014b))

## 3.3 Data collected

Physically based hydrological models are very data intensive. SHETRAN requires:

- A digital elevation model (DEM)
- A map describing the subsurface properties of a catchment
- A land cover map
- Rainfall
- Potential evapotranspiration (PET)
- A mask delineating the watershed of the catchment.

See Figure 3.2. The datasets used are outlined in Table 3.1 and the details are discussed in subsequent sections.

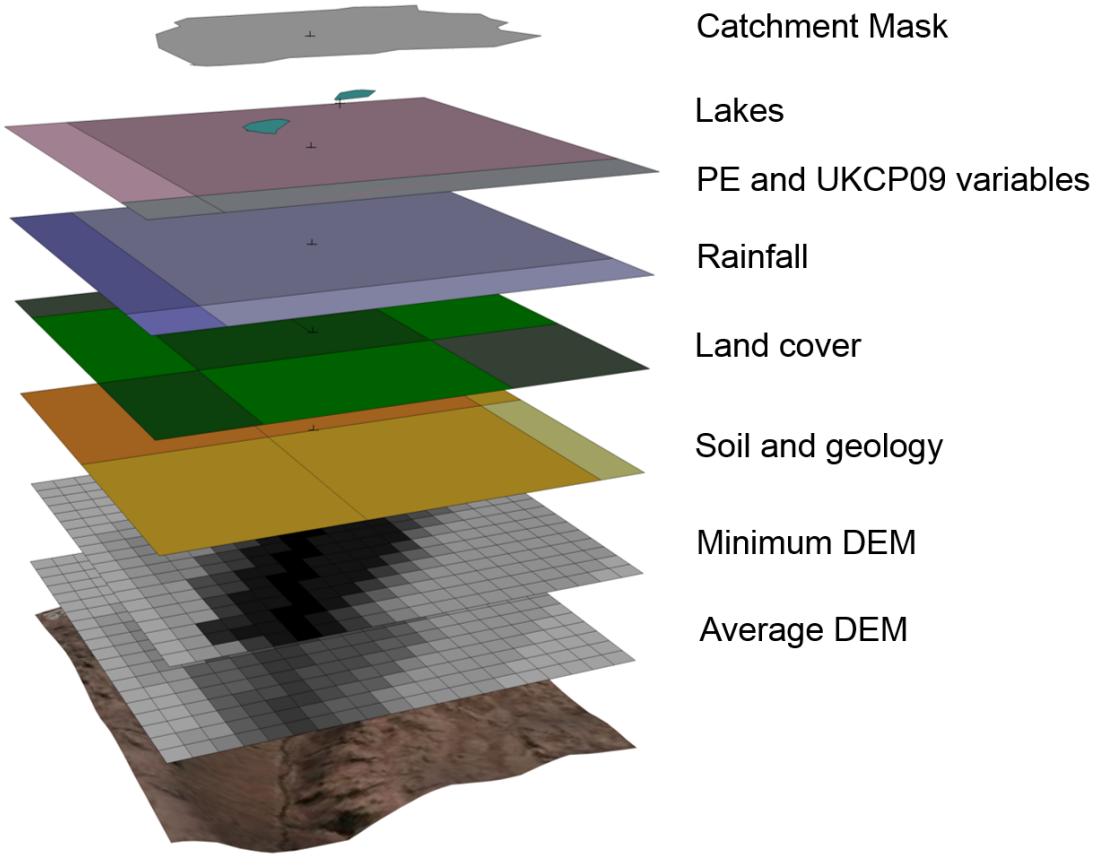



*Figure 3.2: A diagram showing the data layers included in the SHETRAN for GB system.*

The datasets chosen for input to the modelling system each cover the whole of Great Britain, which means that the information supplied for each catchment is consistent. This has the advantage that interpreting variations in model performance will not be additionally complicated by the confounding effects of using multiple datasets for different catchments. With the exception of the rainfall inputs, the datasets are all also freely available to download under an academic licence, making the models suitable for use by any research group. Recent developments are such that an alternative, freely available, gridded daily rainfall dataset for the UK could potentially be applied within the modelling system (Tanguy et al., 2014), although the suitability of this dataset would need to be evaluated first in future work. It should also be noted that the datasets used are the highest quality ones available for this research, as described below for each dataset. On this basis it has been presumed that the quality control processes undertaken in the construction of each dataset make them sufficient for direct application in the modelling system without any further quality review or refinement.

The fully distributed nature of SHETRAN means that model structure and properties need to be specified for each grid cell. All maps required as part of this were resampled to a 1km resolution and aligned with the British National Grid (BNG) for consistency and ease-of-use. This resolution is partly dictated by the scales of the available soil and land cover datasets, but it could be argued of course that higher spatial resolution models may be desirable, in order to more realistically capture the significant heterogeneity that can occur at sub-kilometre scales in catchments. For example, with respect to the representation of soils, it is recognised that effective parameters are required at this comparatively coarse resolution to implicitly account for the many soil types and complex variations in structure and properties that could be present within a single grid square. Some authors, such as Beven (2006), argue that this compensation by effective parameterisation could undermine the physical basis of the model to some extent, as discussed in Chapter 2. It is suggested here that the extent to which this issue represents a problem can be evaluated with respect to model performance, however. If the dynamics of catchment models appear to be conceptually plausible and consistent with available evaluation data, the selected spatial resolution would seem justifiable for the purpose of national-scale modelling. This point is assessed further in Chapter 5.

The spatial resolution adopted here therefore represents a balance between the information content of available data and pragmatic considerations with respect to computing resources and run times particularly, as well as perceived priorities for investigation and evaluation. This was also guided by the recent work of Zhang (2012), who investigated the influence of spatial and temporal resolution in SHETRAN. Zhang (2012) found

that simulations can be significantly improved by increasing the temporal resolution of forcing inputs - particularly rainfall - from daily to hourly intervals. In contrast, improvements due to increasing spatial resolution were found to be more varied and depend upon the level of catchment homogeneity. This highlights the potentially universal importance of temporal resolution of inputs relative to spatial resolution, such that the former was selected as a priority for further investigation, as reported in Chapter 5. Furthermore, the 1km resolution applied here is consistent with other large-scale, national modelling studies, such as the the UK Grid-To-Grid model (Bell et al., 2007a) and the national MIKE SHE model for Denmark (Henriksen et al., 2003), which also use a 1km grid resolution.

The spatial datasets used for model construction were additionally resampled to 500m and 100m for flexibility. These higher resolution datasets were not used in this study, but future work could take advantage of them to investigate the effects of spatial resolution in more detail.

| Data Type | Source | Description | Reference | Licence |
|---|---|---|---|---|
| Catchment boundaries | National River Flow Archive (2012) | Shapefiles of 1170 UK catchment boundaries | Morris et al. (1990) | Centre for Ecology and Hydrology (2014) |
| Lakes | Ordnance Survey (2012) | Shapefiles of lakes. | Ordnance Survey (2013b) | Ordnance Survey (2014) |
| UKCP09 Daily Maximum Temperature | The Met. Office (2012) | 5km ascii files, one per day. Used to calculate PET and snowmelt. | Perry et al. (2009) | Department for Environment, Food and Rural Affairs (2014) |
| UKCP09 Daily minimum Temperature | The Met. Office (2012) | 5km ascii files, one per day. Used to calculate PET and snowmelt. | (Perry et al., 2009) | Department for Environment, Food and Rural Affairs (2014) |
| UKCP09 Monthly Relative humidity | The Met. Office (2012) | 5km ascii files, one per month. Used to calculate PET. | Perry and Hollis (2005) | The National Archives (2014) |
| UKCP09 Monthly Wind Speed | The Met. Office (2012) | 5km ascii files, one per month. Used to calculate PET. | Perry and Hollis (2005) | The National Archives (2014) |
| UKCP09 Monthly Sunshine hours | The Met. Office (2012) | 5km ascii files, one per month. Used to calculate PET. | Perry and Hollis (2005) | The National Archives (2014) |
| UKCP09 Daily Rainfall | The Met. Office | 5km ascii files, one per day | Perry et al. (2009) | Department for Environment, Food and Rural Affairs (2014) |
| Land Cover Map 2007 | Centre for Ecology and Hydrology (2012) | 1km raster | Morton et al. (2011) | Morton et al. (2011) |
| Soil map | European Commission, Joint Research Centre (2012) | Four 1km rasters | Liedekerke et al. (2006) | European Commission, Joint Research Centre (2014) |
| Hydrogeology map | British Geological Survey (2012) | Shapefile | British Geological Survey (2014) | British Geological Survey (2012) |
| Digital Elevation Model | Ordnance Survey (2012) | 50m raster | Ordnance Survey (2013a) | Ordnance Survey (2014) |
| Flow data | National River Flow Archive (2012) | Individual .csv files | National River Flow Archive (2014a) | Centre for Ecology and Hydrology (2014) |

*Table 3.1: Summary table of datasets used and their associated information.*

### 3.3.1 DEM

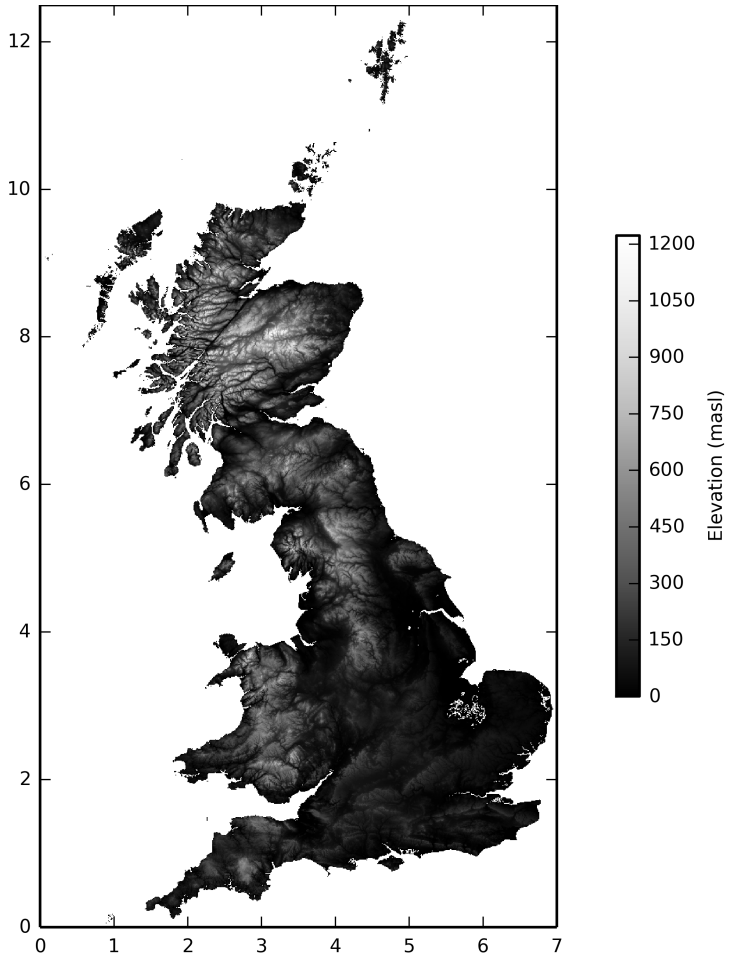



*Figure 3.3: The 1km resolution DEM.*

The digital elevation model (DEM) used in this study was based on the Ordnance Survey (OS) Land-Form Panorama data. This 50m resolution raster was derived from stereo areal photography taken in the 1970s, which was used to create OS Landranger 1:50 000 contours with an absolute accuracy (defined by root-mean-square error (RMSE)) of typically better than 3m. The Land-Form Panorama DEM was created from these contours using interpolation and generally has a vertical accuracy of 5m or better (Ordnance Survey, 2013a). The DEM was downloaded in 20km x 20km tiles, which were then processed in ArcGIS using the mosaic to raster tool to create one raster file covering the whole of Great Britain. This raster was then resampled to a 1km resolution using bilinear resampling to determine the average elevation for each

grid square. Figure 3.3 shows the 1km DEM resulting from this procedure, which illustrates the variation in topography across Great Britain. The figure shows how steep catchments are generally found in areas of Scotland, Wales and the north of England, which may be contrasted with areas such as the south-east of England that are very flat.

One of the applications of the DEM in SHETRAN is in generation of the location of river channels for routing within the model. These are referred to in SHETRAN as river links. Preliminary tests for this work suggested that these links more closely follow actual river paths when the minimum elevations in each grid square - taken from the original resolution of the Panorama DEM - are accounted for (S. Birkinshaw, pers. comm., 2012). A DEM based on the minimum elevation values in each 1km grid square was therefore created, using the same method as described above except with the resampling based on minimum elevations from the Panorama dataset.

It should also be acknowledged that there are many other DEMs available for the UK. For example, the Environment Agency offers much higher spatial resolution DEMs (0.25m to 2m) with vertical accuracy of +/- 0.15m produced from airborne Light Detection and Ranging (LiDAR) (Environment Agency, 2014). However, these data only cover around 50% of England and Wales, and such high resolution is not considered relevant for catchment modelling at the scales used in this study. The Ordnance Survey have also recently released a new height product for GB, which is known as OS Terrain 50 (Ordnance Survey, 2013c). This product, with a reported absolute accuracy of 4m RMSE, originates from a different capture process to the Panorama data used here. However, OS Terrain 50 was released in April 2013, too late for incorporation in this study. Future versions of SHETRAN for GB may benefit from using OS Terrain 50 as the basis for the DEM, as the latter is an actively maintained dataset, in contrast to the Panorama dataset. However, the differences between the two products are likely to be small and of limited significance when modelling at the spatial resolutions used in this study, although further work could test this hypothesis.

### 3.3.2 Geology and Soil

SHETRAN represents the catchment subsurface as a column containing multiple layers of soil or rock. There is one such column for each grid cell in the model, and the parameters required for each layer of the column are as follows:

- The depth of the layer

- Saturated water conductivity, $K_s$
- vanGenuchten-$\alpha$, relating to the inverse of air entry suction (Van Genuchten, 1980)
- vanGenuchten-n, a measure of the pore-size distribution
- The residual moisture content, $\theta_r$
- The saturated moisture content, $\theta_s$.

Soil data obtainable for free under an academic licence are not widely available. There are some national soil data sets, such as HOST (Boorman et al., 1995) and the LandIs Digital Soil Dataset (National Soil Resources Institute, 2001), although licences to use both of these datasets need to be purchased at reasonable expense. The LandIs Digital Soil Dataset from Cranfield University contains 297 distinct soil associations (classes), each containing multiple soil horizons. It is possibly the most detailed, widely available soil dataset for the UK. However, it only covers England and Wales, making it unsuitable for this project. The HOST dataset is the most widely used soil data source, but the parameters associated with the soil types do not easily map on to SHETRAN parameters. This is because HOST data is based on interpretation of raw soil data with consideration to the dominant runoff processes using a very simplistic, conceptual model (Ali et al., 2012). Another option is the Food and Agriculture Organization (FAO) harmonized world soil database (Nachtergaele and Batjes, 2012), which is comprehensive and provides a good basis for conducting global comparisons, containing as it does a globally consistent set of parameters.

However, an alternative data product, the European Soil Database (ESDB) v2.0 (Liedekerke et al., 2006) was identified as the most suitable dataset for meeting the requirements of this study and modelling system. This based on the fact that the ESDB dataset contains all of the relevant data required for model input and is the most UK-focussed, freely available dataset. It is Europe wide with a 1km resolution and hydraulic properties were assigned by a collaboration of 12 European countries. Both the particle size and the hydraulic data were standardized across Europe by fitting the Mualem-van Genuchten model parameters (Van Genuchten, 1980) to the individual (h) and K(h) hydraulic properties stored in the ESDB.

The ESDB was downloaded as several separate raster layers. The ESDB relates the parameters required by SHETRAN to soil texture, for example a coarse soil a fine soil, peat, etc. The layers containing the information that SHETRAN uses - dominant topsoil texture, depth to textural change, dominant subsoil texture and depth to rock - were selected and combined into one raster file of unique soil classes using a Python script.

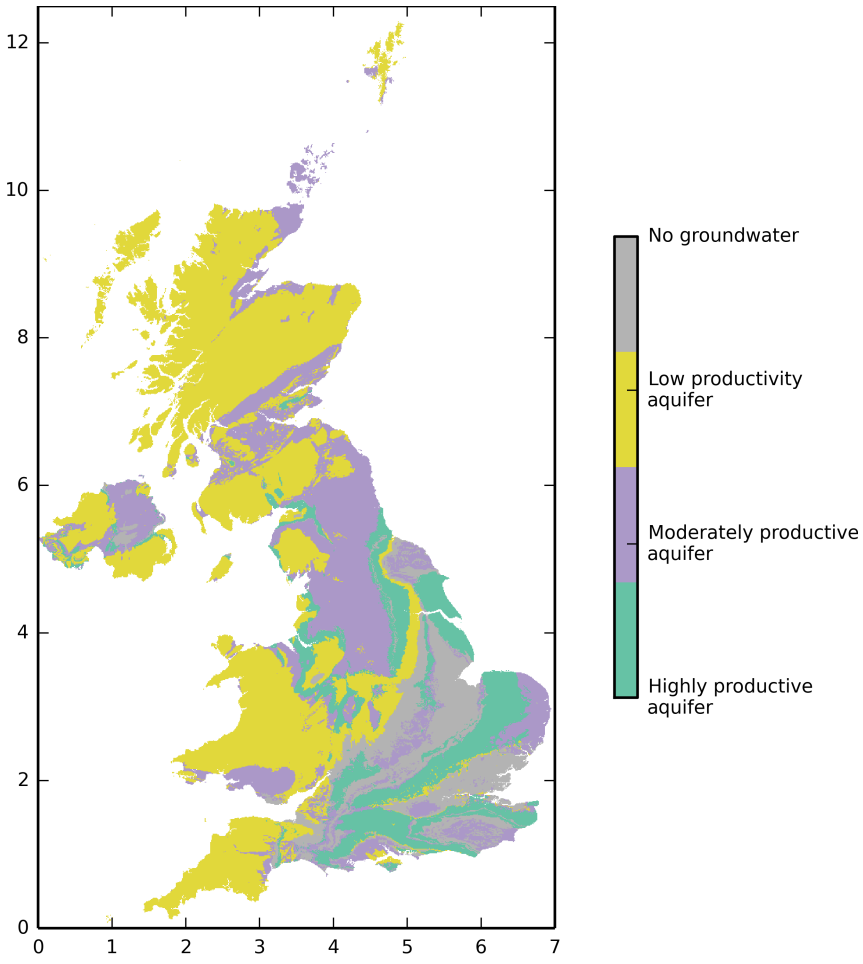



*Figure 3.4: The BGS hydrogeology map.*

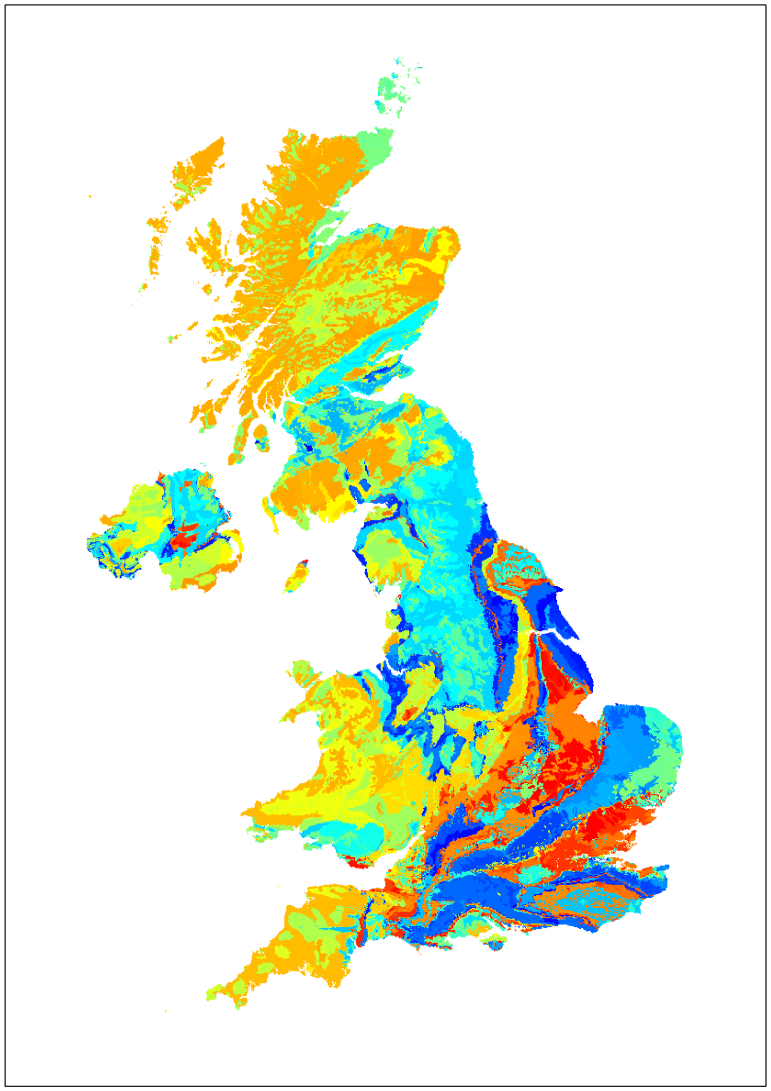


*Figure 3.5: The subsurface map of the UK created from combining layers from the ESDB and the BGS hydrogeology layer. The different colours indicate different combinations of properties.*

In an attempt to setup physically realistic catchment models, a geological layer was added to the bottom of each soil column. The data for this were taken from the British Geological Survey (BGS) 1:625 000 scale digital hydrogeological map (British Geological Survey, 2014) (see Figure 3.4). This map classifies the bedrock of the UK into a small number of aquifer types based on productivity and is the only GB-wide map describing aquifer properties, such that it formed the only real option for including aquifer types into the model (although see Chapter 7 for further recent developments). It can be seen from Figure 3.4 that there is notable variation in aquifer types across the country. The highly productive aquifers are mainly found in the chalk of south-east England and East Yorkshire, whereas most of Scotland and Wales are underlain by impermeable igneous rocks and mudstones, resulting in low productivity aquifers.

The BGS hydrogeology map is available in vector format and was converted to a raster for combination with the soil layer derived from the ESDB. 236 unique subsurface column types were identified on this basis and coded for use in SHETRAN. The distribution of these column types can be seen in Figure 3.5. As the parameters that SHETRAN uses are primarily associated with soil and not included with the BGS hydrogeology map, representative hydraulic property values were therefore assigned to the aquifer descriptions based on typical SHETRAN parameters for similar rock types from previous modelling experience (S. Birkinshaw, pers. comm. 2012). Each subsurface type was assigned the parameters shown in Table 3.2. SHETRAN is given hydraulic properties for each layer on a column based on its type where the soil depths are determined by the ESDB layers and bedrock depth was assigned an arbitraty depth of 20m.

### 3.3.3 Land cover

The CEH land cover map (LCM) 2007 (Morton et al., 2011) is the most up-to-date land cover map freely available for academic use for the UK. The map is derived from satellite images and digital cartography, using land cover classifications based on the UK Biodiversity Action Plan Broad Habitats that lead to the definition of 23 land cover types. The map was simplified into the 7 basic land cover types typically used with SHETRAN: arable, bare ground, grass, deciduous forest, evergreen forest, shrub and urban, as the differences between the original 23 groups are not easily distinguishable by SHETRAN parameters (see Table 3.3 and Figure 3.6) (Birkinshaw, 2011). Figure 3.7 summarises the proportion of each catchment assigned to different land cover types, which indicates that most catchments are generally only dominated by grass, shrub or

| EU code | soil type | $\theta_s$ | $\theta_r$ | $K_s$ (m/day) | vG-$\alpha$ ($cm^{-1}$) | vG-n |
|---|---|---|---|---|---|---|
| Texture class of the dominant surface soil type | | | | | | |
| 0 | Peat | 0.766 | 0.010 | 8 | 0.0130 | 1.2039 |
| 9 | Peat | 0.766 | 0.010 | 8 | 0.0130 | 1.2039 |
| 1 | Coarse (<18%clay And <65%sand) | 0.403 | 0.025 | 60.000 | 0.0383 | 1.3774 |
| 2 | Medium (18-35%clay And <15%sand Or <18%clay And 15-65%sand) | 0.439 | 0.010 | 12.061 | 0.0314 | 1.1804 |
| 3 | MediumFine (<35%clay and <15%sand) | 0.430 | 0.010 | 2.272 | 0.0083 | 1.2539 |
| 4 | Fine (35-60%clay) | 0.520 | 0.010 | 24.800 | 0.0367 | 1.1012 |
| 5 | VeryFine (>60%clay) | 0.614 | 0.010 | 15.000 | 0.0265 | 1.1033 |
| Texture class of the dominant subsurface soil type | | | | | | |
| 0 | Peat | 0.766 | 0.010 | 8 | 0.0130 | 1.2039 |
| 9 | Peat | 0.766 | 0.010 | 8 | 0.0130 | 1.2039 |
| 1 | Coarse (<18%clay And <65%sand) | 0.366 | 0.025 | 70.000 | 0.0430 | 1.5206 |
| 2 | Medium (18-35%clay And <15%sand Or <18%clay And 15-65%sand) | 0.392 | 0.010 | 10.755 | 0.0249 | 1.1689 |
| 3 | MediumFine (<35%clay and <15%sand) | 0.412 | 0.010 | 4.000 | 0.0082 | 1.2179 |
| 4 | Fine (35-60%clay) | 0.481 | 0.010 | 8.500 | 0.0198 | 1.0861 |
| 5 | VeryFine (>60%clay) | 0.538 | 0.010 | 8.235 | 0.0168 | 1.0730 |
| Texture class of the hydrogeology layer | | | | | | |
| 1 | Highly Productive Aquifer | 0.3 | 0.2 | 0.1 | 0.01 | 5 |
| 2 | Moderately Productive Aquifer | 0.3 | 0.2 | 0.01 | 0.01 | 5 |
| 3 | Low Productivity Aquifer | 0.3 | 0.2 | 0.001 | 0.01 | 5 |
| 4 | No Groundwater | 0.3 | 0.2 | 0.0001 | 0.01 | 5 |

*Table 3.2: Summary table of the soil and rock types used in with SHETRAN and their properties.*

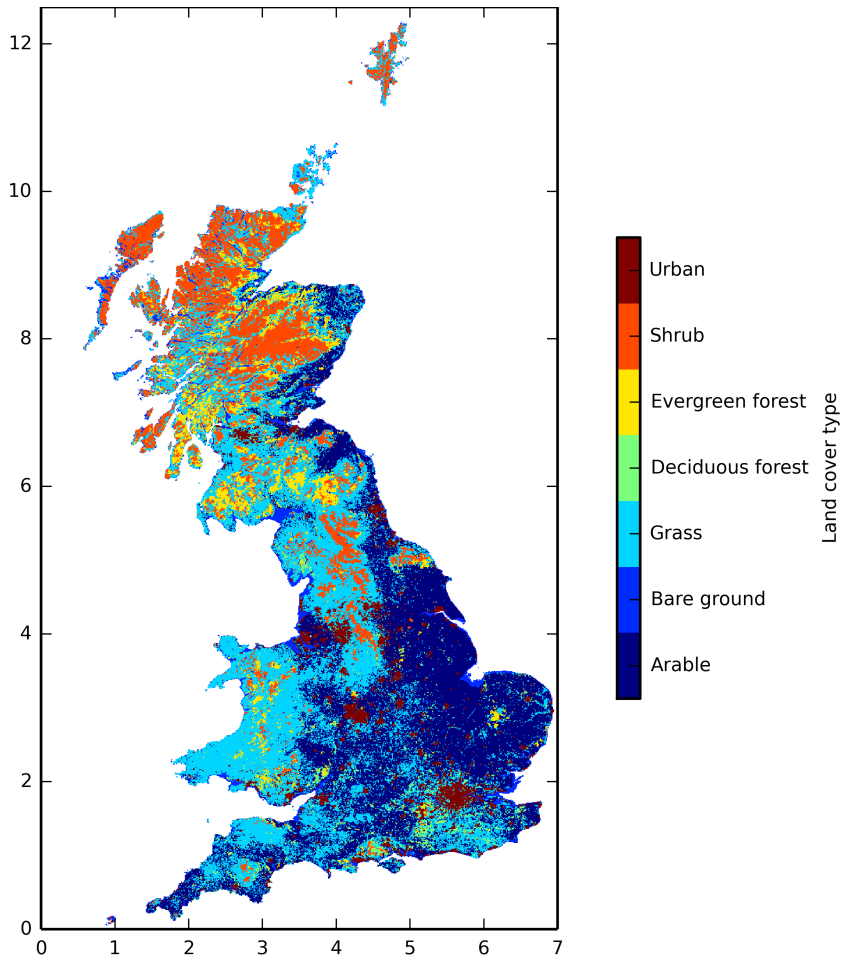



*Figure 3.6: A map showing the variation of land cover types across the UK. The dominant types are arable (dark blue), grassland (pale blue), urban (red), shrub (orange) and forest (green and yellow).*

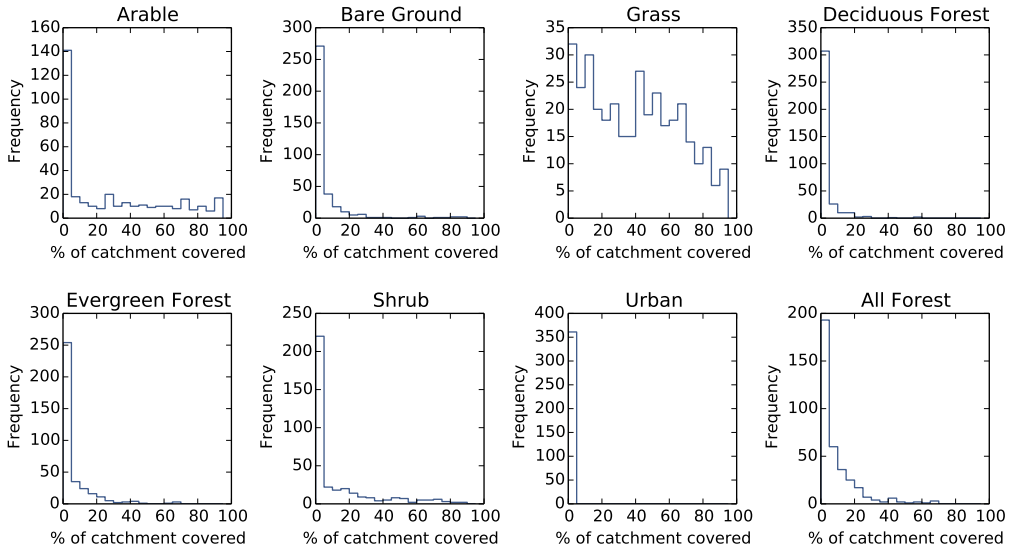



*Figure 3.7: Histograms showing % area of catchment of each land cover type. They show that the most dominant land cover types within catchments are arable, grass and shrub.*

arable land covers. The 7 land cover classes used in SHETRAN were assigned the parameters detailed in Table 3.4, which are based on the SHETRAN documentation (Birkinshaw, 2011).

Highly detailed maps of land cover and vegetation types are available from other sources. For example, Natural England hold maps of agricultural land, ancient woodland, heathland and grassland amongst other habitat types at a 100m resolution for England only (Natural England, 2014). The Joint Nature Conservation Committee also hold very detailed vegetation data for some areas of GB (Joint Nature Conservation Committee, 2014). However, this level of detail is not likely to be required in the SHETRAN modelling system when using a spatial resolution of 1km.

### 3.3.4 Rainfall

In most previous work with SHETRAN, point rainfall data has tended to be used, with rainfall time series obtained from the Environment Agency and the Met Office for example. This approach is adequate for modelling individual catchments, but a more consistent data set is necessary for setting up a nationwide system. SHETRAN was therefore updated to take gridded rainfall as input and the UKCP09 5km gridded precipitation dataset was selected as the fundamental rainfall input for the modelling system (Perry et al., 2009). This dataset created by the Met Office is based on a large

| SHETRAN type | SHETRAN code | LCM2007 class | LCM2007 number |
|---|---|---|---|
| Arable | 1 | Arable and Horticulture<br>Arable | 3 |
| Bare ground | 2 | Inland Rock<br>Supra-littoral Rock<br>Supra-littoral Sediment<br>Littoral Rock<br>Littoral sediment<br>Saltmarsh<br>Saltwater<br>Freshwater | 14, 15, 16, 17, 18, 19, 20, 21 |
| Grass | 3 | Improved Grassland<br>Rough grassland<br>Neutral Grassland<br>Calcareous Grassland<br>Acid grassland<br>Fen, Marsh and Swamp | 4, 5, 6, 7, 8, 9 |
| Deciduous Forest | 4 | Broadleaved woodland | 1 |
| Evergreen Forest | 5 | Coniferous Woodland | 2 |
| Shrub | 6 | Heather<br>Heather grassland<br>Bog<br>Montane Habitats | 10, 11, 12, 13 |
| Urban | 7 | Urban<br>Suburban | 22, 23 |

*Table 3.3: Table showing the conversion of LCM2007 land cover types to standard SHETRAN land cover types.*

| Veg No | Vegetation Type | Canopy storage capacity (mm) | Leaf area index | Maximum rooting depth(m) | AE/PE at field capacity |
|---|---|---|---|---|---|
| 1 | Arable | 1.5 | 1 | 0.8 | 0.6 |
| 2 | BareGround | 0 | 0 | 0.1 | 0.4 |
| 3 | Grass | 1.5 | 1 | 1.0 | 0.6 |
| 4 | DeciduousForest | 5 | 1 | 1.6 | 1.0 |
| 5 | EvergreenForest | 5 | 1 | 2.0 | 1.0 |
| 6 | Shrub | 1.5 | 1 | 1.0 | 0.4 |
| 7 | Urban | 0.3 | 0.3 | 0.5 | 0.4 |

*Table 3.4: Summary table of the land cover types used in SHETRAN and their properties.*

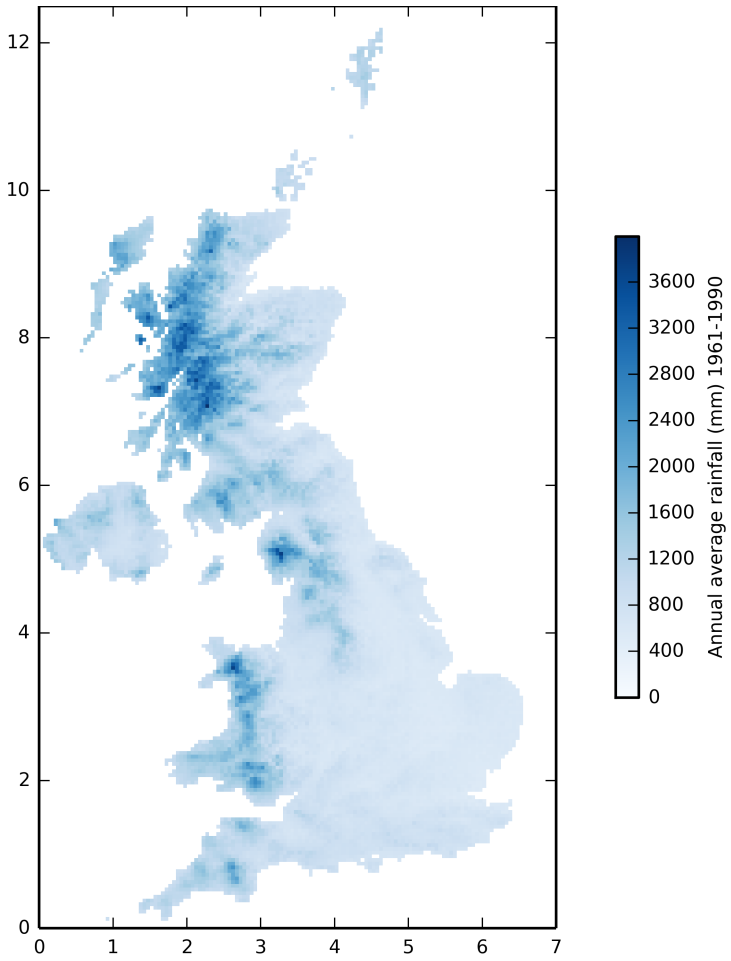



*Figure 3.8: Map showing average annual rainfall. The west coast experiences much higher rates of rainfall than the east.*

amount of data from the UK's comparatively dense gauge network and provides full coverage of the UK at a daily resolution for the period 1961 to 2007. Figure 3.8 shows average annual rainfall calculated from the dataset for the period 1961-1990, which highlights how the west coast and Scotland, where average annual rainfall can exceed 3600mm, are considerably wetter than the east coast and south-east of England, where average annual rainfall can be less than 600mm.

The UKCP09 daily rainfall grids were derived using regression and interpolation (inverse-distance weighting) of data from irregularly distributed stations across the UK. It should be noted that the procedures used to create the precipitation grids do not account explicitly for the influences of latitude, longitude, altitude, terrain, proximity to the coast and urban land use (the latter being considered in the development of the equivalent temperature products) (Perry et al., 2009). In areas of high gauge density this may result in accurate gridded estimates, however, in mountainous areas higher elevations are typically under-sampled and so the gridded estimates may be too low. The data were originally downloaded as 5km ascii raster grids, one for each day, and a Python script was written to convert the gridded data into a set of time series for each 5km grid cell as this is the format required for SHETRAN.

Gridded precipitation datasets for the UK are still relatively small in number, but CEH have very recently developed a new 1km gridded daily rainfall product, CEH-GEAR (Gridded Estimates of daily and monthly Areal Rainfall for the United Kingdom (1890-2012)) (Tanguy et al., 2014). Unlike the UKCP09 5km gridded dataset, which uses inverse-distance weighting to interpolate station data into a gridded format, CEH-GEAR uses natural neighbour interpolation. The gridded values are also adjusted by a monthly correction factor. This new rainfall product could easily be formatted for use with SHETRAN for GB, which would then allow for it to be tested relative to the UKCP09 grids. As the UKCP09 5km gridded rainfall is the only dataset used in this study that is not freely available to download and CEH-GEAR will be freely available for all purposes, this might form a useful additional exercise.

### 3.3.5 PET

The Met Office Rainfall and Evaporation Calculation System (MORECS) is often used as an evaporation data source for hydrological modelling in the UK (Hough and Jones, 1997), for example the Environment Agency regional groundwater models (Shepley et al., 2012). It provides nationwide, real-time assessments of rainfall, potential evapotranspiration (PET) and soil moisture (Hough and Jones, 1997), but the data are not

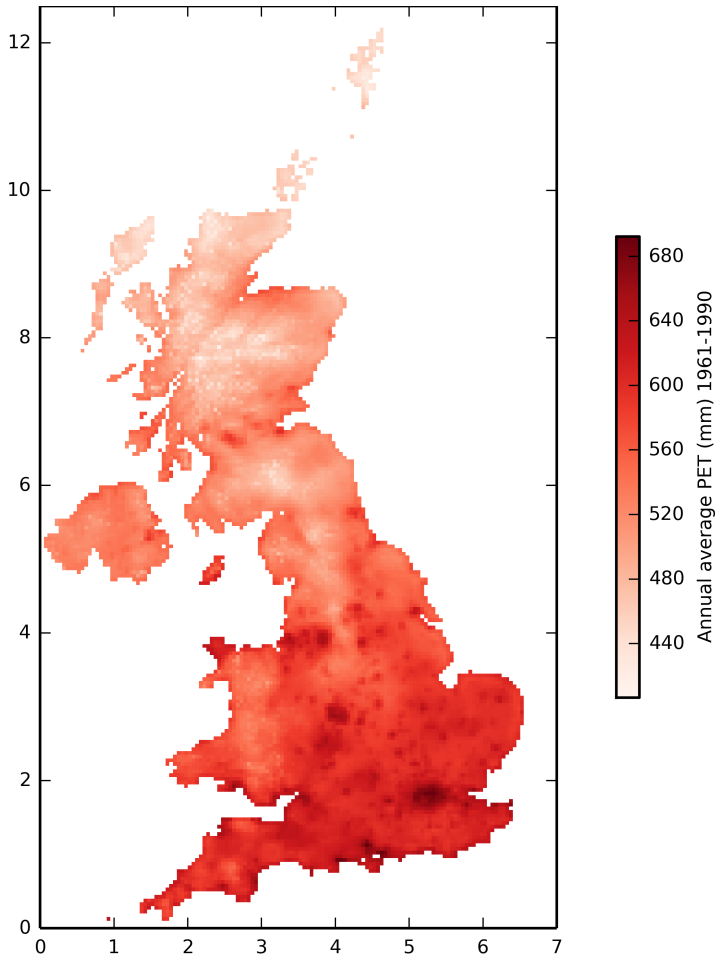



*Figure 3.9: Map showing average annual PET calculated from UKCP09 meteorological variables. The south east of England experiences much higher rates of PET than the north. PET rates are also higher in urban areas indicated by the dark red spots on the map.*

freely available. Other distributed estimates of PET are very limited, such that the approach taken in this work was to calculate PET directly from the gridded variables available within the UKCP09 dataset. This method also allows the SHETRAN for GB system to be more directly compatible with UKCP09 weather generator outputs. Validation of the approach of calculating PET in this way is examined in Chapter 5, where the sensitivity of model performance to PET is assessed. However, it would be an interesting additional check to compare the calculated PET dataset against MORECS or MOSES, while future plans to couple SHETRAN for GB with the land surface scheme JULES (Best et al., 2011) may provide another option.

Potential evapotranspiration time series were calculated from the UKCP09 5km gridded climate variables using the FAO Penman-Monteith method (Allen et al., 1998) to produce a UK wide 5km x 5km grid. The variables incorporated were:

- Maximum temperature 1960-2006 (daily timestep)
- Minimum temperature 1960-2006 (daily timestep)
- Sunshine hours 1961-2006 (monthly timestep)
- Relative humidity 1961-2006 (monthly timestep)
- Mean wind speed 1969-2006 (monthly timestep).

The daily grids for these variables were downloaded as ascii files, one for each day, and so the same process was used to convert the ascii files to time series data for each grid square. Maximum temperature in the UKCP09 dataset has been corrected for the effects of latitude/ longitude, altitude and coastal effects, with minimum temperature corrected for urban effects in addition to these (Perry et al., 2009). The monthly data for the other variables were obtained as large spreadsheets containing time series for each 5km grid square for a 20 year period. The monthly data were converted to daily time series files for each grid square using the simple approach of applying the average daily value provided for each month as the actual daily value. For the years in which monthly data were unavailable, the average value for that month was calculated from the rest of the data set and applied to fill in the gaps. The resulting data is therefore based on some approximations and could be sensitive to this however, no better data/methods were available.

The daily data series were then used to calculate daily PET using the Penman-Monteith equation:

$$\gamma ET = \frac{\Delta(R_n - G) + \rho_a c_p \frac{(e_s - e_a)}{r_a}}{\Delta + \gamma(1 + \frac{r_s}{r_a})}$$

where:

$ET_0$ is the reference evapotranspiration (mm/day),

$R_n$ is net radiation at the crop surface $[MJm^{-2}day^{-1}]$,

G soil heat flux density $MJm^{-2}day^{-1}]$,

T mean daily air temperature at 2 m height $[^oC]$,

$u_2$ wind speed at 2 m height [m $s^{-1}$],

$e_s$ saturation vapour pressure [kPa],

$e_a$ actual vapour pressure [kPa],

$e_s - e_a$ saturation vapour pressure deficit [kPa],

$\Delta$ slope vapour pressure curve $[kPa^oC^{-1}]$,

$\gamma$ psychrometric constant $[kPa^oC^{-1}]$. (Allen et al., 1998)

The resulting spatial patterns of PET are shown in Figure 3.9, which shows average annual PET derived from the UKCP09 dataset using the method described above for 1961-1990. This figure shows that the south-east of England experiences much higher rates of PET (up to 700mm in London) than the north (less than 400mm). PET rates are also higher in urban areas as indicated by the dark red spots on the map due to the urban heat island effect (Howard, 1818). The spatial variation in PET is closely related to differences in temperature regimes between regions in Great Britain.

It should also be noted that the daily maximum and minimum temperature data are used to provide input to the snowmelt module of SHETRAN. There are options for both temperature-index and energy balance modelling of snowpacks within SHETRAN, but the former, simpler method is used here on account of the lower input data requirements. Given the climatology of the UK, this modelling decision is likely to have most bearing on upland and mountainous regions, for which estimation of all the inputs required for energy balance is particularly complicated due to topographic complexity. The commonly used temperature-index method is therefore considered to be a more appropriate starting point for the modelling system.

### 3.3.6 Lakes

SHETRAN was modified to allow for input of a map showing the position of lakes in a catchment, in order to improve their representation in catchment simulations. The details of the method used for simulating the effects of lakes in SHETRAN is discussed briefly below and also in Chapter 5. The data layer used as input is the OS Meridian 2 lakes layer (Ordnance Survey, 2013b). This dataset is available as a vector file, which was converted to a 1km raster.

### 3.3.7 National River Flow Archive

Each SHETRAN simulation requires a mask to delineate the catchment boundary , in order to specify which grid squares should be taken into consideration in calculations. Catchment boundary shapefiles were downloaded for all 1537 catchments described in the National River Flow Archive (NRFA) (Morris et al., 1990) and converted to individual ascii files for each catchment. It is worth noting that these boundaries

are all for the surface water catchment to gauging stations, which does not take into account any differences in underlying groundwater catchments areas where aquifers are present or show different flow divides. This consideration only applies to some of the catchments simulated and is discussed further in Chapters 5 and 7.

For evaluation of the modelling system, all freely available historic daily flow data was downloaded from the NRFA and stored as time series files for each gauging station (National River Flow Archive, 2014a). The data had been largely quality controlled by the NRFA however some small formatting errors were identified and removed.

## 3.4 Modifications to SHETRAN

During the process of building and testing the system presented here, SHETRAN has been modified in several ways in order to improve its performance (changes to the SHETRAN code itself were made by S. Birkinshaw). These changes and their effects are discussed more fully in Chapter 5, but an overview of the fundamentals is presented here for information. The modifications include the following:

- SHETRAN now accommodates spatially varying rainfall and potential evapotranspiration data, as opposed to only point rainfall data as used in most prior work. This allows for use of recently developed gridded products and provides more realistic representation of important variability within catchments
- There is now also a better process within SHETRAN for removing sinks in the DEM - i.e. grid squares at a lower elevation than all neighbouring grid squares - to prevent unrealistic levels of ponding and surface storage, which would act to reduce flows in an unrealistic way
- A minimum DEM (describing the minimum elevation in a grid square) is now used in combination with an average DEM (describing the average elevation in a grid square) to more accurately route the river links calculated in SHETRAN
- SHETRAN has been modified to accept a map of lake locations so that they can be accurately represented within catchments. If a lake grid cell intersects a river link, it is treated as an open water body by reducing the Strickler coefficient controlling surface roughness from 20 to 3. This acts to effectively slow flow and increase storage of water in the channel
- Changes have also been made so that it is possible to assign Strickler coefficients as a function of land cover, rather than applying one parameter value for the whole

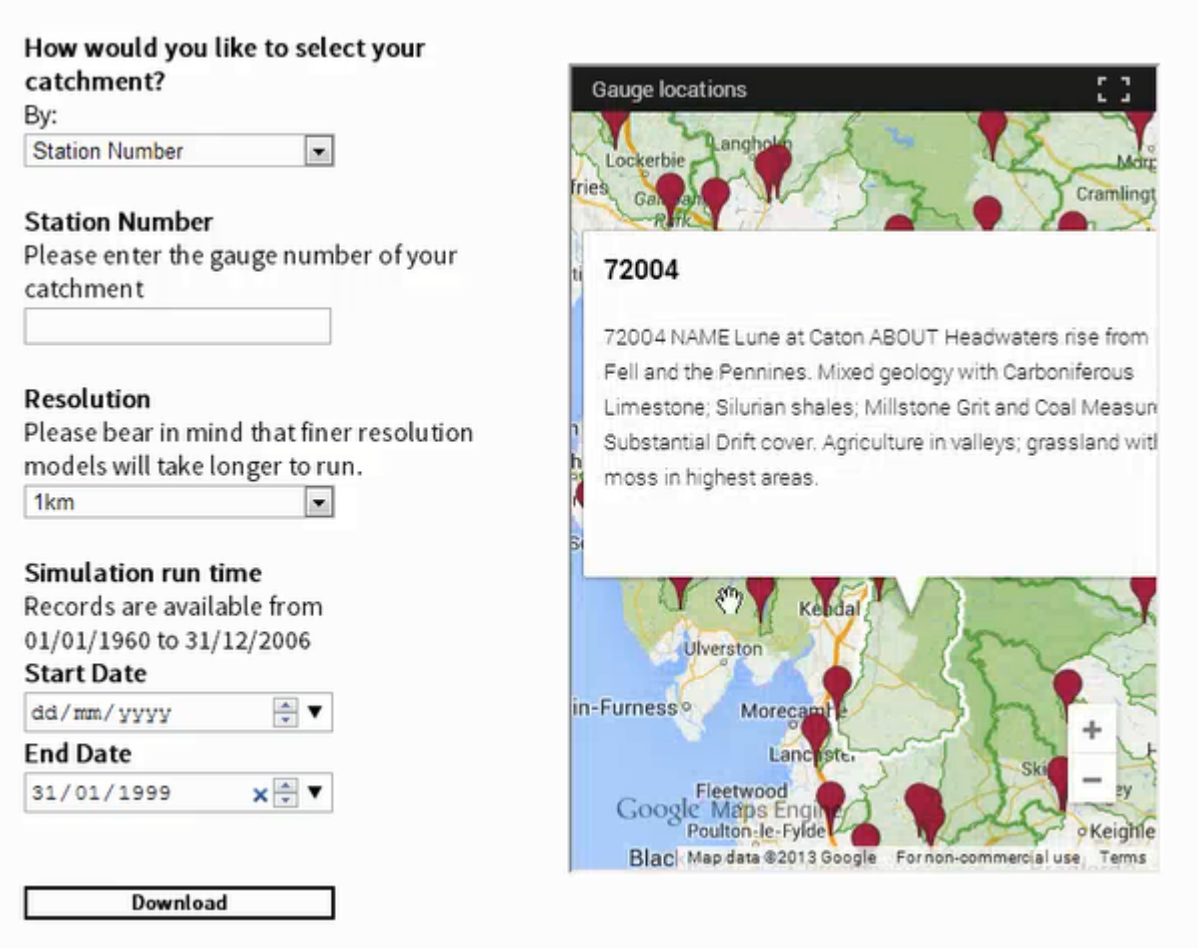



*Figure 3.10: The user interface of SHETRAN for GB.*

catchment, as has generally been the case in the past. This allows roughness to vary with vegetation, which seems a conceptually reasonable option. (This is not explored in this thesis).

## 3.5 Automatic set up of Shetran

In order to develop a national modelling system based on SHETRAN, a large array of data for the whole of Great Britain and the period 1960-2006 described above has been integrated into a framework that features a new, user-friendly graphical interface, which extracts and prepares the data required for a SHETRAN simulation of any catchment in Great Britain. This has vastly reduced the time it takes to set up and run a model from months to seconds. The system represents very substantial progress in the ability to deploy SHETRAN for individual or very large numbers of catchments, which is a significant contribution of this thesis to ongoing developments in physically-based

modelling.

The underlying Python scripts of the system take input from a user interface (scripted in HTML and Javascript) to automatically set up a model for a catchment (see Figure 3.10). For gauged catchments, an existing catchment boundary can be selected from a map, while for ungauged catchments a shapefile of the catchment boundary can be uploaded to extract an appropriate model. The process of setting up a catchment model using this interface takes only a few seconds as opposed to the several weeks it can take using a manual approach (Birkinshaw, 2010).

The algorithm underpinning catchment setup goes through the following steps:

- Takes a boundary shapefile or gauge number as input to delineate the catchment
- Takes the start and end date of the simulation as input
- Creates a project directory
- Creates or selects the mask
- The bottom left coordinates and the number of rows and columns of the mask file are used to seek and copy the corresponding area of the larger DEM, minimum DEM, soil, land cover and PET and rain map. These selections are saved as ascii maps in the project file
- The PET and rainfall map for the catchment is read and the grid IDs covering the catchment are identified. The appropriate time series data (using the start and end date) are then selected and written to a .csv file for use in SHETRAN (one each for PE, rainfall, maximum temperature and minimum temperature)
- The flow data for this time period are also retrieved using the gauge number. Any missing data are written to file as a blank, for example in an ungauged catchment
- A library file is generated for this catchment. The library file provides the SHETRAN data pre-processor with information on file locations, parameter values and timestep information
- The SHETRAN data pre-processor is run using the library file as input, generating the input files for SHETRAN.

## 3.6 Condor

Conducting multiple sensitivity tests and scenario runs for hundreds of models requires computational resources for conducting large numbers of simulations, ideally in paral-

lel. Depending on the size of the catchment, running SHETRAN simulations for 12 years at a 1km resolution can take anywhere between a few minutes to several hours, with a typical run for a single catchment taking between 30 and 60 minutes. Run times of this order are of course not insignificant when large numbers of runs for each catchment are desired. In order to efficiently conduct large numbers of runs for the several hundred catchments simulated in this work, the runs were distributed using Newcastle University's Condor Network. The University's Condor Network is based on the original Condor Network described in Litzkow et al. (1988), and allows applications to run on idle computers around the university.

To run multiple models using the Condor Network, the input files for each catchment model must be zipped up and sent to a folder on the main Condor server, along with a batch file to control execution and a shell script. A zipped copy of SHETRAN, its .dll library file and 7z.exe are also submitted. Experience has shown that Condor can be unreliable, so jobs (i.e. catchment simulations) are submitted in batches of 10 with a 5 minute wait in between batches, which tends to give a much higher success rate than attempting to submit too many jobs at once. Each job is then sent from the main Condor server to an idle computer located within the university. Once the job has completed, the results are sent back to the Condor server and are automatically retrieved onto a local machine to avoid using up storage on the Condor server, which can result in crashes. This process can handle 1-2 sensitivity tests overnight.

## 3.7 Conclusion

This chapter has detailed the datasets, processing and software development involved in setting up the SHETRAN for GB national modelling system. Freely available data - with the exception of rainfall, at this stage at least - were collated and processed into 1km rasters aligned with the British National Grid. The datasets together describe the country's physical characteristics. The north and west of GB are particularly steep, wet, cool and impermeable whilst the south east of England is typically flat, dry, hot and underlain by chalk aquifers. Modifications were made to SHETRAN itself (by S. Birkinshaw) to improve the realism of input data, such as permitting distributed precipitation inputs, as well as representation of some hydrological processes. For example, channel delineation and therefore flow routing was improved by using a minimum DEM, with modifications also made to handle DEM sinks and improve the treatment of lakes in SHETRAN. The snowmelt module was activated to simulate snowpacks. A graphi-

cal user-interface was developed so that a non-expert can rapidly set up a SHETRAN catchment model. This user-interface is designed to work in a browser, as it is ultimately intended to be used on a university server so that other modellers can have access to the system.

# Chapter 4

# Creating an hourly rainfall dataset

## 4.1 The need for hourly rainfall data

Many hydrological applications require high temporal resolution meteorological data. One important example of this is flood risk management, with flooding in the UK highly dependent on sub-daily rainfall intensities amongst other factors. Knowledge of sub-daily rainfall intensities is therefore critical to designing hydraulic structures or flood defences to appropriate levels of service. Sub-daily rainfall rates are also essential inputs for flood forecasting, allowing for estimates of peak flows and stage for flood warning and response.

As well as its importance in water management applications, sub-daily rainfall patterns are also of significance for the accuracy and physical realism of SHETRAN simulations. Zhang (2012) show that SHETRAN is more sensitive to temporal than spatial resolution, and that moving from daily to hourly rainfall data can substantially improve model performance. Hourly intervals are indeed used as the basic calculation time step within SHETRAN, but the default method is to distribute daily rainfall inputs uniformly over each hour in any given day. This is clearly unrepresentative of observed rainfall patterns, and often leads to underestimation of rainfall intensity at sub-daily timesteps. It also has the potential to lead to a slight overestimation of actual evapotranspiration in the model, as the higher more realistic rainfall intensities are omitted. In order to provide more realistic hourly rainfall inputs for SHETRAN and maximise use of its physical basis, the UKCP09 5km daily gridded rainfall dataset was disaggregated to an hourly interval using available rain gauge records. The effects of using this dataset for model input are then examined in Chapter 5 to see if more realistic simulations are attained, although the evaluation is limited to daily flows on the basis

of available data. This should still provide a basis for seeing if the timing of peak flows and the water balance improve with more accurate sub-daily rainfall inputs.

In addition, an hourly gridded rainfall dataset has significant potential for practical applications (Blenkinsop et al., 2016):

- Better representation of extremes and pluvial flash flooding
- Validation of high resolution climate models such as the model used in the CONVEX project (Kendon et al., 2012)
- Improving the representation of sub-daily rainfall in weather generators (Jones et al., 2009).

## 4.2 Hourly rainfall data

Hourly rainfall data for gauges with records spanning at least 10 years and held in the UK Meteorological Office (Met Office) Integrated Data Archive System (Met Office Integrated Data Archive System, 2012) were downloaded from the British Atmospheric Data Centre (BADC). These data include land surface observations from the UK Met Office meteorological station network. Additionally, data from 11 gauges in Scotland were collected from the Scottish Environment Protection Agency (SEPA). Tipping bucket rain gauge (TBR) data from 1,300 gauges across England and Wales was also obtained from the UK Environment Agency (EA) (see Figure 4.1). The data held by each organisation had been partially quality controlled, but Blenkinsop et al. (2016) undertook further quality control procedures, particularly by checking the TBR data against their associated check gauges. The following checks were used:

- Accumulated totals were removed, for example totals occurring at regular daily or monthly intervals
- Spuriously high values from high frequency tipping were removed
- Long dry periods (>1 month) were identified and removed if a similar dry period was not observed in a nearby station
- Negative values - mainly occurring as a consequence of erroneous coding of missing data - were removed
- Potential daily accumulated values were noted to be most likely to occur at 9am and 12pm and were identified and removed
- Identification and removal of duplicate rainfall totals in consecutive hourly periods.

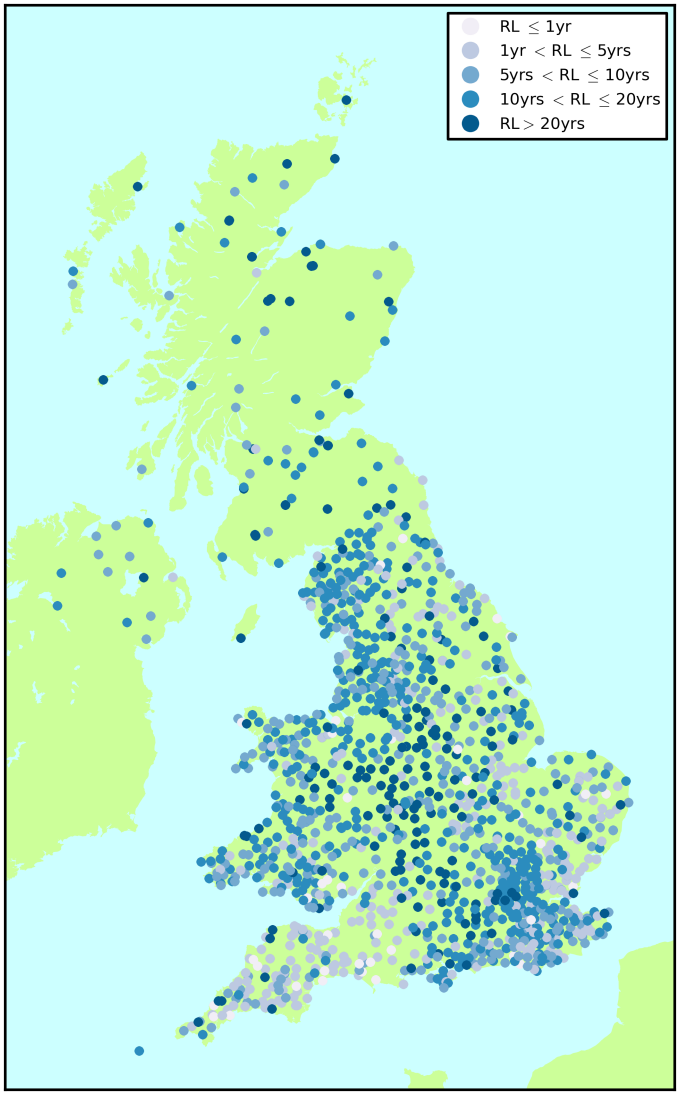



*Figure 4.1: Hourly rain gauge locations coloured by record length. Network density is poor in Scotland but the records there are long. In contrast, the South East of England has good network density but consists of mainly very short records.*

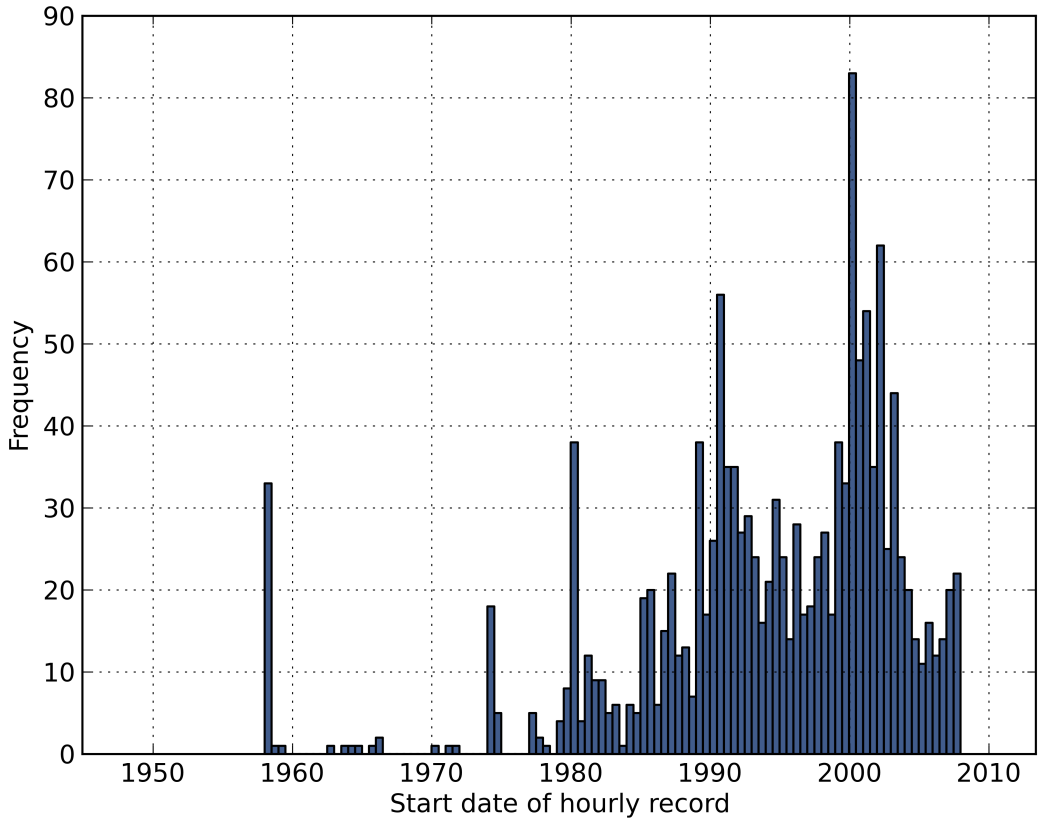



*Figure 4.2: A histogram showing the start dates of the hourly rainfall records. The number of hourly records coming on-line begins to increase in the 1980s and peaks in 2000.*

Figure 4.1 shows the locations of all of the rain gauges used for disaggregation. This illustrates that coverage in terms of gauge density and record length is good for Wales, the Midlands and northern England. However, hourly rainfall records in Scotland are sparse, while hourly rain gauges in the south-west of England tend to have generally short record lengths. None of the records obtained were 100% complete; all had some data missing between the record start and end dates. Figure 4.2 shows that most gauges begin recording in the 1990s. These characteristics of the dataset have implications for both the period of disaggregation feasible using available hourly data, as well as uncertainties likely to arise from station density and record completeness or quality. This may be a particularly important consideration for regions of Scotland, where hourly records are very limited but mountainous topography could result in complex variation in spatial and temporal rainfall patterns.

## 4.3 Correlation to the daily record

As an additional quality control measure, the hourly rainfall records were aggregated to daily intervals and compared to the UKCP09 5km gridded daily rainfall dataset. This dataset was created from the Met Office archive of daily observations and covers the UK at a 5km x 5km resolution over the period 1958 to 2006. The dataset has already been quality controlled (Perry and Hollis, 2005) and used in research (Rahiz and New, 2013b; Simpson and Jones, 2014), so it is considered to form an appropriate reference against which to evaluate the likely accuracy of the hourly measurements.

The procedure for this comparison was to aggregate each hourly rainfall record to provide a daily total for the period 0900 to 0900 the next day, following the conventions for rainfall measurement in the UK. Any days for which daily totals could not be calculated from the hourly data were omitted from the analysis. Each of the aggregated time series were then compared to those of the UKCP09 5km grid square in which the gauges are located. This comparison was initially conducted using Pearson's correlation coefficient. This statistic can exhibit bias when applied to correlation of rainfall time series, as a result of the high number of zeros often found in rainfall series, which leads to a highly skewed dataset (Ha and Yoo, 2007; Serinaldi, 2008). As investigated by Habib et al. (2001), the Pearson correlation coefficient is therefore not an informative measure of correlation between rainfall datasets, but in this work it was certainly found to be useful for highlighting errors in the hourly dataset. Every station with a correlation coefficient value of less than 0.9 turned out to have a very obvious error when inspected visually.

Three main types of error were found using this method:

- Spurious rainfall rates up to 100 times greater than any other value in the record (see Figure 4.4)
- An excess of zeros at the beginning of a record
- A change point in the record where the values increase by a factor of 5, potentially associated with a change in rain gauge (see Figure 4.5)

In order to account for the high frequency of zeros in the distributions of daily rainfall, binary matching statistics were also applied to compare the gauge data with the UKCP09 dataset. This follows an established method demonstrated in Yoo and Ha (2007) for example. The matching statistics are calculated for each gauge location by finding: the percentage of days on which it rains in both the gauge and UKCP09

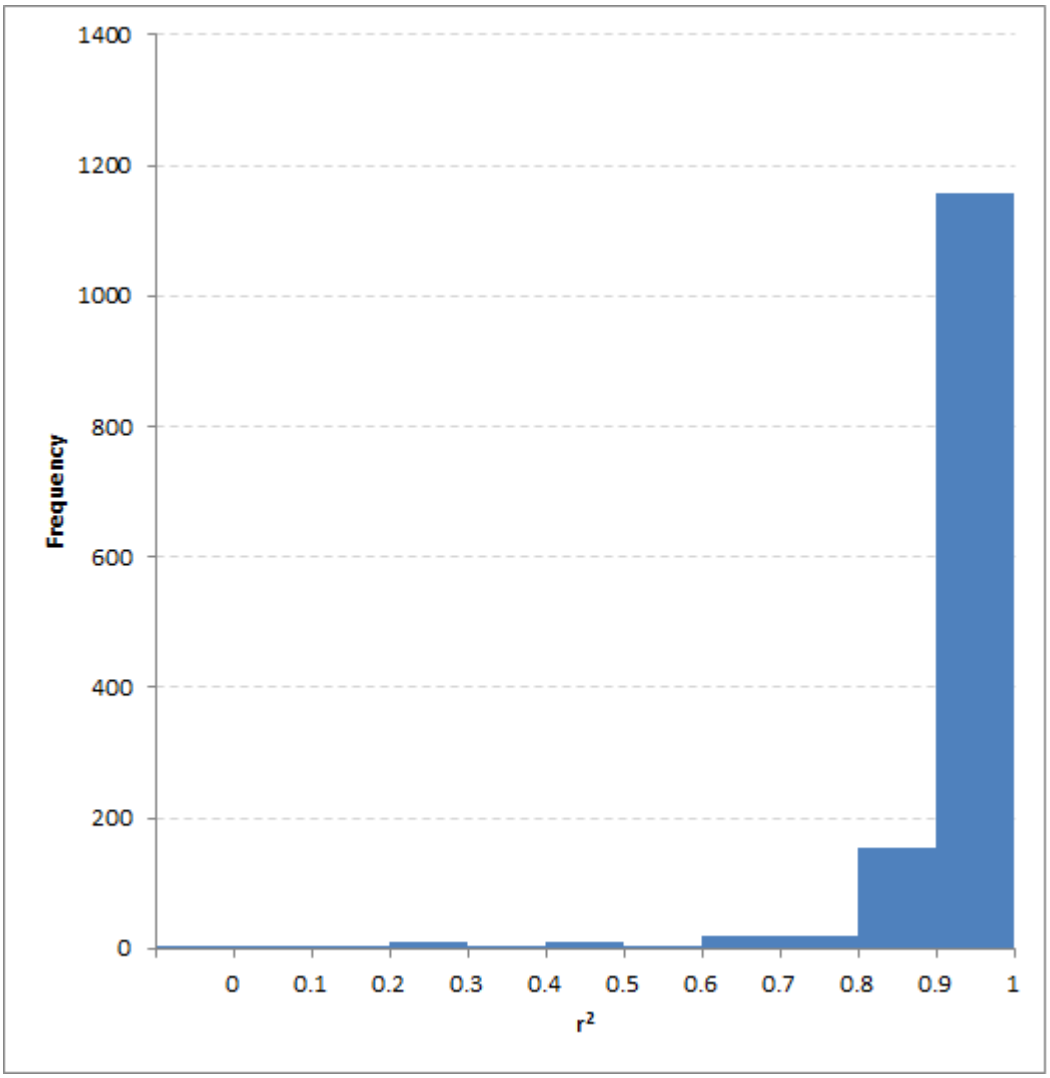



*Figure 4.3: Histogram showing the frequency distribution of $R^2$ calculated for each gauge and the daily record corresponding grid square. The majority of records show a correlation of 0.8 or higher.*

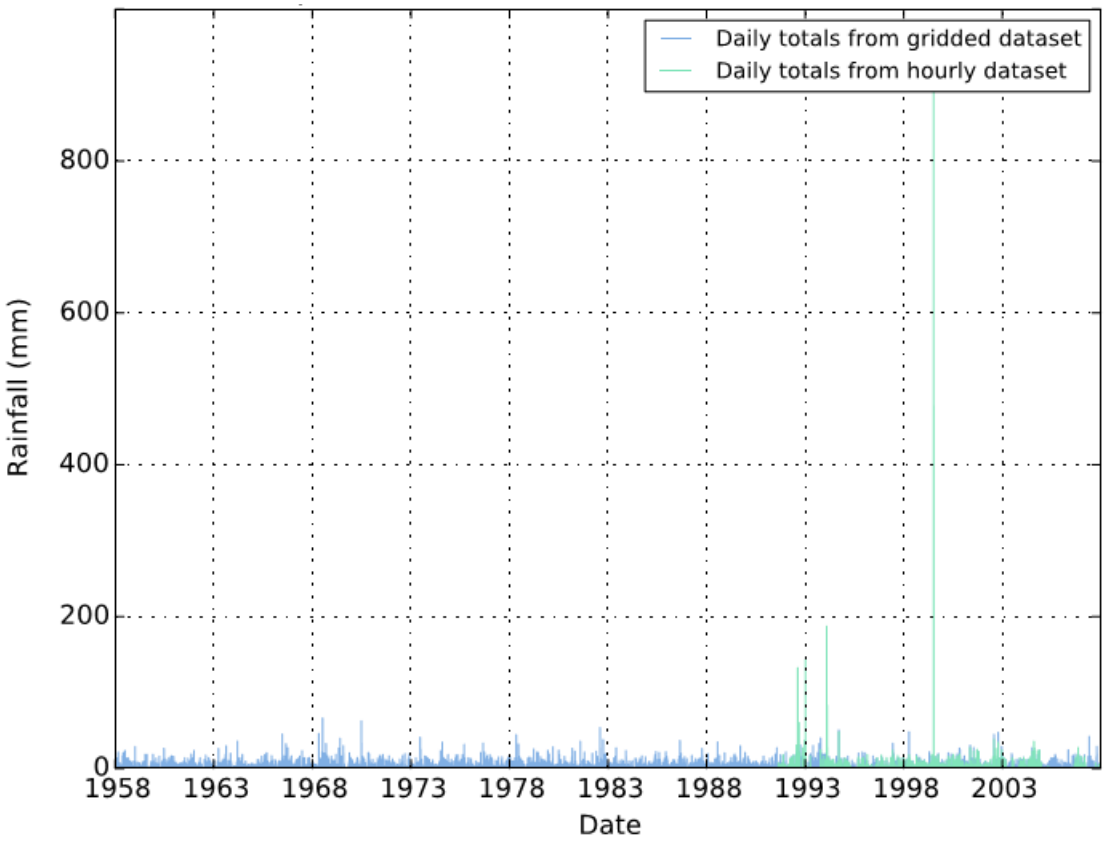



*Figure 4.4: Time series from the gauge in March showing a spuriously high result.*

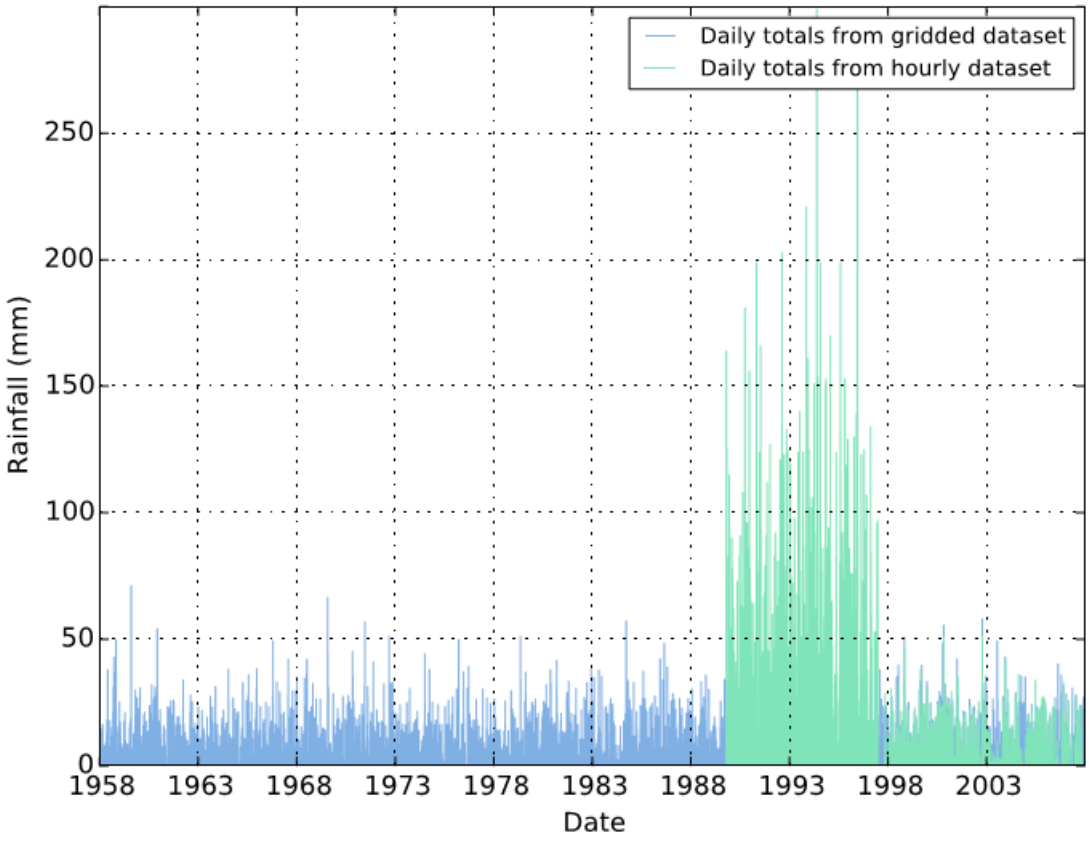



*Figure 4.5: Time series from the gauge at Cowbridge showing a change in the measurements by a factor of 5, likely to be caused by a change in rain gauge.*

records (P11); the percentage of days in which it rains in the UKCP09 daily series but not the aggregated hourly record (P10); the percentage of days that it rains in the aggregated hourly record but not the UKCP09 daily record (P01); and the percentage of days where it does not rain in either data series (P00). P00 + P11 therefore shows how concordant the two records are. Using this approach shows that on average the records were concordant (P00+P11) 82% of the time. Removing the records subject to the errors identified above increases the degree of consistency between the aggregated hourly data and the UKCP09 series to 91% on average. P00 was calculated as 0.38, P10 as 0.07, P01 as 0.03 and P11 as 0.53.

Overall the analysis suggests that almost all gauge locations show an excellent correlation between the hourly rainfall record and daily rainfall taken from the UKCP09 gridded dataset. This finding was confirmed by visual inspection of time series to further validate the good correlations and identify unexpected anomalies. Stations with erroneous values were then removed in preparation for using the hourly data in disaggregation.

### 4.3.1 Extreme values

Once the erroneous values were removed, comparison of the extreme values in the aggregated hourly and UKCP09 daily datasets was undertaken. For each gauge, the highest annual value was found in the aggregated hourly record and compared to the highest value in the surrounding 3-day period in the UKCP09 daily gridded dataset. The 3-day window was used in order to allow for any minor timing error. The two series were compared by calculating $R^2$ for each gauge and corresponding grid cell. This revealed good correlation between the two data series. This supports the conclusion that the hourly data shows a high degree of consistency with the UKCP09 dataset, which supports the approach of disaggregating the latter using the former. Spatial interpolation of rainfall data typically results in the underestimation of extreme heavy rainfall events and overestimates the frequency and intensity of very light rainfall events. It is therefore expected that the correlation of annual maxima will be reduced.

## 4.4 Disaggregation

Multiple methods for temporally disaggregating rainfall exist that rely upon statistical relationships between rainfall intensity, frequency and duration (see Pui et al. (2012),

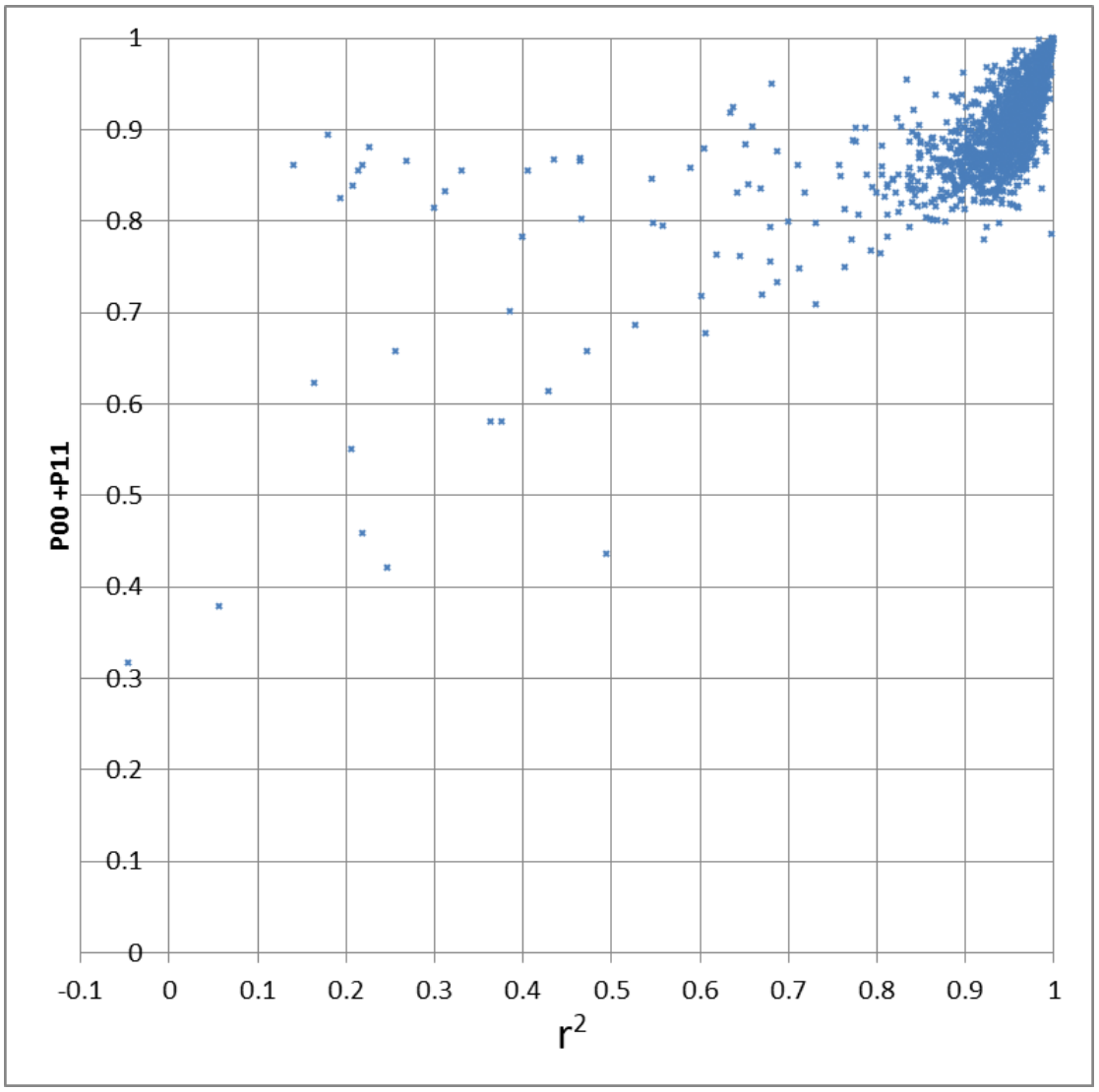



*Figure 4.6: Graph showing Pearson $R^2$ and matching statistics for each gauge. Points with low $R^2$ and low P00+P11 are records that start with an excess of zero values. Points with low $R^2$ but high P00+P11 are records that contain a spuriously high value.*

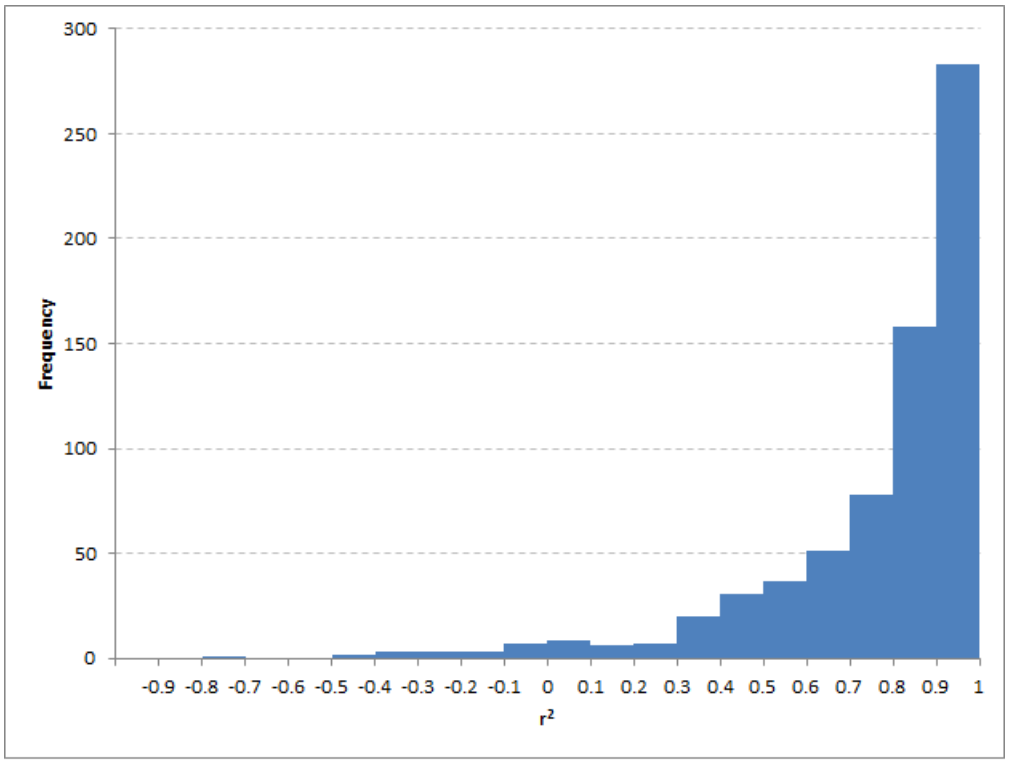



*Figure 4.7: Histogram showing the frequency distribution of $R^2$ calculated for the annual maximum rainfall for each gauge and the daily record corresponding grid square before removing the spurious records. Correlation is still generally high but lower correlations are more frequent due to spatial averaging in the daily record.*

Debele et al. (2007) and Koutsoyiannis (2003)). However, due to the availability of hourly rainfall timeseries that appear to be suitable for use with the UKCP09 dataset, simple nearest neighbour disaggregation using real data was possible in this case. This method is predicated on examining the observed hourly rainfall record for the gauge closest to a given cell (on a given day) and applying the shape of the hourly record to distribute the daily total to hourly intervals. Additional rules were defined to account for missing data. The process is therefore based on an assumption that the nearest gauge to a given cell is sufficiently well correlated to the cell at hourly intervals that the sub-daily distribution can be applied, with no other corrections such as interpolation of hourly gauge records required. This is likely to be a reasonable approach for those regions with good gauge density, but uncertainties are expected to be higher in Scotland and other areas with limited records. In such areas, hourly records are so limited that it is considered unlikely that more complex methods would yield more accurate disaggregation, although this is certainly something that would benefit from further research.

The following algorithm describes the procedure used to disaggregate the daily rainfall time series for each 5km grid square:

- Calculate the coordinates of the grid square
- Create a list of hourly rainfall gauges, ordered by distance from the grid square
- Create an empty hourly time series
- Fill this with data from the nearest hourly gauge
- If the nearest hourly gauge did not contain all of the information, fill any gaps with values from the next nearest gauge
- Continue this process until the new hourly time series is full of values or until there are no more data to fill the time series with
- For each day (0900-0900), distribute the daily total by giving it the shape of the hourly record whilst preserving the daily total. If there are missing hourly values in that day, exponentially distribute the values to hourly intervals.
- The results are recorded in a text file as a time series. As well as the disaggregated value for each hour, the following are also recorded:
  - The original hourly value
  - The exponentially distributed daily value for that hour
  - The gauge that the hourly value came from

* The preference value (whether it was the first choice gauge, second choice etc.)
* The distance of this gauge from the centroid of the grid square
* Whether statistical disaggregation was used for this day.

### 4.4.1 Distance limits

The relationship between rain gauge records as a function of the distance between gauges was investigated, with a view to ensuring that appropriate spatial search limits were defined for finding gauges from which sub-daily distributions could be transferred. This is based on the idea that correlation between rainfall is likely to decrease with distance between gauges, eventually to the point where a gauge may be of little help in disaggregating rainfall for a given grid cell. For example, data from a gauge in Scotland will be of little use when disaggregating daily rainfall in Cornwall in the south-west of England. Introducing a distance limit also has the practical benefit of reducing processing time for creating the hourly gridded dataset, as the number of gauges to be examined for disaggregating each cell can be constrained.

To determine appropriate distance limits, time series from 50 randomly selected hourly gauges were each compared against another time series for another 50 gauges, each of which is an increasing distance from the original gauge. The correlation coefficients were found for each pair, which can be seen in Figure 4.8). This shows that the correlation between hourly gauge records measured by $R^2$ tends to decrease exponentially with increasing distance from the focal gauge. On average, $R^2$ fell below 0.5 with a separation distance of 27km, but a distance limit of 50km was considered to be a useful pragmatic choice for application in the disaggregation process to provide reasonable coverage across the country. However, as spatial correlation of hourly rainfall depends on factors other than just distance and has the potential to vary in both space and time, some additional investigation was conducted with respect to patterns of regional and seasonal variation in the relationships between hourly gauge records. This is discussed in the next section.

### 4.4.2 Regional and seasonal differences

In order to examine regional dependencies in the spatial correlation between hourly rain gauges, separate tests of correlation between records were carried out for the north-west and south-east of England. These regions were selected as examples on the basis of known contrasts in their precipitation climatologies, with the north-west of

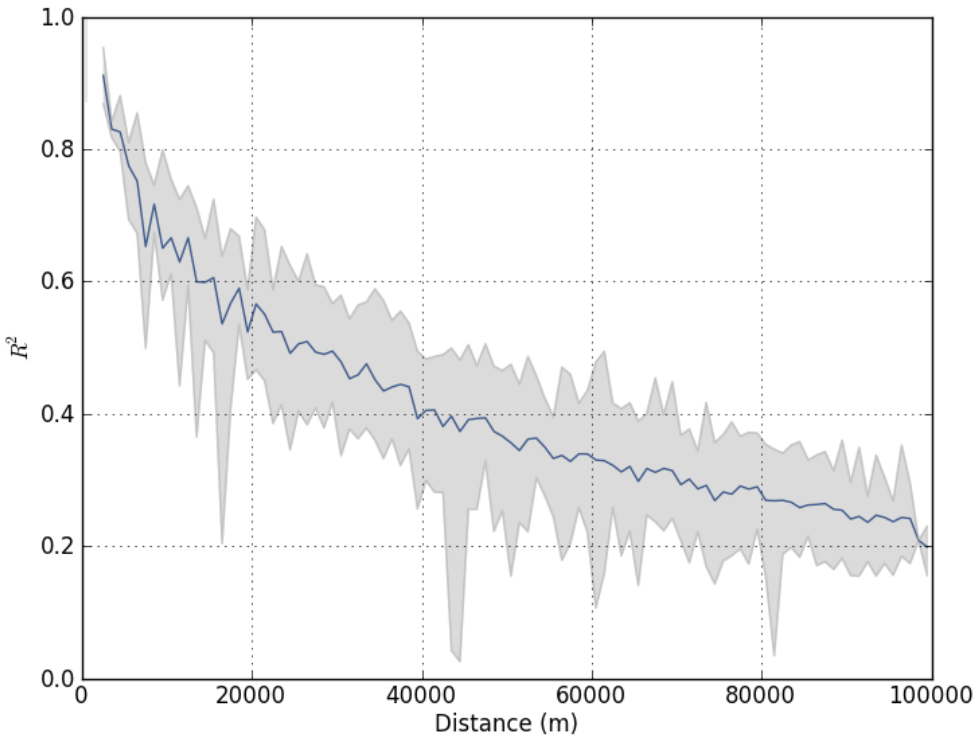



*Figure 4.8: Graph showing how correlation between two gauges decreases with distance. The graph summarises 50 decay relationships where the blue line is the average correlation per km and the grey area shows the variation of all of the runs.*

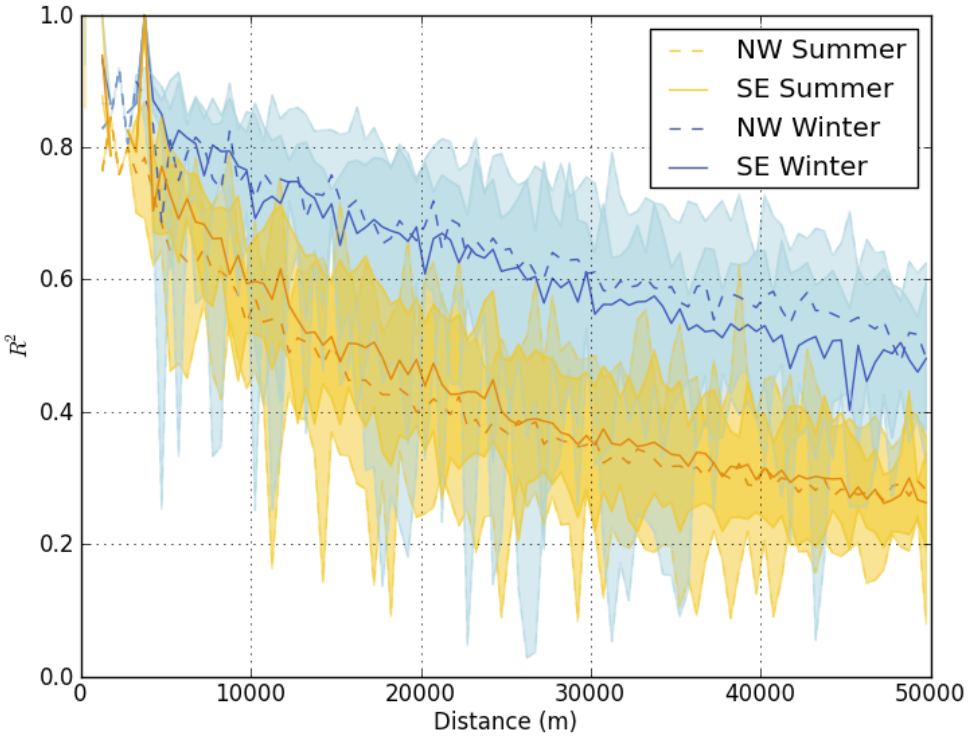



*Figure 4.9: Graph showing how correlation between two gauges decreases with distance and how this varies between region and season. Correlation decays with distance faster in summer (yellow) than in winter (blue) as summer storms are local convective events and winter storms occur as large frontal systems. The differences in correlation decay between regions (solid vs dashed lines) is negligible.*

England recieving higher average annual rainfall and more wet days than the south-east of England. The two areas were arbitrarily divided by a line using the equation $y = 0.82x + 307008$ (in British National Grid (BNG) coordinates) and 50 gauges were chosen in each region for comparison using the same process described above. The results of this are shown in Figure 4.9, which illustrates that there is very little difference in the decay of correlation with distance between the two regions.

However, completing separate analyses for each month of the year indicates that correlation between hourly gauges decays faster with distance in summer months (June, July, August) than in winter months (December, January, February). Correlation is acceptable up to around 50km in winter, whereas in summer correlation tends to fall beneath 0.5 at distances of approximately 15km. These differences appear to be physically realistic when seasonal variation in mechanisms underpinning precipitation in the UK are considered. Rainfall during the winter months is often from large frontal systems, from which spatial correlation of rainfall over substantial distances is to be expected. In contrast, summer rainfall is often associated with convective systems that are less extensive spatially, leading to lower correspondence over large distances than may be observed in winter. These seasonal differences in spatial correlation were not included in the disaggregation method at this stage, however. Instead a limit of 50km was chosen to pragmatically balance results from the correlation analysis, gauge density variability and processing time, as well as to avoid introducing additional complexities before a simple approach was tested. Further analysis might be required to incorporate more complicated disaggregation rules accounting for variability in spatial correlation, although it would be interesting to examine the differences resulting from this in future work.

### 4.4.3 Temporal limits

The record lengths and periods of the available hourly rain gauge data were analysed to determine the most appropriate period for disaggregation. Figure 4.10 shows that the length of the hourly records is highly variable, with a mean record length of approximately 11 years. However, there is a trend for more gauges to come online as time progresses, as can be seen in Figure 4.11). On this basis, the time period for disaggregation was selected to be 01/01/1990-31/12/2006, which coincides with the best record period in terms of gauge numbers. This period is also useful for comparing with the baseline simulations and sensitivity tests described in Chapter 5.

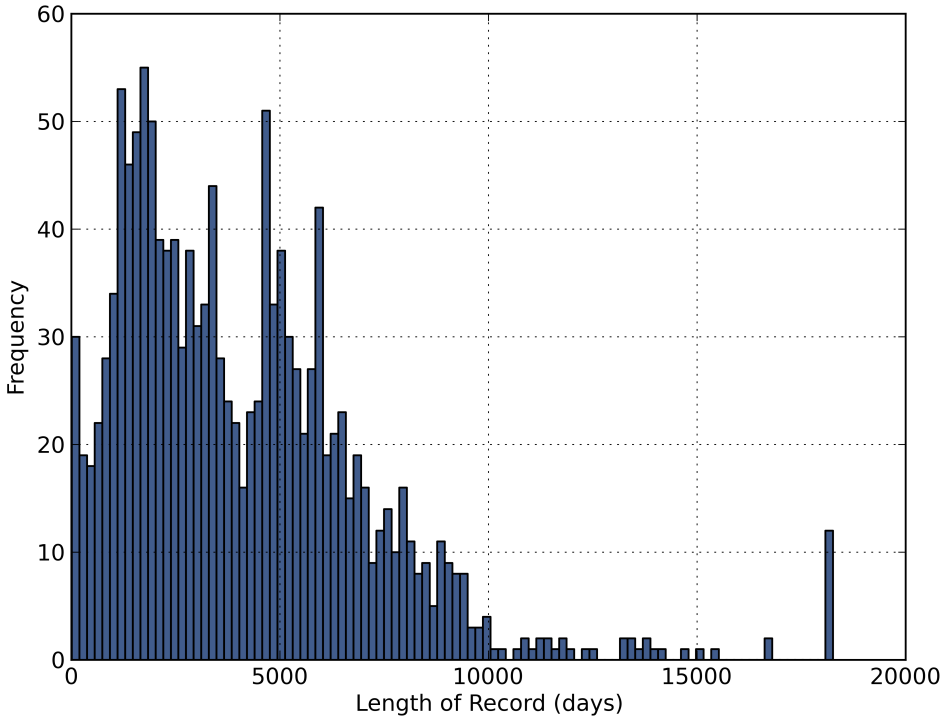



*Figure 4.10: Histogram showing the range of lengths of hourly rainfall records.*

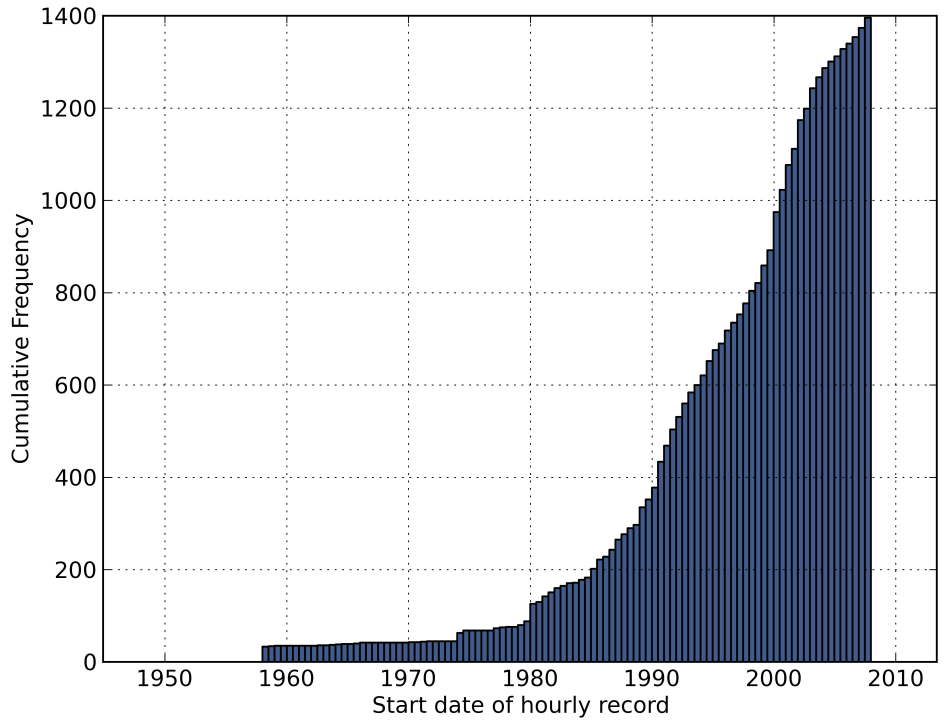



*Figure 4.11: Cumulative frequency histogram of hourly rain gauge record start dates. At least 400 gauges are available at the beginning of the chosen time period of 1990-2006 with more coming online throughout the period.*

### 4.4.4 Intensity-duration relationships

For days when no hourly data are available within the 50km limit for a given cell - even if only one hour is missing or when rainfall occurs in the UKCP09 daily record but not in the hourly record - the daily total is instead disaggregated by statistical methods. The method applied takes the daily total and distributes it exponentially over a number of hours that depends on the rainfall total for the day. The maximum hourly rainfall is arbitrarily taken to be at 10am. The reasoning behind this was to mimic in very general terms the pattern of a storm event, with the exponential distribution defined by:

$$P(X \leq x) = 1 - e^{-\lambda x}$$

where

$$\lambda = \frac{1}{\bar{x}}$$

Therefore for an hours worth of rainfall:

$$rainfallathourh = dailytotal(e^{\frac{-h}{\bar{x}}}) - e^{\frac{-(h+1)}{\bar{x}}}$$

$\bar{x}$ was determined by intensity-duration relationships that were calculated using all of the available hourly data. For each gauge, the daily total rainfall and the number of wet hours were calculated for each day of the record. The results were binned into daily rainfall quantities: less than 1mm, 1-5mm, 5-10mm, 10-25mm and over 25mm. The median durations were found for each bin and used as $\bar{x}$ in the above equation. The results are shown in Figure 4.12.

The intensity-duration relationships were calculated over all available data but are expected to vary seasonally, where more intense, short duration events are expected in summer. The timing of events is also likely to change seasonally, as demonstrated in Blenkinsop et al. (2016), where summer storms were found to typically occur at 4pm. This could be an avenue for further investigation and refinement of the gridded product.

This statistical disaggregation is rarely used because of missing data but is often implemented due to there being no rainfall recorded in the nearest hourly gauge but rainfall occurring in the gridded daily data.

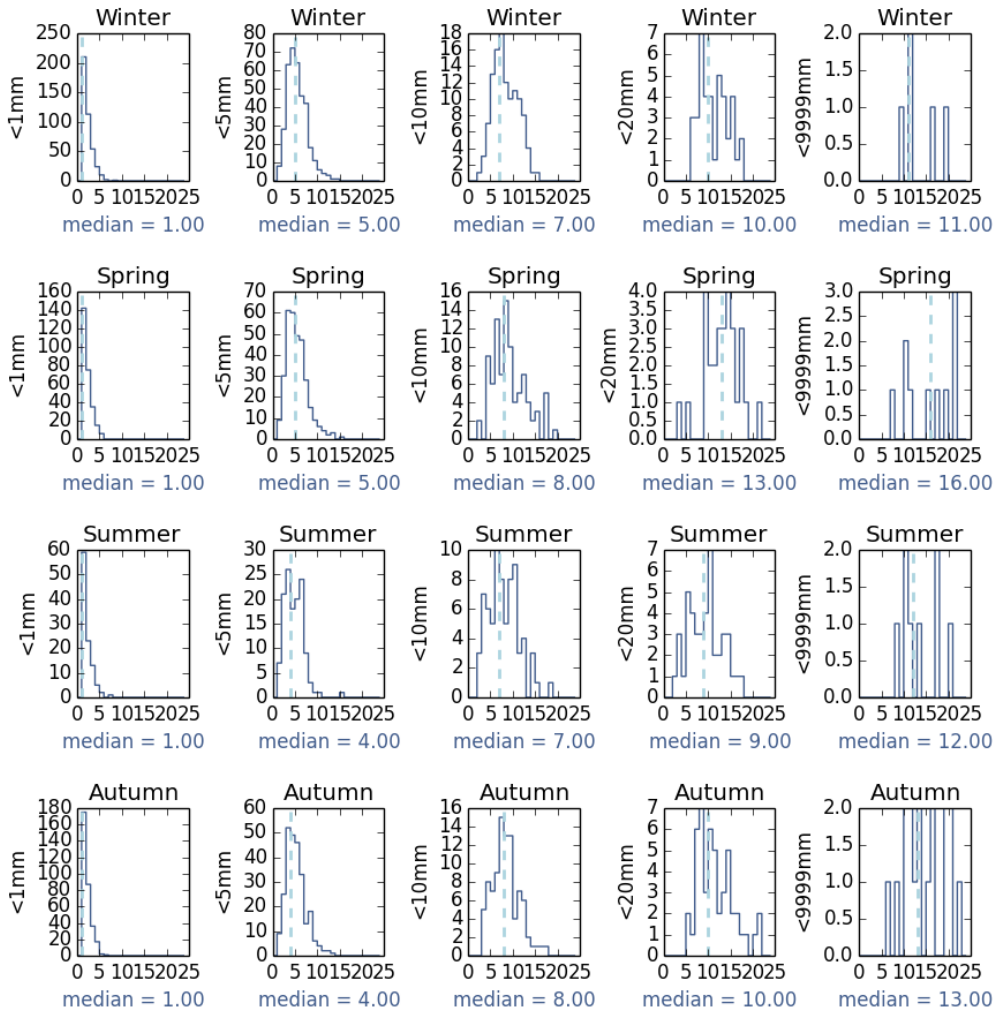



*Figure 4.12: Histograms showing the intensity/duration relationships of hourly rainfall records. Storm duration increases with intensity. Seasonal differences are negligible.*

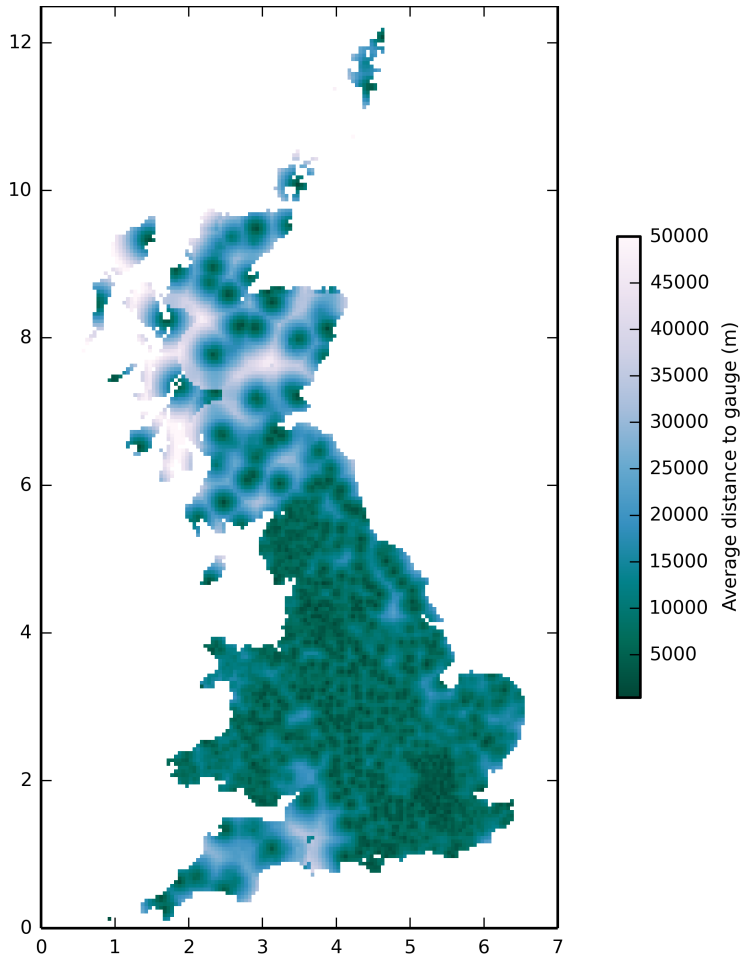



*Figure 4.13: Map Showing the average distance between grid square and hourly gauge used for disaggregation. Most grid squares accessed data from a nearby gauge (<20km), however areas of Scotland and South West England have to look further afield.*

## 4.5 Disaggregation results

The disaggregated hourly gridded rainfall dataset for the period 01/01/1990 to 31/12/2006 took 5 days to process on a desktop PC. The script outputs a time series file for 4313 5km grid squares and also text files containing the time series for all cells in each of the catchments modelled in this work. The time series are ordered in ascending order according to their ID codes for input to SHETRAN.

The quality of this newly disaggregated record was assessed first in terms of the average distance between grid cells and the rain gauges used for disaggregation. Figure 4.13 shows the average distance of rain gauges used in disaggregation from each grid square, which clearly reflects the lower density of rain gauges in Scotland. The small number of dark spots in Scotland in Figure 4.13 correspond to the locations of these gauges. It can

also be seen from this figure that there is generally good hourly record coverage over the north of England, while the south-west of England performs particularly poorly as a result of local gauges having only very short records.

The percentage of the time series for each grid cell for which hourly data were unavailable for disaggregation was also examined. The majority of grid squares had no missing hourly data in their records. This confirms that hourly disaggregation was conducted with observed data rather than a statistical method for almost all hours and grid cells, excluding those areas described above. In conjunction with the average distances of gauges used in disaggregation, this suggests that the gridded hourly rainfall record should be of good quality overall, as a result of the availability of proximate, well-correlated gauge records for the majority of cells. However, limitations in terms of comparatively data-poor regions and simplifications of temporally varying spatial correlations of rainfall should be noted.

## 4.6 Conclusions

This chapter outlines the creation of a national hourly gridded rainfall dataset for the period 01/01/1990 to 31/12/2006 from disaggregation of the UKCP09 5km daily rainfall grids using SEPA, BADC and EA hourly raingauge data. Quality control measures for the sub-daily gauged data were undertaken, both through site-specific analysis of errors indicated by statistical checks and comparison to daily rainfall totals from the UKCP09 grids. The resultant dataset shows good quality overall for England and Wales, where local (<20km) gauges were available for use in disaggregation. More uncertainties are present for Scotland in particular, where only a few hourly gauge records exist. Further evaluation of this dataset through its use as input to the SHETRAN modelling system is described in Chapter 5.

# Chapter 5

# Sensitivity Testing

In chapter 3, the configuration of a national, physically based spatially distributed hydrological model for Great Britain was outlined. This chapter proceeds to assess the performance of this modelling system in terms of its capacity to accurately simulate the behaviour of 306 gauged catchments of varying characteristics distributed across the country. The performance metrics underpinning this analysis are first introduced and then applied to evaluate an initial simulation, which is intended to provide a preliminary benchmark against which subsequent structural changes to SHETRAN made in this work can be compared. Description and evaluation of these structural changes follows, providing details of the modifications considered necessary to improve physical process representation in the model and resolve some initial model stability issues. Thereafter, the results from a series of sensitivity tests are presented, focusing on the implications of global changes to a number of SHETRAN's key parameters. Finally, the relationships between catchment characteristics and model performance are explored through cluster analysis and assessment of the relative performance of nested catchments, which highlights some strengths and weaknesses of both the SHETRAN modelling system applied to catchments across Great Britain and the currently available datasets.

## 5.1 Performance measures

Simultaneously simulating several hundred catchments presents a challenge in terms of how to assess the extent to which the modelling system produces accurate streamflow simulations and whether it does so for the correct reasons, i.e. in a conceptually plausible way, by faithfully emulating physical catchment processes. In studies fo-

cussing on a smaller number of catchments, it is typical to combine both quantitative performance metrics with qualitative judgement and experience in model evaluation, although there are currently no standard procedures for assessing model performance (Hrachowitz et al., 2013). However, as the emphasis of this study is on developing a system for national scale hydrological modelling of large numbers of catchments, intensive assessment of each catchment by the modeller is clearly not feasible. Therefore, in conjunction with qualitative analysis of representative hydrographs, this study has sought to define a range of performance metrics in order to assess goodness-of-fit across the sample of catchments in a systematic way to indicate aggregate performance.

It should be noted that this study focuses on metrics for evaluation of simulated streamflows and catchment water balances with respect to observations, without consideration of other catchment state variables that could be used to diagnose model accuracy, such as groundwater heads. Restricting the analysis to streamflows is a pragmatic decision based on data availability and retaining tractability of the analysis within pragmatic constraints on the project. Clearly assessment of the system with respect to other available data would be a desirable avenue for further study, given the integrated representation of hydrological processes provided by SHETRAN.

An array of metrics have been employed to evaluate hydrological models with respect to observations. One such metric is the Nash-Sutcliffe Efficiency (NSE), which is widely utilised as a general indicator of the degree to which observed and simulated flow time series correspond (Nash and Sutcliffe, 1970). NSE is calculated as:

$$NSE = 1 - \frac{\sum_{i=1}^{n}(O_i - S_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2}$$

with O representing observed flows and S representing simulated flows. Maximum NSE values of 1 indicate a perfect correspondence between simulated and observed flow time series, with lower NSE values reflecting a lesser degree of fit. A NSE of 0 suggests that the model is no better a predictor of observed streamflow than simply taking the mean of the observations, with NSE values less than 0 implying that the model generally provides worse estimates of flow than the mean of the observations.

In this work, NSE values for each simulated catchment were classified into a small number of intervals to facilitate easier interpretation. This follows the approach taken in other studies investigating large numbers of catchments (e.g. Henriksen et al., 2003; Crooks et al., 2014; Habets et al., 2008). NSE values greater than 0.7 were taken to generally indicate good overall simulation of the streamflow hydrograph here, with

values between 0.5 and 0.7 indicating reasonable performance and values below 0.5 considered poor. These intervals were defined to be in line with the interpretations of NSE typically found in the literature and are essentially in line with the classification used by Henriksen et al. (2003) in one of the few other examples of national-scale modelling using a physically based spatially distributed approach.

While NSE is frequently used to assess model performance, the metric is known to have some notable limitations. Firstly, NSE is often considered to be biased towards high flows, where absolute residuals are often largest and the implications of using squared differences between observed and simulated values in the calculation are significant (Beven, 2012). Beven (2012) also highlights the sensitivity of NSE to errors in timing in simulated flows and the potential temporal autocorrelation of residuals, while Gupta et al. (2009) demonstrate how decomposing NSE into constituent components of correlation, bias and variability could indicate that underestimation of variability occurs if the NSE value is maximised. Moreover, Gupta et al. (2009) find that comparison of NSE values between catchments is made more complicated by the way in which bias is scaled according to the standard deviation of the observations. It is therefore suggested that NSE needs to be used in conjunction with other performance metrics to circumvent some of its limitations.

As such, a number of other measure are calculated, as in Crooks et al. (2014), including the bias in the water balance:

$$Bias = 100(\frac{\bar{S}}{\bar{O}} - 1)$$

the average of the monthly bias in the water balance:

$$MM = \frac{1}{12}\sum_{m=1}^{12} 100\frac{|O_m - S_m|}{O_m}$$

the average bias in the flow duration curve at percentiles 1, 10, 25, 50 and 75:

$$FD = \frac{1}{5}\sum_{n\in\{1,10,25,50,75\}} 100\frac{|O_n - S_n|}{O_n}$$

and a measure combining both monthly water balance biases and those in the flow duration curve:

$$mmfd = MM + FD$$

Annual water balance bias (%) is calculated by finding the ratio between observed and simulated total flows for all years of the simulation. This provides an indication of how well the balance between aggregate catchment inputs, outputs and storage

| Band | NSE | Bias (%) | mmfd |
|---|---|---|---|
| 1 | NSE $\geq$ 0.8 | -10 $\leq$ Bias $\leq$ 10 | mmfd $\leq$ 20 |
| 2 | 0.6 $\leq$ NSE $<$ 0.8 | -20 $\leq$ Bias $<$ -10 or 10 $<$ Bias $\leq$ 20 | 20 $<$ mmfd $\leq$ 40 |
| 3 | NSE $<$ 0.6 | Bias $<$ -20 or Bias $>$ 20 | mmfd $>$ 40 |

*Table 5.1: Bands for performance statistics.*

changes (which may be limited over the longer term) match between observations and the simulation. Water balance bias therefore allows for assessment of whether or not the simulated flow volumes are reasonable overall, something which cannot be easily inferred from visual inspection of hydrographs or indeed quantitative characterisation using NSE.

However, one of the limitations of analysing the water balance at only the annual timescale is that seasonal variation is not accounted for and any intra-annual compensating errors remain hidden. As such, monthly water balance bias is also calculated, which will show any deficiencies in performance related to seasonal variation linked to the catchment flow regime. This is further complemented by evaluating bias at selected flow percentiles, which indicates how well the model simulates the overall flow regime, with less emphasis on timing errors than inherent in NSE. Biases in monthly water balances and at selected percentiles across the flow range are then combined into a composite metric, mmfd, which usefully describes overall model performance with respect to both seasonally-varying water balance patterns and the catchment flow regime.

Similarly to Crooks et al. (2014), bands to categorise performance across a large number of catchments have been defined for these metrics (see Table 5.1). The bands provide a useful framework for large scale, overall assessment of model simulations, but of course categorising scores in this way represents a compromise between pragmatism and detail.

It would also be possible to define yet more performance measures to cover other areas of interest, such as peaks over threshold or flood frequency distributions. For this study, however, combining NSE, water balance bias, bias at flow percentiles across the catchment regime and mmfd with visual inspection of as many hydrographs as feasible is considered to provide enough information for evaluation of overall model performance.

## 5.2 Initial Simulation

In order to identify a benchmark level of performance against which to compare the structural improvements made to SHETRAN and sensitivity tests described shortly, an initial run was conducted utilising the model configuration described in chapter 3 in conjunction with best-estimate parameters based on expert knowledge of SHETRAN (Birkinshaw, 2011). The results of this initial simulation are presented below with respect to the performance metrics outlined above. Locations of the specific catchments discussed can be found in figure 5.1.

| Measure | NSE | Bias | MM | FD | mmfd |
|---|---|---|---|---|---|
| Median | 0.67 | -2.2 | 20.9 | 25.2 | 45.2 |
| $25^{th}$ Percentile | 0.35 | -13.3 | 13.5 | 16.7 | 30.5 |
| $75^{th}$ Percentile | 0.77 | 10.4 | 32.8 | 40.2 | 71.6 |
| % Band 1 | 19.4 | 41.1 | NA | NA | 4.4 |
| % Band 2 | 40.8 | 28.2 | NA | NA | 36.7 |
| % Band 3 | 39.8 | 30.7 | NA | NA | 58.9 |

*Table 5.2: Summary table of performance statistics for the initial run.*

The first point to note regarding the initial simulation is that 69% of catchments are associated with NSE greater than 0.5 and just under 50% of catchments attained NSE greater than 0.7. This provides a preliminary indication that the initial model configuration results in an encouraging level of simulation of the dynamics of streamflow hydrographs across the wide range of catchment types in Great Britain. In addition, Table 5.2 shows that the annual water balance bias is within a range of +/- 10% for approximately 50% of the catchments, which represents a fairly high degree of consistency between observed and simulated runoff volumes. However, the degree of spread and fairly wide overall range of WB biases are notable. Further insights into the patterns of WB bias may be drawn from Figure 5.6, which indicates that spring/summer flows may be typically overestimated, whereas October/November/December flows may be underestimated. This pattern of seasonally-varying bias refers to the overall situation across all catchments, roughly commensurate with the median of the monthly/seasonal distribution of WB biases for all catchments. However, the degree of spread in the water balance performance between catchments is large, with the interquartile range spanning positive to negative balances in each month. This indicates a degree of spatial variability in water balance bias, as reflected in Figure 5.4. Particularly noticeable are the large biases associated with geographical areas containing a number of chalk catchments, which are considered in more detail shortly.


*Figure 5.1: Locations of the example catchments discussed in this chapter.*

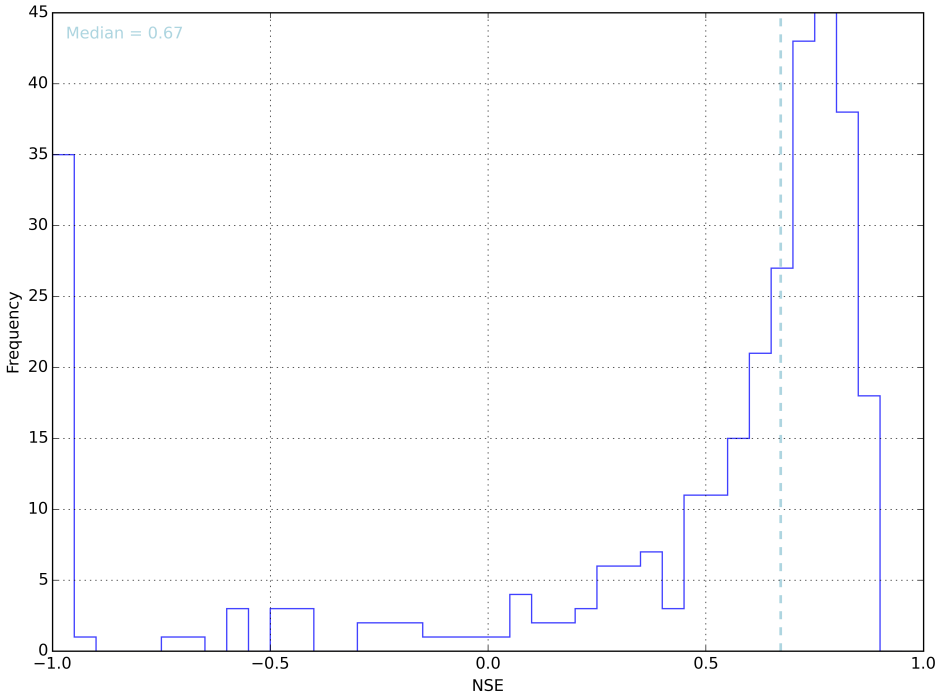



*Figure 5.2: The distribution of NSE for the initial run. Even with no modification, the majority of models perform to an acceptable standard.*

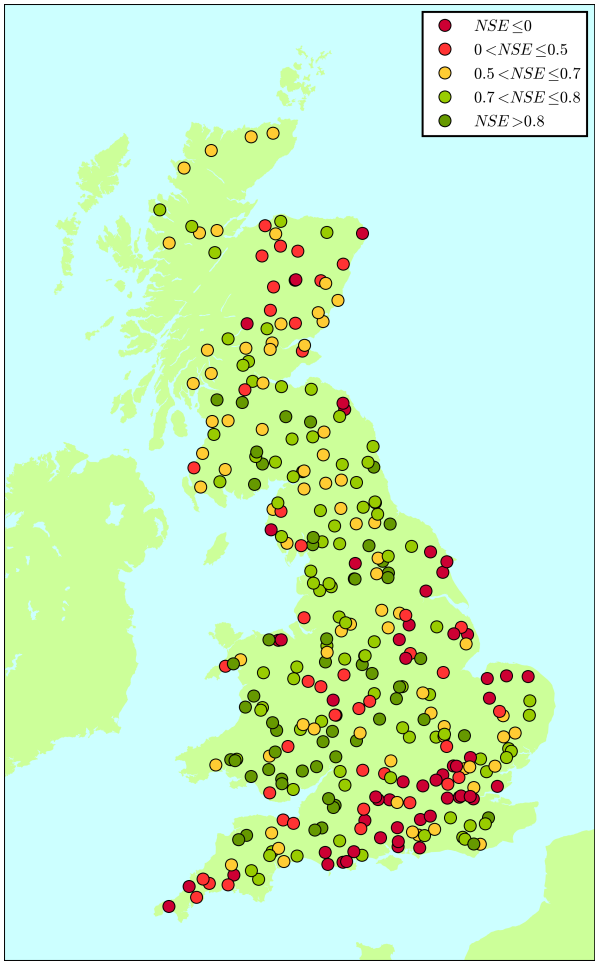



*Figure 5.3: Map showing NSE of initial SHETRAN run. The points represent gauge locations in each catchment. Green indicates good model performance, yellow moderate and red poor. The majority of poorly performing catchments are groundwater catchments (Affected by chalk in the south east) or affected by snowmelt in Scotland.*

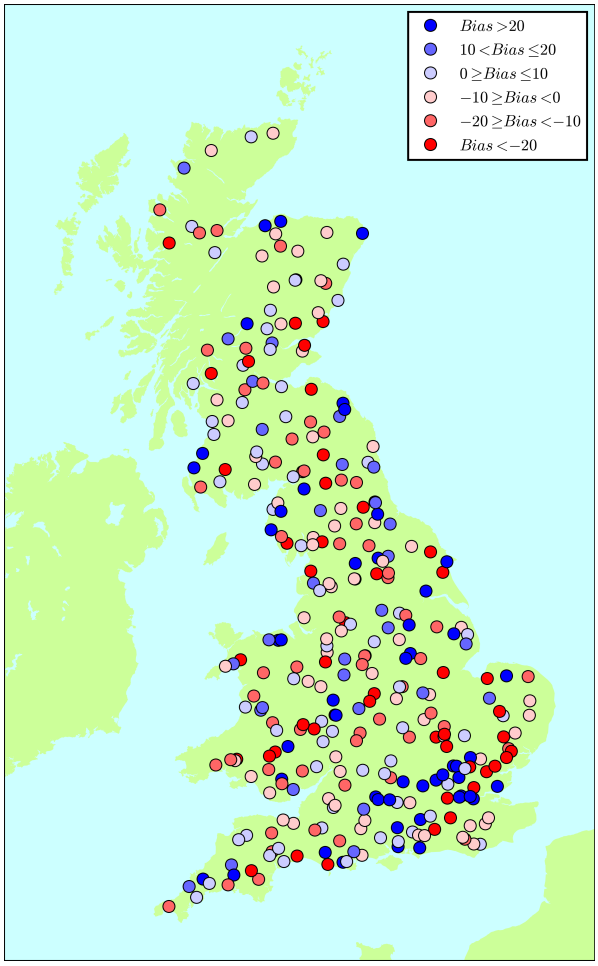



*Figure 5.4: Map showing water balance bias of initial SHETRAN run. The points represent gauge locations in each catchment. A pale colour indicates a small bias (good), darker indicates a larger bias (bad), red indicates that the simulation is under-predicting flow, blue, over-predicting. There is no obvious spatial pattern to the bias.*

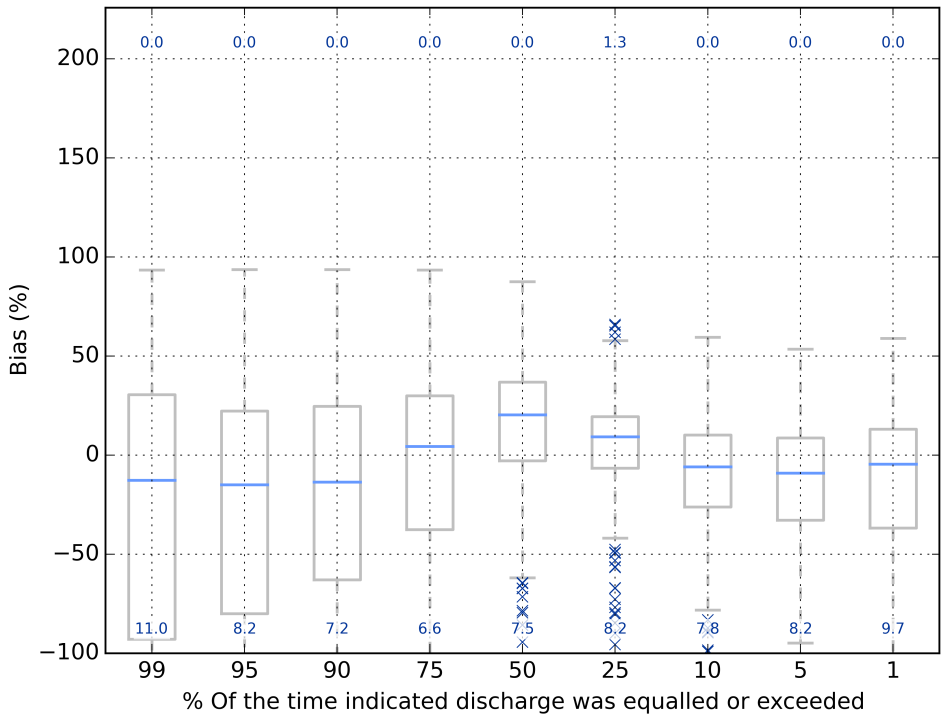



*Figure 5.5: A set of box plots showing the spread of percentage difference between simulated and observed flows at points on the flow duration curve. The small blue crosses are outliers and the blue numbers are the percentage of points that are outliers. Midflows are often overestimated whereas high and low flows are often underestimated.*

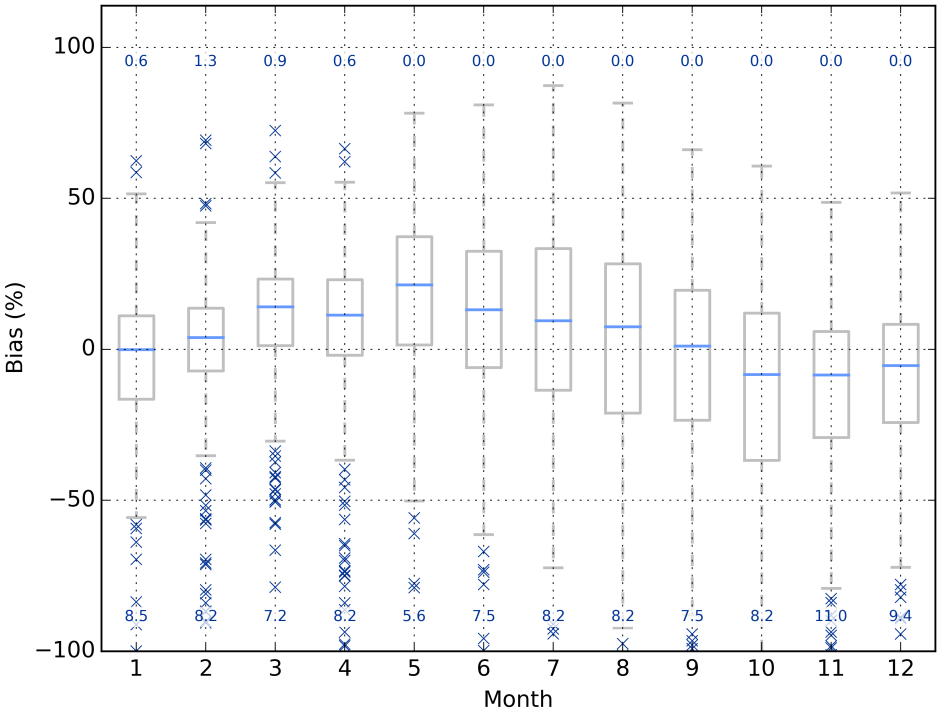



*Figure 5.6: A set of box plots showing the spread of percentage difference in water balance between simulated and observed flows. The small blue crosses are outliers and the blue numbers are the percentage of points that are outliers. The models generally under predict water balance.*

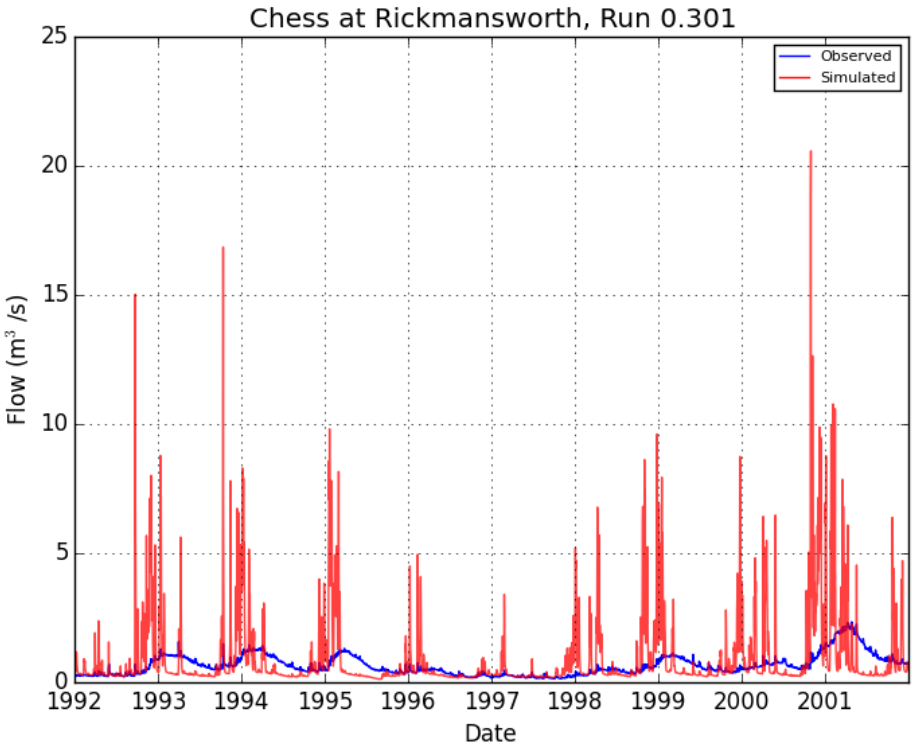



*Figure 5.7: Hydrograph for the Chess at Rickmansworth, initial run. SHETRAN does not capture the baseflow dominance in the hydrograph. NSE = -12.7.*

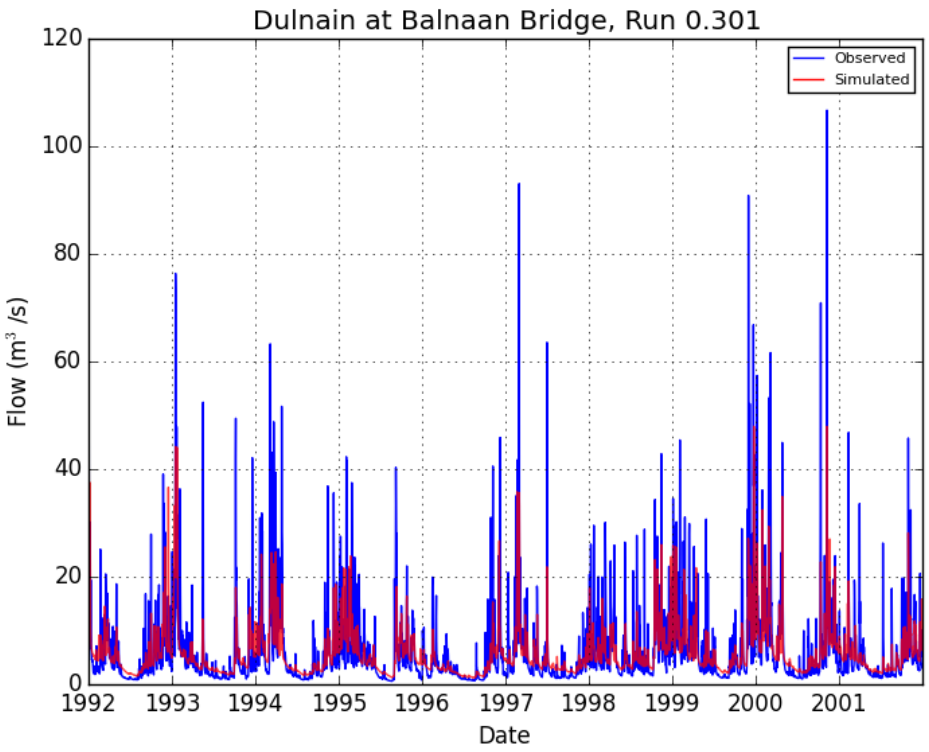



*Figure 5.8: Hydrograph for the Dulnain at Balnaan Bridge, initial run. SHETRAN generally under-predicts flows, especially in the spring months with snowmelt. NSE = 0.5.*

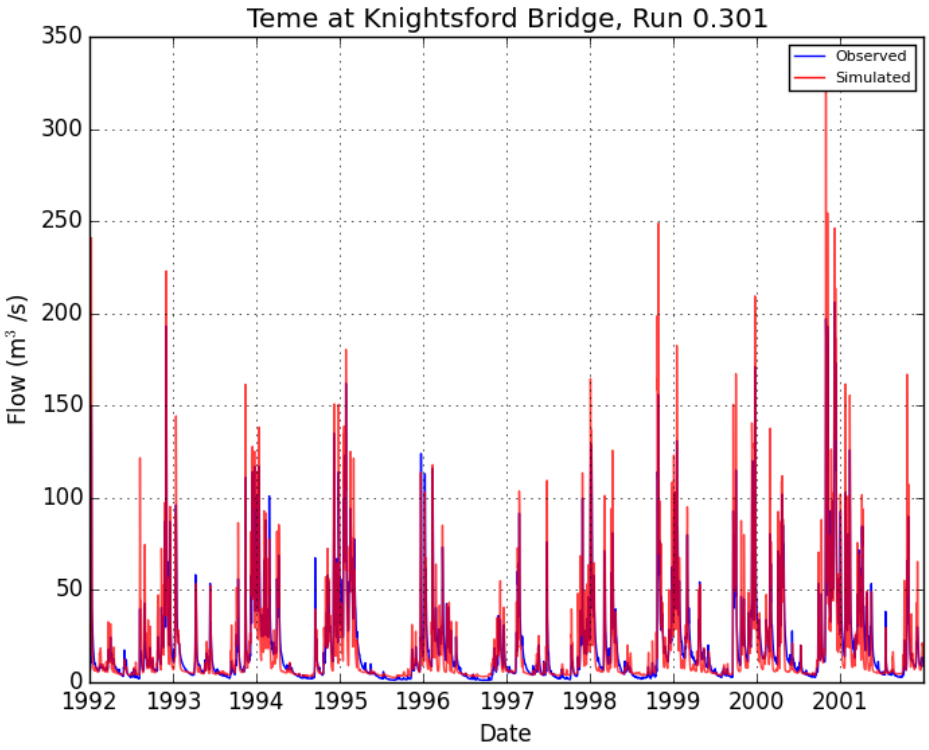



*Figure 5.9: Hydrograph for the Teme at Knightsford Bridge, initial run. Generally very good. the recessions are a little steep and the peaks are generally over estimated. NSE = 0.8.*

As discussed in chapter 3, precipitation inputs are taken from the UKCP09 dataset (Perry et al., 2009), the best source of distributed precipitation data available for this study, while PET is calculated using the FAO Penman-Monteith method (Allen et al., 1998). Uncertainties in both of these datasets could be at least partially responsible for apparent water balance biases where they occur, although the difference in sign of typical WB biases between summer and winter potentially indicates that there is not a seasonally consistent bias in the inputs. The initial simulation therefore highlights the need to explore the sensitivity of the SHETRAN national modelling system to uncertainty in rainfall/PET meteorological inputs. This has been investigated and is discussed later in this chapter. There may also be an influence on the water balance bias from abstractions and discharges in some catchments but this is not easily investigated without abstraction/discharge data and so has not been examined in this study.

The accuracy of the modelling system in simulating catchment flow regimes across the country can be further investigated through comparison of observed and modelled flow duration curves. Focusing on aggregate patterns, Figure 5.5 shows that high flows are often underestimated in catchments in the initial simulation, as to a lesser extent are low flows. In contrast, mid-range flows tend to be slightly overestimated overall. The underestimation of runoff at higher flow percentiles is consistent with the overall under-prediction of flow during the winter months apparent in the analysis of water balance bias above. A possible explanation of the underestimation of low flows could be that the simulations appear to exhibit flashier responses and steeper flow recessions relative to observed hydrographs, which could skew the frequency distribution of lower flows towards apparent underestimation (see Figure 5.9). The combination of underestimation at low flow percentiles and slight overestimation at medium flow percentiles could be responsible for the overall pattern of slight overestimation of the water balance in summer months. However, these general patterns should be viewed in the context of a substantial range of difference in observed and simulated runoff across the flow range between catchments, as indicated by the spread in Figure 5.5. This is reflected in the relatively high percentage of catchments with band 3 mmfd (59%) which indicates that the variation of differences at points on the flow duration curve and monthly water balance bias is higher than the annual/ overall bias.

From analysing the initial simulation, it is therefore apparent that many of the catchments are simulated adequately using a basic configuration of SHETRAN without calibration of parameters, when performance is judged in terms of the range of metrics outlined above. However, there are certain locations where simulated hydrographs, water balances and flow regimes do not correspond well with observations. Two geographical

areas subject to comparatively poor performance are the south east of England and northern Scotland. The south east of England is dominated by flat, relatively dry and warm catchments, in which groundwater derived from hydrogeologically varying and complex aquifers forms a highly significant component of stream flow in a number of cases. For example, many of the catchments in this area that exhibit fairly poor performance in the initial simulation are underlain particularly by major chalk aquifers. The hydrogeological properties and behaviour of these aquifers varies spatially as a result of factors such as weathering and water table fluctuations that significantly heighten transmissivity in river valleys and result in highly heterogeneous hydraulic property distributions, as well as variation in overlying deposits such as glacial till and Eocene clays that affect aquifer confinement and interactions with the surface water system (Shepley et al., 2012).

For reasons such as these, chalk aquifers are notoriously difficult to model (Bell et al., 2009; Crooks et al., 2014) and are clearly poorly represented in the initial simulation using this modelling system. The Chess at Rickmansworth (see Figure 5.7) is an example of a typical chalk catchment exemplifying this problem. Its hydrograph demonstrates how SHETRAN does not capture the baseflow-dominated regime of the catchment when using the parameter values and/or subsurface representation applied in the initial simulation. The integrated surface-subsurface formulation of SHETRAN may in theory provide a basis for modelling catchments of this nature, but it is clear that insufficient information on structural geology and hydraulic properties is contained within the datasets used in the initial simulation to achieve good performance (attempts to improve this can be found in chapter 7). In addition to this, appreciable abstraction takes place in many groundwater catchments; the absence of national abstraction data available for this work therefore adds an additional complication with respect to modelling these areas effectively.

The apparent biases and errors in the initial simulation of some Scottish catchments are perhaps easier to quantify. Firstly, seasonal snow is an important feature of the observed climatology and hydrology across large areas of Scotland. Snow accumulation and melt is not incorporated in the initial simulation, such that flows may be overestimated in the simulation in winter. Rather than precipitation being stored in the seasonal snowpack, as occurs in reality, it would be routed to streamflow much faster, particularly given the absence of soil moisture deficits during winter and the extensive low permeability geology present in Scotland. The initial simulation therefore also underestimates spring flows by not incorporating snowmelt processes as temperatures rise. Secondly, Scotland contains a number of large lochs, which are inadequately represented in the initial

SHETRAN configuration. This omission is manifest in incorrect timing of peak flows in some catchments, as well as overestimation of peaks where attenuation and storage of runoff is not represented. In addition, low flows may be underestimated if lochs are not properly accounted for, as the comparatively steady drainage of water from the loch under dry conditions is absent.

## 5.3 Structural changes

From the analysis of the initial simulation described above, it was identified that a number of structural improvements to the configuration of SHETRAN were required. These structural changes are the inclusion of the snowmelt module, the inclusion of a lake map and a way to fix sinks in the DEM that result in some numerical instabilities. Each of these changes to the model configuration are described in turn.

### 5.3.1 Representation of lakes

The incorporation of improved lake representation in the modelling system is an important modification, given that 60% of catchments contain at least 1 lake grid cell. Indeed, 51% of modelled catchments have lake areas representing more than 1% of total catchment areas, while 18% of catchments have lake areas accounting for 5% or more of total catchment areas. The extreme case is the Blackwater at Loch Dee in Scotland, where 29% of the catchment is covered by lake or loch area. In order to better account for the influence of these features, a lake map was created following the procedure described in chapter 3. For grid cells corresponding to lake areas in this map, the internal Strickler coefficient parameter in SHETRAN for bare ground was reduced from the initial spatially uniform value of 20 to a lower value of 3 . This adjustment effectively acts to slow flow routed through lakes, thereby inducing attenuation of flow and longer storage of water in lake cells.

In order to evaluate the changes to streamflow simulation arising from incorporation of lakes with an improved parameterisation, hydrographs comparing observed and simulated flows for several catchments containing significant lake areas have been examined, of which one example is presented here. The Leven at Newby Bridge is an upland, impervious catchment completely containing Lake Windermere, a large lake of 14.7km$^2$ (6% of the catchment area). As Windermere constitutes a significant fraction of the catchment area, it will clearly play a large role in modulating overall catchment be-

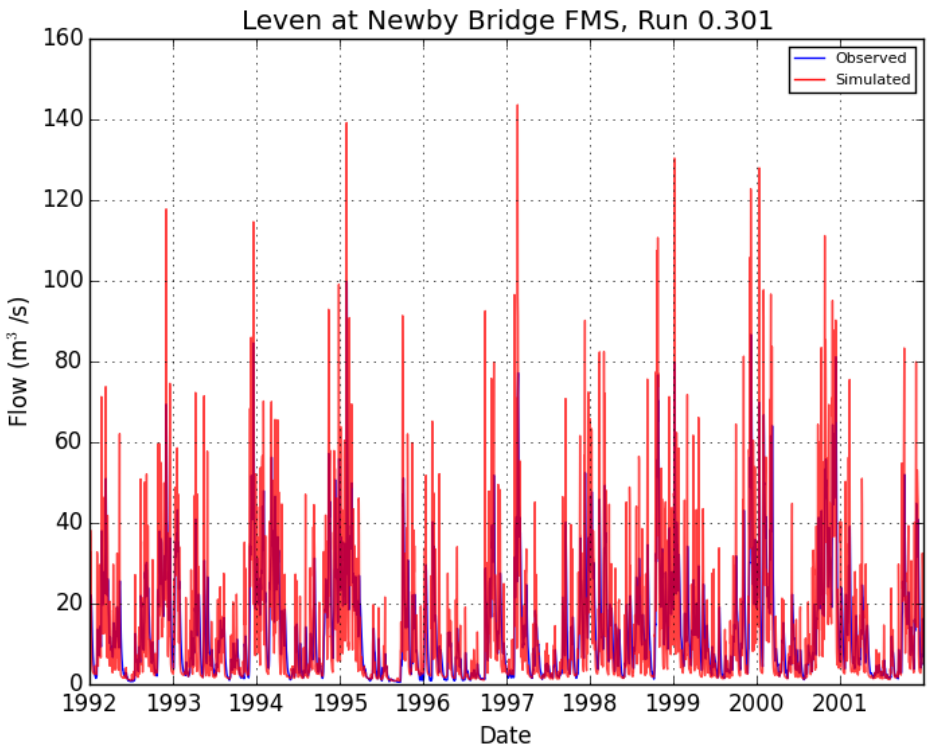



*Figure 5.10: Hydrograph showing initial run of Leven at Newby Bridge. Without the inclusion of a lake, the simulated peaks are too large, the recessions are too steep and the low flows are too low.*

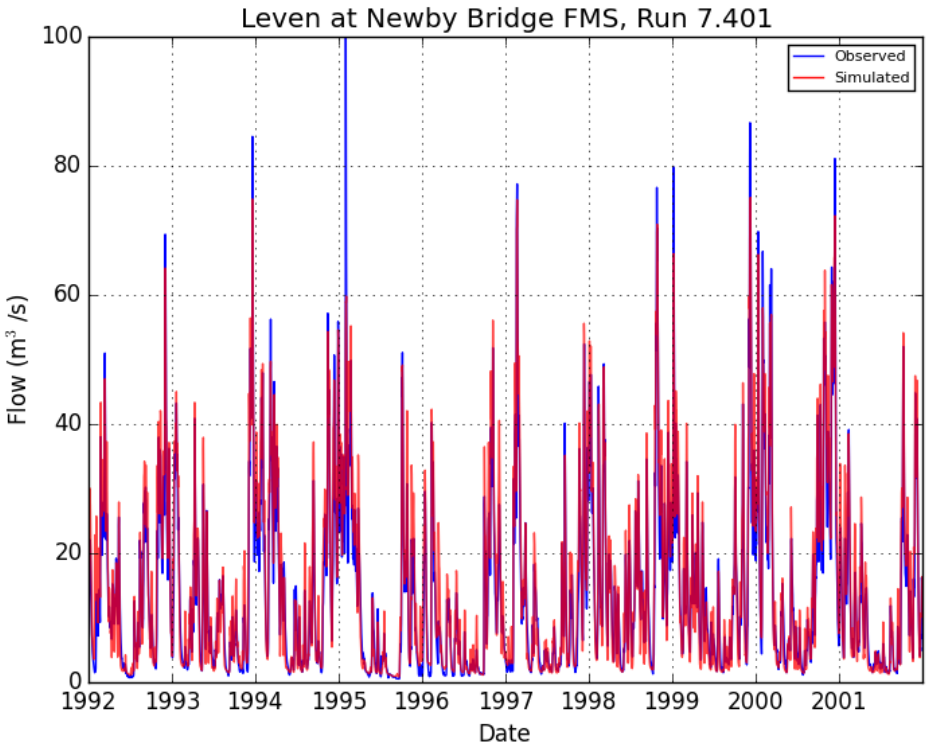



*Figure 5.11: Hydrograph showing revised run of Leven at Newby Bridge. The introduction of a lake has attenuated the peaks and regulated the low flow runoff.*

haviour, which is determined from the flow gauging station located a short distance downstream of the lake outlet. From Figure 5.10 it is apparent that the initial simulation (i.e. without the lake map) is associated with substantial overestimation of peak flows, as well as underestimation of low flows. The simulated catchment response is much flashier compared with that seen in the observations. Yet when the lake map is introduced, peak flows are much more accurately simulated, as can be seen in Figure 5.11. This is a result of more realistic approximation of the attenuating effect of the large lake. In addition, the role of the lake in temporarily storing water to sustain low flows is clear from comparing Figures 5.10 and 5.11. The consequences of improved lake representation are particularly evident in winter months, during which runoff into the lake is significantly delayed before reaching the catchment outlet.

### 5.3.2 Removing sinks from the digital elevation model

The original DEM applied in the initial simulation was not checked for 'sinks'. Sinks occur where a grid cell is lower than all of its surrounding grid cells, which causes water to artificially accumulate. This sometimes creates numerical instabilities in the program and causes SHETRAN to crash, but the main consequence of this issue is that water is not routed correctly. As such, a new algorithm was added to the SHETRAN pre-processing 'prepare' program, in order to remove sinks from the mean DEM. This is considered to be a conceptually coherent inclusion for modelling in this context and useful for conducting runs across a large number of catchments in an efficient way. Adding in this step in the modelling process results in increased flows and decreased catchment storage. In some catchments flows were increased by a considerable amount when sinks were removed, as reflected in the differences in water balance bias between the initial simulation and the simulation incorporating the sink corrections (see Figures 5.4 and 5.15).

### 5.3.3 Representation of snowmelt processes

Seasonal snowfall represents a significant component of the water balance and dynamics of a number of regions of Great Britain, including upland areas such as the highlands of Scotland, Snowdonia and Cumbria amongst others. As such, incorporating snow accumulation and melt processes is intended to account for the role of snow in modulating catchment runoff through seasonal storage and runoff processes.

SHETRAN has two methods for modelling snow: a full energy balance approach or a simpler degree day method. Employing an energy balance approach requires specification of a number of spatially distributed inputs, including incoming shortwave and longwave radiation, air temperature, relative humidity and wind speed. The full suite of meteorological data required to drive the energy balance method were not available, such that the degree day method was implemented in preference. This approach requires only maximum and minimum daily temperatures as inputs by taking advantage of the strong correlation between air temperature and melt rates, which in turn derives from the relationship between air temperature and components of the energy balance (Zuzel and Cox, 1975; Hock, 2003, 2005). In addition a single parameter, the degree day factor, must be specified to control the rate of melt associated with positive degree days. The degree day factor is an empirical parameter that may vary both spatially (with vegetation, slope and other variables) and temporally (Hock, 2003, 2005). For this project, it was set at a typical value of 0.0002 $mms^{-1}{}^{o}C^{-1}$ for all areas (Birkinshaw, 2011). Snowmelt occurs if the air temperature is greater than 0 degrees C and the rate of melt depends on the temperature, degree day factor and the specific gravity of the snow (Birkinshaw, 2011).

The improvements that occur from including the snowmelt module can be seen in Figure 5.13 relative to Figure 5.12. Comparing these two figures, it is notable that, in the springs of 1994 and 2001 in particular, flows are higher and decrease more gradually when snow processes are included relative to the initial simulation.

## 5.4 Standard set up of SHETRAN for GB

Each of the three structual changes described above are considered to improve the representation of catchment processes in the SHETRAN modelling system in a coherent manner, such that they have all been brought forward from the initial simulation into the standard configuration of SHETRAN for Great Britain. The standard simulation forms the basis for all subsequent sensitivity tests and scenario runs reported in this chapter and also in subsequent chapters examining climate change scenarios (chapter 6) and integration of a 3D geological model created by the British Geological Survey (BGS) (Mathers et al., 2014) (chapter 7). This section proceeds to characterise the overall performance of the standard simulation, in order to evaluate the modelling system as a whole and provide context for the applications and modifications described in later chapters.

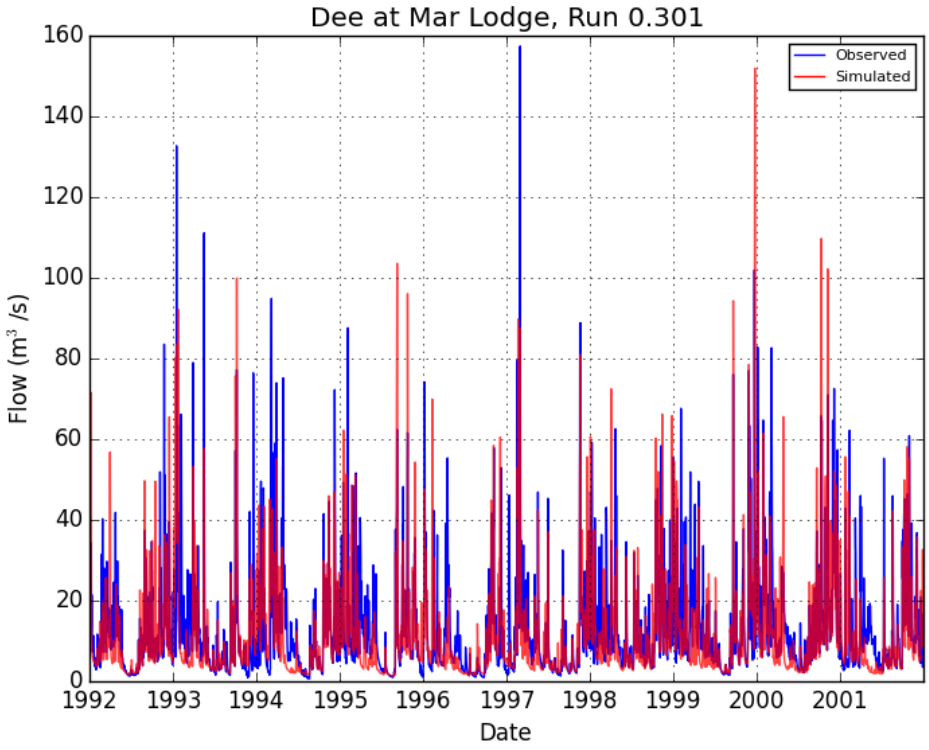



*Figure 5.12: Hydrograph showing initial run of Dee at Mar Lodge. Note the poor performance of the model in the spring of 1994 and 2001 when the snowmelt module is not included.*

| Measure | NSE | Bias | MM | FD | mmfd |
|---|---|---|---|---|---|
| Median | 0.69 | 5.1 | 18.7 | 25.3 | 42.9 |
| $25^{th}$ Percentile | 0.39 | -2.2 | 11.6 | 14.5 | 27.2 |
| $75^{th}$ Percentile | 0.79 | 18.3 | 33.3 | 41.0 | 72.4 |
| % Band 1 | 22.5 | 52.3 | - | - | 12.7 |
| % Band 2 | 40.8 | 22.5 | - | - | 34.3 |
| % Band 3 | 36.6 | 25.2 | - | - | 52.9 |

*Table 5.3: Summary table of performance statistics for standard run.*

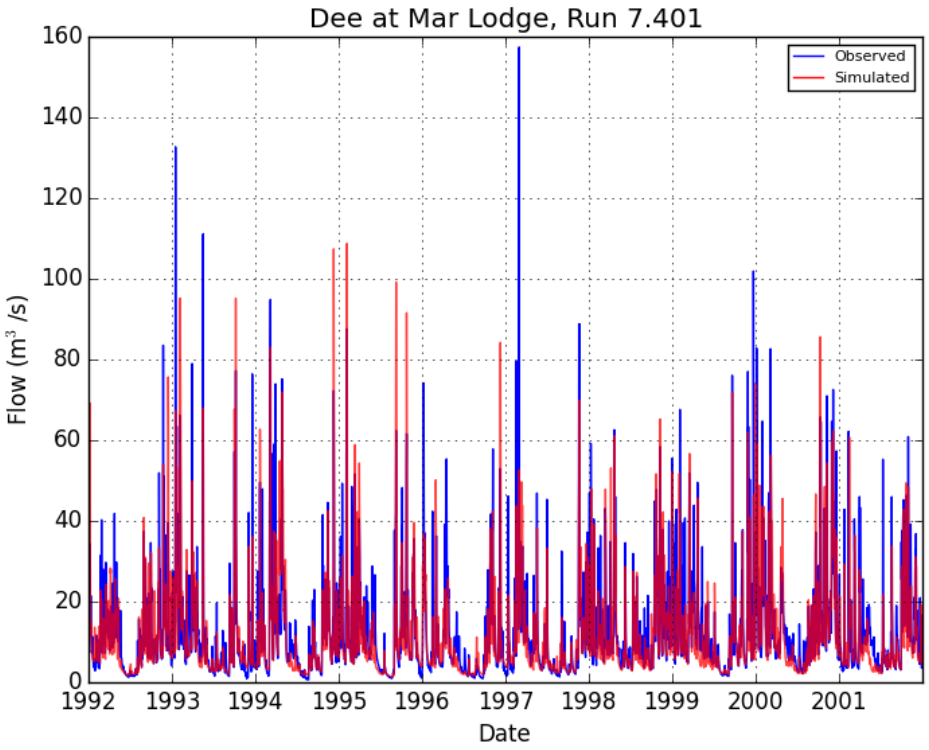



*Figure 5.13: Hydrograph showing revised run of the Dee at Mar Lodge. The introduction of snow module has increased runoff in the springs of 1994 and 2001.*

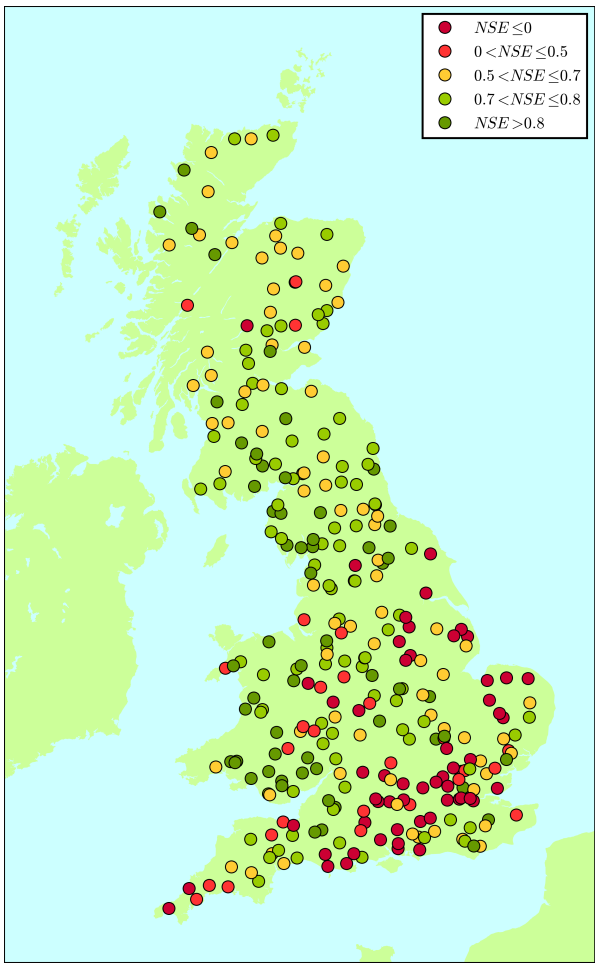



*Figure 5.14: Map showing NSE for the standard SHETRAN for GB. The majority of poorly performing catchments are groundwater dominated in the South East of England.*

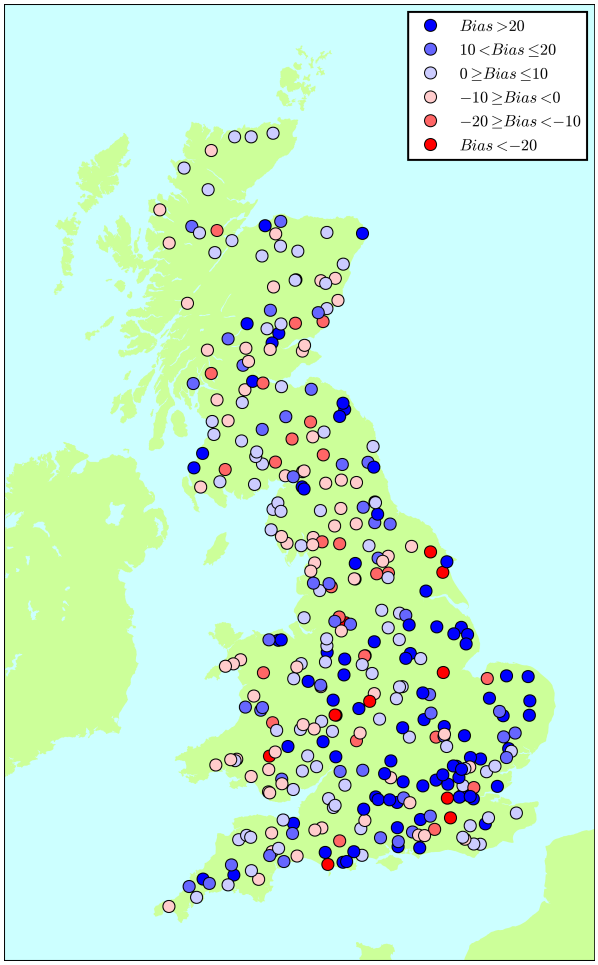



*Figure 5.15: Map showing water balance bias of standard SHETRAN run. The points represent gauge locations in each catchment. A pale colour indicates a small bias (good), darker indicates a larger bias (bad), red indicates that the simulation is under-predicting flow, blue, over-predicting. There is now a mostly small positive bias in the water balance, and now shows a coherent spatial pattern with large overestimations in the South East of England.*

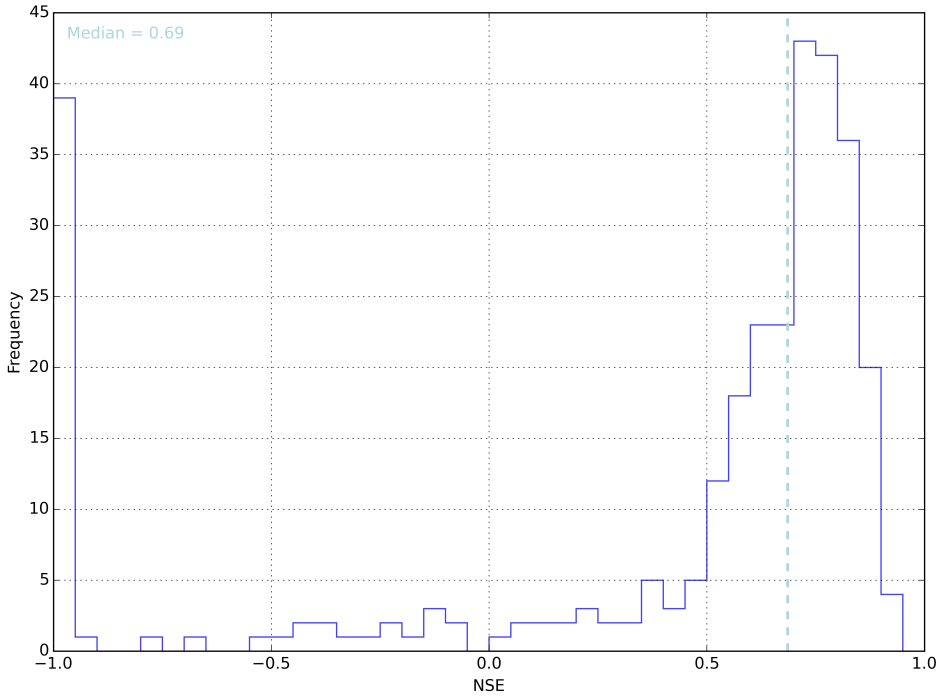



*Figure 5.16: The distribution of NSE values. The majority of catchments have an NSE of 0.5 or greater. the median NSE is 0.69. 39 catchments have an NSE of 0 or lower.*

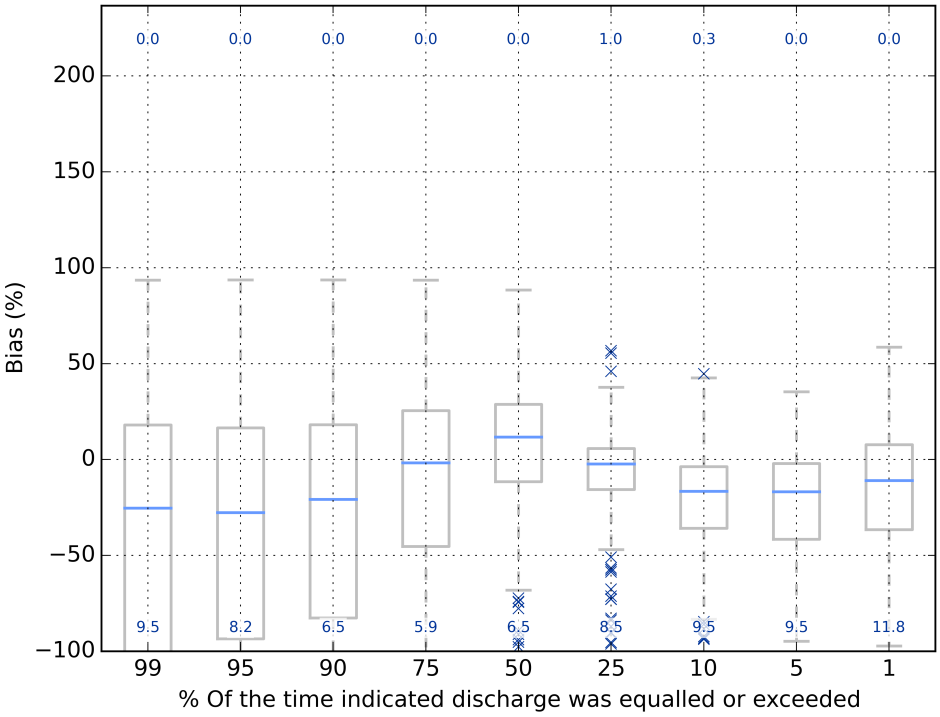



*Figure 5.17: A set of box plots showing the spread of percentage difference between simulated and observed flows at points on the flow duration curve. The small blue crosses are outliers and the blue numbers are the percentage of points that are outliers. Midflows are often overestimated whereas high and low flows are often underestimated.*

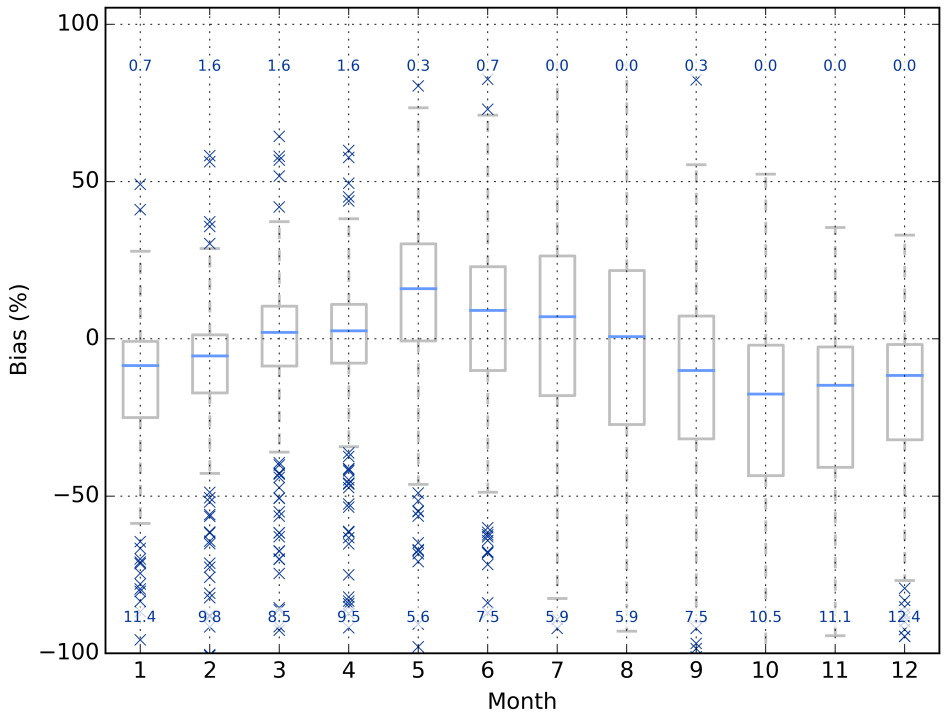



*Figure 5.18: A set of box plots showing the spread of percentage difference in water balance between simulated and observed flows. The small blue crosses are outliers and the blue numbers are the percentage of points that are outliers. The models generally under predict flows in the winter and over predict in the summer.*

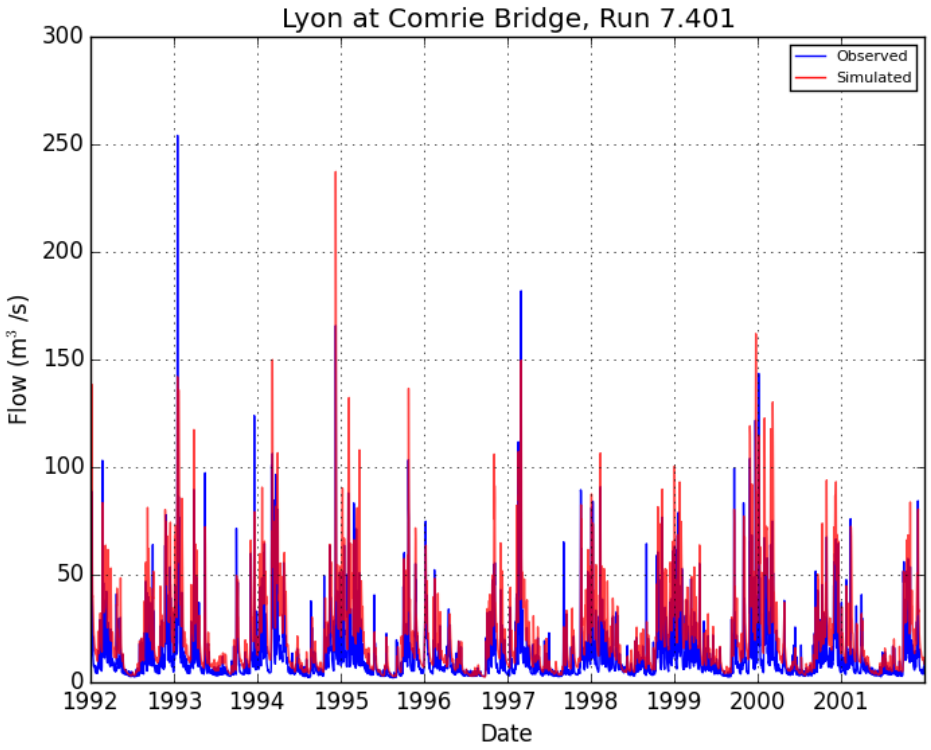



*Figure 5.19: Poor model performance due to runoff affected by hydropower diversions. This information is not captured in the SHETRAN input data.*

From Table 5.3 it can be seen that 74% of catchments in the standard run are associated with NSE values greater than 0.5, while 50% of modelled catchments have NSE values greater than 0.69. This represents a similar level of performance to the initial simulation. both the $25^{th}$ and $75^{th}$ percentiles increase slightly from 0.35 to 0.39 and 0.77 to 0.79 respectively. Reviewing the spatial distribution of NSE values in Figure 5.14 shows that upland catchments in the north of England and Scotland and Wales perform well on the whole. However, Figure 5.14 also shows that the structural improvements implemented between the initial and standard simulations do not resolve all difficulties, particularly in simulating chalk catchments. This is as expected, given that the structural changes made here are not focused on representation of geology, partly as a result of limitations to available national-scale datasets. Some initial work to rectify this deficiency is outlined in chapter 7.

The majority of catchments in Scotland that were subject to relatively poor NSE values in the initial simulation show improvements in the standard run. However, some catchments in this region are still potentially problematic. For example, the Lyon at Comrie Bridge has a poor NSE of -0.06. From the hydrograph (Figure 5.19) it appears that modelled flows are much higher than observations. Investigating this error revealed that it is not easily explained by catchment characteristics derived from the input data. Rather, in the National River Flow Archive (NRFA) catchment description and metadata, this basin is described as having hydropower diversions that greatly reduce catchment runoff, which fits with the pattern of overestimation of flows in the SHETRAN system. This case therefore provides an example of a catchment where SHETRAN has the potential for accurate simulation, but its accuracy is confounded by influences not captured by the available input dataset. Fortunately, it appears that for the majority of catchments selected for analysis in this study where the NRFA details artificial factors affecting runoff, the influence of abstractions and discharges are sufficiently limited that streamflows can be well simulated.

Water balance bias (Figure 5.18) and differences in observed and simulated flow duration curves (Figure 5.17) remain largely unchanged in distribution in the standard run when compared with the earlier initial simulation (see Figures 5.5 and 5.6). Table 5.3 indicates that the median water balance bias is slightly higher in the standard simulation at 5.1% (compared with -2.2% in the initial run). The standard run also shows a larger under-prediction of flow in the winter months overall, which could be the result of storage as snow in the winter months or storage in lakes. The flow duration curve distribution has the same shape as in the initial run, but on average flows are typically higher across the full flow range. This is reflected in the change in interquartile range

of the water balance bias which shifts upwards (from -13.3 to -2.2 at the lower end and from 10.35 to 18.29 at the upper end). This actually increases the percentage of catchments with band 1 water balance bias from 41 to 52%. The percentage of catchments with band 1 mmfd also increases from 4 to 13%. The spatial distribution of water balance bias can be seen in Figure 5.15. There is now a mostly small positive bias in the water balance, and now shows a coherent spatial pattern with large overestimations in the South East of England.

The implication of this discussion is that, prior to any adjustments to the best estimate parameter set, the national modelling system set up in this project appears to perform quite well overall when evaluated with respect to a typical classification of NSE values, as well as various measures of bias that examine different aspects of the flow regime. In addition to the fact that 74% of catchments have a good or reasonable performance when assessed through NSE (i.e. with values greater than 0.5), qualitative assessment of hydrographs confirms a good degree of correspondence between observations and simulations in terms of the timing and amplitude of peaks and the shape of flows recessions (see Figure 5.29). While correspondence between observed and simulated hydrographs is therefore very reasonable, analysis of water balance and flow duration curves reveals notable variation in the match between gauged and simulated flow regimes. Some of these differences may be attributable to confounding factors that cannot be accounted for in the scope of this work, for example artificial influences on catchment flows. However, other explanations for some of this variation may exist, including uncertainties in input data or model parameters. Therefore, a series of sensitivity tests were conducted, the results of which are presented below.

### 5.4.1 Confirmation of results

The sensitivity tests described in this chapter were conducted for the period 1992-2002 (with a 2 year spin-up period of 1990-1991). To confirm that these results are typical of model performance and not just a function of the particular evaluation period, the standard simulation was run again but this time for the period 1982-1992 (again with a 2 year spin-up). Only 245 catchments have flow records that span 1980-2002 and could be compared. as Figure 5.20 shows a scatterplot of the NSE for each catchment, with the NSE values for the 1982-1992 plotted against the scores for the 1992-2002 period. This shows that the system performs similarly for both periods, but with some variation across catchments. This could be due to a number of reasons, such as differences in the quality of rainfall/PET inputs or flow records for the earlier period, as well as the fact

| Metric | Statistic | Standard | Validation |
|---|---|---|---|
| NSE | Mean NSE Change | 0.00 | -1.50 |
| | % Improved | 0.0 | 24.9 |
| | Median | 0.69 | 0.66 |
| | Change in Median | 0.00 | -0.03 |
| | $25^{th}$ Percentile | 0.37 | 0.23 |
| | $75^{th}$ Percentile | 0.78 | 0.77 |
| | % Band 1 | 21.2 | 18.1 |
| | % Band 2 | 42.4 | 40.6 |
| | % Band 3 | 36.3 | 41.4 |
| Water Balance | % Improved | 0.0 | 50.2 |
| | Median | 5.8 | 5.3 |
| | Change in Median | 0.0 | -0.4 |
| | $25^{th}$ Percentile | -1.6 | -1.1 |
| | $75^{th}$ Percentile | 18.9 | 20.7 |
| | % Band 1 | 51.8 | 50.6 |
| | % Band 2 | 22.4 | 22.1 |
| | % Band 3 | 25.7 | 27.3 |
| mmfd | % Improved | 0.0 | 42.4 |
| | Median | 43.8 | 45.1 |
| | Change in Median | 0.0 | 1.3 |
| | $25^{th}$ Percentile | 27.4 | 27.4 |
| | $75^{th}$ Percentile | 73.5 | 81.2 |
| | % Band 1 | 12.7 | 12.9 |
| | % Band 2 | 33.1 | 30.1 |
| | % Band 3 | 54.3 | 57.0 |

*Table 5.4: Summary table of performance statistics for standard run and validation run.*

that 1982-1992 was a generally wetter period. This could result in peakier hydrographs to which NSE scores would be sensitive.

On the whole, performance is similar but very slightly worse in the validation run. The median NSE decreases from 0.69 to 0.66, the $25^{th}$ percentile NSE decreases from 0.37 to 0.23 and the $75^{th}$ percentile NSE decreases from 0.78 to 0.77. Water balance bias over the interquartile range remains very similar only varying by up to 2%. The same is true of mmfd, which only varies noticeably in the $75^{th}$ percentile from 74% to 81%.

## 5.5 General sensitivity of the modelling system

### 5.5.1 Sensitivity to the Strickler coefficient

In SHETRAN, surface water is routed using the diffusive wave approximation of the St. Venant equations. These equations require a friction factor to represent raindrop im-

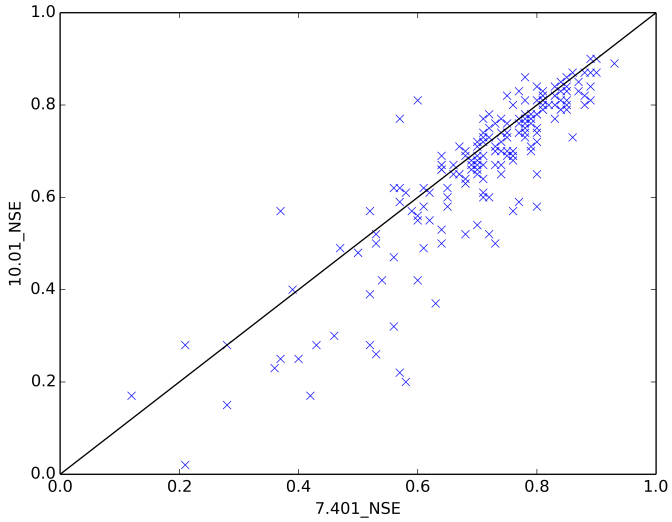



*Figure 5.20: NSE of the standard set up of SHETRAN for GB for the period 1992-2002 on the x axis ,vs 1982-1992 on the y axis. Each point represents a catchment. Model performance is similar in both periods. performance is slightly worse during the 1980s.*

pact, channelization of flow, obstacles such as litter, crop ridges, rocks, and roughness from tillage, the frictional drag over the surface, and erosion and transport of sediment (Engman, 1986). This factor, essentially a roughness coefficient, is represented in SHETRAN by the Strickler coefficient, which is the inverse of the Manning roughness coefficient. As such, the lower the value of the Strickler coefficient, the higher the roughness, which can result in slower flow of water and shallower gradients in hydrograph fluctuations. Conversely, a high value of the Strickler coefficient will cause water to runoff faster resulting in steep recessions in the hydrograph.

The Strickler coefficient can therefore have great control over the shape of the event hydrographs, and prior studies have shown that SHETRAN is very sensitive to this parameter (Zhang et al., 2013). Unlike other parameters in SHETRAN, the Strickler coefficient is not currently tied to any physical catchment characteristics or model input datasets. Given its importance and global application in SHETRAN, the sensitivity of model results to changes in the Strickler coefficient were investigated. Whereas the default Strickler coefficient applied in SHETRAN simulations (including the initial and standard runs presented earlier) on the basis of prior experience is typically 1 (Birkinshaw, 2011), three sensitivity runs were conducted in which the Strickler coefficient across all catchments was set to values of 0.1, 2 and 5 respectively. The results of these sensitivity tests are discussed below.

| Metric | Statistic | Standard | Strickler 0.1 | Strickler 2 | Strickler 5 |
|---|---|---|---|---|---|
| NSE | Mean NSE Change | 0.00 | 1.23 | -0.51 | -1.16 |
| | % Improved | 0.0 | 36.9 | 30.7 | 24.4 |
| | Median | 0.69 | 0.55 | 0.66 | 0.6 |
| | Change in Median | 0.00 | -0.14 | -0.03 | -0.09 |
| | $25^{th}$ Percentile | 0.39 | 0.43 | 0.22 | 0.02 |
| | $75^{th}$ Percentile | 0.79 | 0.69 | 0.78 | 0.76 |
| | % Band 1 | 22.5 | 12.0 | 22.0 | 16.4 |
| | % Band 2 | 40.8 | 28.9 | 37.1 | 34.2 |
| | % Band 3 | 36.6 | 59.1 | 41.0 | 49.5 |
| Water Balance | % Improved | 0.0 | 64.1 | 30.3 | 31.9 |
| | Median | 5.1 | 3.3 | 6.0 | 6.2 |
| | Change in Median | 0.0 | -1.9 | 0.9 | 1.1 |
| | $25^{th}$ Percentile | -2.2 | -4.3 | -2.0 | -2.1 |
| | $75^{th}$ Percentile | 18.3 | 15.0 | 19.1 | 18.9 |
| | % Band 1 | 52.3 | 53.9 | 52.1 | 51.3 |
| | % Band 2 | 22.5 | 25.0 | 22.3 | 23.3 |
| | % Band 3 | 25.2 | 21.1 | 25.6 | 25.5 |
| mmfd | % Improved | 0.0 | 50.2 | 24.3 | 22.2 |
| | Median | 42.9 | 43.3 | 43.7 | 45.7 |
| | Change in Median | 0.0 | 0.4 | 0.7 | 2.7 |
| | $25^{th}$ Percentile | 27.2 | 28.9 | 29.0 | 31.6 |
| | $75^{th}$ Percentile | 72.4 | 71.7 | 74.5 | 76.0 |
| | % Band 1 | 12.7 | 10.1 | 10.8 | 7.6 |
| | % Band 2 | 34.3 | 34.7 | 32.8 | 32.7 |
| | % Band 3 | 52.9 | 55.2 | 56.4 | 59.6 |

*Table 5.5: Summary table of performance statistics for Strickler coefficients of 1, 0.1, 2 and 5.*

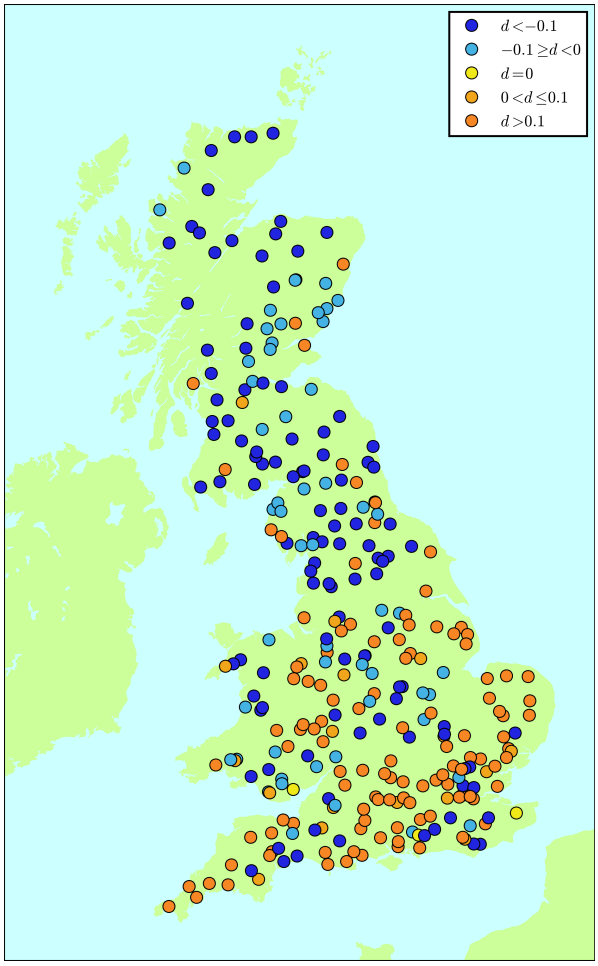



*Figure 5.21: Map showing the difference (d) to NSE due to reducing the Strickler coefficient from 1 to 0.1. Blue indicates a decrease in NSE and orange an increase. Model performance for southern catchments improves as the decrease in Strickler reduces the peakiness of the hydrograph.*

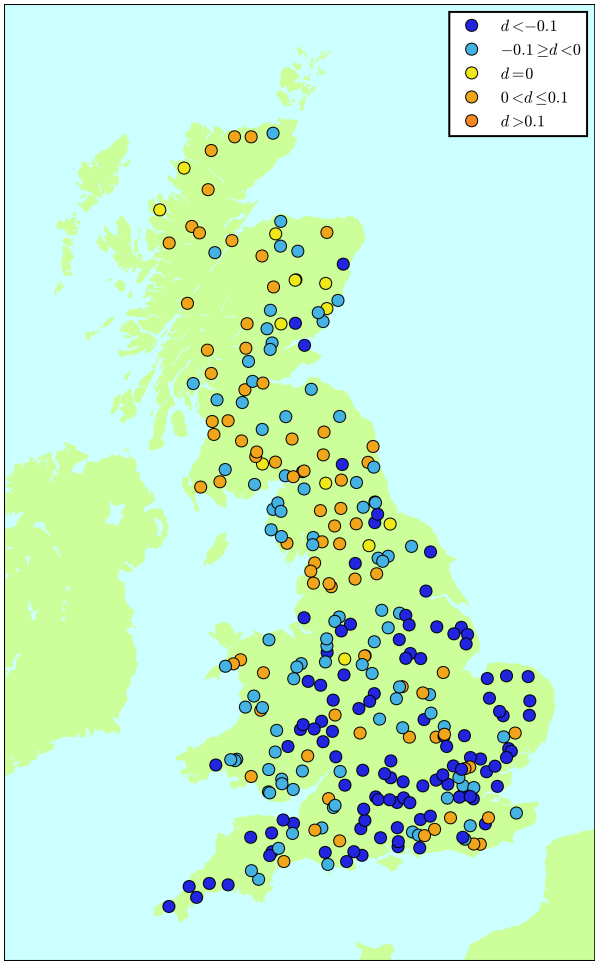



*Figure 5.22: Map showing the difference (d) to NSE due to increasing the Strickler coefficient from 1 to 2. Blue indicates a decrease in NSE and orange an increase. Model performance in southern catchments decreases with an increase of Strickler as the hydrographs are now peakier.*

The results of altering the Strickler coefficient across all catchments are summarised in Table 5.5, which shows changes in each of the performance metrics with respect to the standard run. From this table it can be seen that, in aggregate terms, neither increasing nor decreasing Strickler globally has a positive impact on NSE scores. The median NSE value for all catchments decreased in each of the simulations with trial Strickler coefficients relative to the standard run, with the percentage of catchments in the lowest classification interval (band 3) increasing in all cases. Applying a Strickler coefficient of 0.1 showed the largest number of improved catchments in terms of NSE, with 37% of catchments having higher values relative to the standard run. Using a Strickler coefficient of 5 results in improvement for only a handful of catchments in the north of Great Britain - and deterioration in other areas - such that this run is not analysed further.

Analysing the spatial distribution of NSE changes that result from varying the Strickler coefficient indicates some interesting geographical patterns. Figure 5.21 shows that improvements in NSE arising from application of a lower Strickler coefficient are largely seen in the southern part of Great Britain. In general terms, a significant proportion of catchments in this region are typically flatter, with notable contributions from groundwater in a number of areas. Reducing the Strickler coefficient in this region could have the effect of lowering catchment response times and inducing shallower recessions in the hydrograph, i.e. compensating for the physical processes occurring in the catchment that are not accurately represented in the model, particularly with respect to groundwater-surface water interactions.

When the Strickler coefficient is increased from 1 to 2, an inverse pattern in the spatial distribution of NSE changes is apparent compared with that described above. Figure 5.22 shows how increasing the coefficient exerts an apparently positive effect on the NSE values of catchments in the north of Great Britain. Generally there are a larger number of steep, upland catchments in this region, and it appears that the higher Strickler coefficient inducing faster overland flow could be favourable here. However, not all catchments in Scotland improve as a result of this adjustment, which could be a result of lower roughness conflicting with the representations of lochs/lakes or snowmelt runoff described above. This could therefore reflect the potential for significant interdependencies between parameterisations and parameter values in a complicated model such as SHETRAN.

It is also of note that the median water balance bias did slightly reduce in the run where the Strickler coefficient was set to 0.1 relative to the standard run. This is consistent

with the way in which a lower Strickler coefficient slows runoff and thereby enables more evaporation, which would reduce the overall negative bias slightly. However, the converse is not apparent, i.e. applying a higher Strickler coefficient does not appear to degrade water balance bias substantially further, which suggests that the implications on the water balance of changing Strickler may have some dependence on the direction of change. With a Strickler of 0.1 low flows were largely underestimated and high flows were over estimated and with a Strickler of 2 low flows and high flows were both underestimated.

Therefore, the sensitivity tests on the Strickler coefficient appear to confirm that a value of 1 performs reasonably overall as a universal parameter, in line with prior applications of SHETRAN (Birkinshaw, 2011). Crucially, the analysis here demonstrates that further improvements in the specification of this parameter are unlikely to be attainable through global adjustment of the parameter, i.e. no single value is likely to be optimal for all of the different types of catchments in the sample used in this study. However, it may be possible in future work to relate Strickler to catchment properties, such as steepness (where it would act as a calibration factor) or land cover, which would seem intuitively to have some physical relationship to roughness properties. Indeed, reviewing the land cover distribution presented in chapter 3 shows some geographical patterns that are at least partially consistent with the sensitivity test results discussed here. For example, the majority of land cover is arable in the south of the country compared with predominantly grassland in the north, which fits to some degree with the differences in model performance arising from increasing or decreasing the Strickler coefficient, but in reality defining the relationship between land use and roughness may need to account for other factors. This would be an interesting avenue for further work.

### 5.5.2 Sensitivity to rainfall rates

There are many errors associated with rain gauge records, the most significant of which is wind-induced undercatch (Michelson, 2004). These errors were not accounted for in the construction of the UKCP09 5km gridded dataset (Perry et al., 2009) and so a sensitivity test was conducted to investigate whether or not any systematic biases in the model input dataset were present and potentially attributable to undercatch. The two sensitivity tests used to explore possible undercatch consist of increasing the daily rainfall by 10 and 20% for all grid squares covering Great Britain. However, from the evaluation described above, it is apparent that the median water balance bias in the standard run is an overestimation of 5%. This direction of bias could be interpreted as

| Metric | Statistic | Standard | -10% rain | -20% rain | +10% rain | +20% rain |
|---|---|---|---|---|---|---|
| NSE | Mean NSE Change | 0 | 1.42 | 2.19 | -2.04 | -4.66 |
| | % Improved | 0.0 | 59.5 | 40.3 | 13.0 | 6.0 |
| | Median | 0.69 | 0.71 | 0.61 | 0.55 | 0.25 |
| | Change in Median | 0 | 0.02 | -0.08 | -0.14 | -0.44 |
| | $25^{th}$ Percentile | 0.39 | 0.53 | 0.45 | -0.11 | -0.87 |
| | $75^{th}$ Percentile | 0.79 | 0.79 | 0.71 | 0.72 | 0.6 |
| | % Band 1 | 22.6 | 23.4 | 7.0 | 7.9 | 2.0 |
| | % Band 2 | 40.9 | 46.2 | 46.7 | 38.4 | 24.0 |
| | % Band 3 | 36.6 | 30.5 | 46.3 | 53.8 | 74.0 |
| Water Balance | % Improved | 0.0 | 40.1 | 20.3 | 12.7 | 5.0 |
| | Median | 5.1 | -12.1 | -28.3 | 23.9 | 41.9 |
| | Change in Median | 0.0 | -17.2 | -33.4 | 18.8 | 36.8 |
| | $25^{th}$ Percentile | -2.2 | -18.0 | -35.5 | 13.0 | 28.4 |
| | $75^{th}$ Percentile | 18.3 | -2.5 | -21.4 | 42.4 | 65.9 |
| | % Band 1 | 52.3 | 33.3 | 3.5 | 17.4 | 2.0 |
| | % Band 2 | 22.6 | 38.5 | 12.5 | 22.0 | 9.2 |
| | % Band 3 | 25.2 | 28.2 | 84.0 | 60.7 | 88.8 |
| mmfd | % Improved | 0.0 | 52.3 | 28.5 | 8.7 | 3.0 |
| | Median | 43.0 | 46.1 | 65.6 | 63.3 | 92.7 |
| | Change in Median | 0.0 | 3.1 | 22.7 | 20.4 | 49.8 |
| | $25^{th}$ Percentile | 27.2 | 32.5 | 53.6 | 41.8 | 67.7 |
| | $75^{th}$ Percentile | 72.4 | 63.0 | 83.2 | 104.2 | 144.0 |
| | % Band 1 | 12.8 | 3.5 | 0.6 | 3.0 | 0.0 |
| | % Band 2 | 34.3 | 36.9 | 4.8 | 20.7 | 5.6 |
| | % Band 3 | 52.9 | 59.6 | 94.6 | 76.4 | 94.4 |

*Table 5.6: Summary table of performance statistics for changes to input rainfall of -10%, -20%,+10%, +20%.*

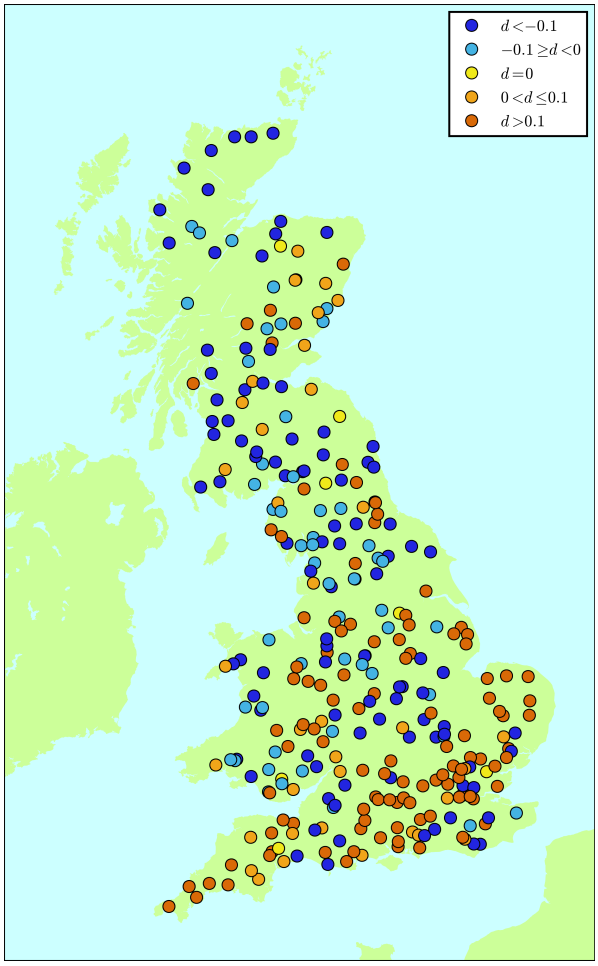



*Figure 5.23: Map showing changes to NSE due to decreasing rainfall by 20%. Blue indicates a decrease in NSE and orange an increase. Model performance improves in the south of England where flow is normally over predicted.*

running counter to the hypothesis of possible undercatch, but it is difficult to attribute the bias to rainfall inputs immediately, given that uncertainties also exist in PET inputs and model parameters. In order to explore the issue of water balance bias further, additional tests were also carried out in which universal reductions in rainfall of 10 and 20% for all grid squares were also tested, which in effect allow for exploration of possible limitations in interpolation of point rainfall in creation of the UKCP09 gridded dataset. These initial sensitivity tests are clearly crude and do not account for factors such as spatial variation in the probability of undercatch amongst other issues, but they do provide an initial means of checking rainfall inputs. More sophisticated treatments of uncertainties in rainfall inputs are discussed in subsequent sections.

Of the four sensitivity tests, reducing the rainfall by 10% actually exerts the largest positive impact on model performance when judged in terms of NSE. In this test, NSE in 60% of catchments improved, with the median NSE increasing from 0.69 to 0.71. These improvements are mainly associated with a number of catchments moving from band 3 to band 2 in the classification scheme. This increase in performance of previously poorly performing catchments is also reflected in the change in the $25^{th}$ percentile of the NSE distribution from 0.39 to 0.53, which is principally related to improvements in a number of baseflow-dominated catchments in the southern part of the country. In these cases, a reduction in rainfall reduces the magnitude of hydrograph peaks, thus improving NSE without actually rectifying the more fundamental problem of appropriate representation of the groundwater system in these catchments. This is emphasised in Figure 5.23, which shows that with a large reduction in rainfall, the groundwater dominated catchments show the largest improvement (along the warmest areas of the country, for example, Cornwall) whilst catchments that previously had good model performance worsen.

Reducing rainfall uniformly by 10% leads to a median reduction in flow of 17% (i.e. the water balance bias decreases by 17%), whereas decreasing rainfall by 20% leads to a median reduction in flow of 33%. Conversely, an increase in rainfall of 10% results in flows approximately 19% higher, while increasing rainfall by 20% results in an increase in flow of 37%. The changes in flow are therefore clearly proportionate to the changes in rainfall. There is normally a multiplying effect between change in rainfall and resultant change in flow (often approximately double) due to the complex interaction of catchment processes in the model.

It is apparent from these tests that there is not likely to be any simple, uniform spatial patterns of bias in rainfall inputs, with particular complexity apparent in upland

catchments. This is consistent with the spatial variation in NSE, water balance bias and other metrics described for the standard run, which exhibit some geographical clustering in terms of performance but also notable variability. In aggregate terms, the water balance appears to be approximately correct, which supports the notion that biases for individual catchments could be related to factors other than undercatch or simple systematic rainfall measurement biases, such as PET estimation, poorly constrained abstractions/discharges or locally sub-optimal parameters. However, it should be noted that the level of performance obtained with the standard rainfall inputs could be partly an artefact of the development history of SHETRAN and consequent model parameter choices that compensate for undercatch in rainfall or other systematic errors in inputs. Further exploration of this issue would be interesting but is beyond the scope of this work.

### 5.5.3 Sensitivity to evapotranspiration rates

In addition to analysing the sensitivity to rainfall rates, the sensitivity of the SHETRAN for GB modelling system to daily total potential evapotranspiration was also assessed in a similar way to daily total rainfall inputs. The rate of evaporation can be controlled in two ways in SHETRAN. Firstly by changing the input time series of PET and secondly through the actual evaporation to potential evaporation (AE/PE) ratio, which is a function of land cover. Two tests were conducted where the PET input timeseries was decreased externally to the model by 5% and 10%, with one additional test of the effects of increasing PET by 5%. Four further tests were carried out in which PET was also changed internally by adjusting the actual evaporation to potential evaporation (AE/PE) ratio parameter within SHETRAN. The AE/PE ratio was increased by 0.1 and 0.2 and also decreased by 0.1 and 0.2 (i.e. the AE/PE ratio for grass is 0.6 and it was changed to 0.7, 0.8, 0.5 and 0.4 respectively).

Changing the PET externally to SHETRAN by adjusting the input series gives similar patterns to adjusting the internal AE/PE ratio, with overall improvements in model performance found in tests with higher evapotranspiration and general deterioration in simulation skill when less evapotranspiration takes place, as reflected in Table 5.7 and elaborated below. This pattern is consistent with the results from sensitivity analysis of rainfall inputs, which suggests that errors in rainfall or PET may compensate for each other in some cases. The improvements from applying higher evapotranspiration tend to be greatest in England rather than Scotland and Wales, as the latter two countries have quite high proportions of upland catchments with comparatively high mean

| Metric | Statistic | Standard | -5% | -10% | +5% | +0.1 | +0.2 | -0.1 | -0.2 |
|---|---|---|---|---|---|---|---|---|---|
| NSE | Mean NSE Change | 0.00 | -0.39 | -0.79 | 0.37 | 0.53 | 1.01 | -0.61 | -1.16 |
| | % Improved | 0.0 | 12.1 | 16.4 | 54.0 | 55.2 | 54.3 | 8.5 | 11.6 |
| | Median | 0.69 | 0.68 | 0.66 | 0.69 | 0.7 | 0.7 | 0.66 | 0.64 |
| | Change in Median | 0.00 | -0.01 | -0.03 | 0 | 0.01 | 0.01 | -0.03 | -0.05 |
| | $25^{th}$ Percentile | 0.39 | 0.36 | 0.28 | 0.46 | 0.47 | 0.49 | 0.32 | 0.23 |
| | $75^{th}$ Percentile | 0.79 | 0.78 | 0.77 | 0.79 | 0.79 | 0.79 | 0.78 | 0.76 |
| | % Band 1 | 22.5 | 21.5 | 19.9 | 21.9 | 24.0 | 22.0 | 20.8 | 19.1 |
| | % Band 2 | 40.8 | 39.2 | 38.8 | 42.8 | 41.6 | 46.0 | 38.9 | 35.5 |
| | % Band 3 | 36.6 | 39.2 | 41.4 | 35.3 | 34.4 | 32.0 | 40.4 | 45.4 |
| Water balance | % Improved | 0.0 | 27.9 | 23.0 | 61.9 | 63.8 | 56.5 | 27.4 | 24.8 |
| | Median | 5.1 | 9.3 | 12.8 | 1.9 | 1.5 | -1.1 | 9.1 | 11.4 |
| | Change in Median | 0.0 | 4.2 | 7.7 | -3.2 | -3.6 | -6.2 | 4.0 | 6.3 |
| | $25^{th}$ Percentile | -2.2 | 0.7 | 3.2 | -5.3 | -4.6 | -8.2 | 0.2 | 1.5 |
| | $75^{th}$ Percentile | 18.3 | 23.5 | 29.9 | 13.2 | 13.0 | 9.2 | 24.9 | 30.9 |
| | % Band 1 | 52.3 | 45.7 | 37.8 | 54.8 | 54.2 | 55.0 | 45.3 | 40.8 |
| | % Band 2 | 22.5 | 23.8 | 23.7 | 25.0 | 27.0 | 26.5 | 24.5 | 22.1 |
| | % Band 3 | 25.2 | 30.6 | 38.5 | 20.2 | 18.8 | 18.5 | 30.2 | 37.0 |
| mmfd | % Improved | 0.0 | 29.2 | 23.9 | 60.6 | 63.5 | 59.2 | 26.6 | 22.9 |
| | Median | 42.9 | 45.6 | 50.0 | 41.3 | 39.4 | 38.8 | 45.5 | 50.3 |
| | Change in Median | 0.0 | 2.6 | 7.0 | -1.7 | -3.5 | -4.2 | 2.5 | 7.3 |
| | $25^{th}$ Percentile | 27.2 | 29.1 | 30.7 | 27.5 | 26.1 | 26.6 | 28.9 | 30.4 |
| | $75^{th}$ Percentile | 72.4 | 77.6 | 87.5 | 66.2 | 65.2 | 63.3 | 81.0 | 87.0 |
| | % Band 1 | 12.7 | 11.9 | 10.6 | 12.7 | 13.6 | 13.7 | 12.1 | 10.7 |
| | % Band 2 | 34.3 | 30.9 | 26.9 | 36.6 | 36.7 | 37.4 | 29.8 | 28.6 |
| | % Band 3 | 52.9 | 57.2 | 62.5 | 50.7 | 49.7 | 48.9 | 58.1 | 60.7 |

*Table 5.7: Summary table of performance statistics for increased/decreased PET and increased/decreased AE/PE ratio.*

annual rainfall, which may form a more important control on their overall simulation performance than evapotranspiration.

Decreasing the PET by 5% (10%) results in a mean NSE decrease of 0.4 (0.8) and improves only 12 (16)% of catchments respectively, while the median NSE decreases slightly in both cases. With respect to water balance bias, only 28% (23%) of catchments improve as a result of decreasing PE; these catchments are spread across the country and do not appear to show any consistent spatial patterns. Median mmfd increases in both runs, which confirms that decreasing PET is not beneficial to flow simulation, at least in combination with the rainfall dataset applied. Intuitively, a very similar pattern emerges when the AE/PE ratio for each of the vegetation parameters is decreased by 0.1 (0.2). NSE decreases on average by 0.6 (1.2) and NSE only improves in 9 (12)% of catchments respectively. These improvements again do not show any clear geographical signals. However, Scotland is noticeably dominated by catchments showing no overall change as a result of decreasing the AE/PE ratios, again indicating that those catchments are less sensitive to decreases in PE, which is not unexpected given that PET is relatively low there compared to the rest of the country and rainfall relatively high. The median water balance bias increased from 5% to 9 (11)% through decreasing the AE/PE ratios by 0.1 (0.2), while the median mmfd also increased from 42 in the standard run to 46 (50).

Therefore a decrease in evapotranspiration worsens overall model performance when other inputs and parameters are held constant, and this reinforces the findings of the rainfall sensitivity tests described earlier. Conversely, an increase in PET slightly improves results. Approximately 54% of catchments obtain higher NSE scores as a result of increasing PET input by 5%, and 55 (54)% of catchments improve when the AE/PE ratio is changed by 0.1 (0.2). The mean NSE change is an increase of 0.4 for increasing PET by 5%, 0.5 for increasing AE/PE by 0.1 and 1.0 for increasing AE/PE by 0.2. The median NSE increases slightly across all increased PET tests, as does the $25^{\text{th}}$ percentile, while the lack of change in the $75^{\text{th}}$ percentile shows that the improvements take place largely in those catchments that obtained relatively poor NSE scores in the standard run. This is reinforced by noting that the percentage of catchments in band 3 for NSE reduces with increased PE.

Over all tests, reducing PET shifts the distribution of water balance bias towards more negative bias, resulting in lower median and $75^{\text{th}}$ percentile bias but also lower $25^{\text{th}}$ percentile bias. Increasing PET also improves mmfd slightly, again mainly by improving band 3 catchments. Such changes are indeed not unexpected for many of

the chalk catchments, as an increase in PET will reduce the peaks in the hydrograph and so increase the NSE significantly, although of course this is not necessarily a physically correct reason for the observed improvement.

These results therefore suggest that there is a careful balance to be struck between between rainfall and PE, with simulated hydrographs depending closely on the interplay of both inputs. It appears that some systematic improvements to model performance may be possible if sufficiently robust adjustments to both datasets can be determined on the basis of further analysis. The direction of appropriate adjustments is clearly shown by the sensitivity tests undertaken, with lower rainfall or higher PET required to produce overall improvements across a range of metrics and for different catchment types.

### 5.5.4 Sensitivity to rainfall scaling with elevation

The effects of altitude were not compensated for explicitly in the derivation of the UKCP09 5km gridded dataset and whilst the raingauge data used in the interpolation covers some elevation range, it is biased towards lower elevations, particularly in mountain areas. In order to assess whether or not any signals of error are apparent as a result of elevation-dependency in precipitation, two approaches to compensation for the effects of elevation on precipitation intensity were tested to see if correction might be justified. The first approach was to increase rainfall linearly with elevation by 5% with every 400m elevation for each grid square i.e.

$$FactoredRainfall = (5elevation/40000 + 1)OriginalRainfall$$

The second approach tested was to increase rainfall by 15% for every grid square with an elevation over 150m. In both approaches the multipliers were subjectively chosen, but at this stage the main point of the experiments was to investigate whether compensation for elevation might be a worthwhile line of investigation to follow.

The results in Table 5.8 demonstrate that changing rainfall inputs in both of the ways described above leads to the majority of catchments performing worse than in the standard run. Scaling rainfall linearly with elevation increases NSE in only 15% of catchments, and increasing rain over an elevation threshold increases NSE in only 9% of catchments. The number of catchments with band 1 NSE reduced slightly from 23% to 19% with linearly scaled rainfall but dramatically (down to 9%) with threshold-scaled rainfall. The interquartile range of NSE reduced in both cases also. The median water balance bias increased in both cases (from 5% in the standard run to 10% and

| Metric | Statistic | Standard | Scaled rain | Step rain |
|---|---|---|---|---|
| NSE | Mean NSE Change | 0.00 | -0.28 | -0.53 |
| | % Improved | 0.0 | 14.7 | 9.1 |
| | Median | 0.69 | 0.67 | 0.59 |
| | Change in Median | 0.00 | -0.02 | -0.1 |
| | $25^{th}$ Percentile | 0.39 | 0.33 | 0.16 |
| | $75^{th}$ Percentile | 0.79 | 0.78 | 0.73 |
| | % Band 1 | 22.5 | 19.3 | 8.9 |
| | % Band 2 | 40.8 | 40.9 | 40.7 |
| | % Band 3 | 36.6 | 39.9 | 50.3 |
| Water Balance | % Improved | 0.0 | 24.7 | 11.5 |
| | Median | 5.1 | 10.4 | 19.8 |
| | Change in Median | 0.0 | 5.3 | 14.7 |
| | $25^{th}$ Percentile | -2.2 | 2.3 | 9.5 |
| | $75^{th}$ Percentile | 18.3 | 23.7 | 33.9 |
| | % Band 1 | 52.3 | 43.8 | 22.2 |
| | % Band 2 | 22.5 | 22.9 | 26.8 |
| | % Band 3 | 25.2 | 33.3 | 51.0 |
| mmfd | % Improved | 0.0 | 18.0 | 9.8 |
| | Median | 42.9 | 45.7 | 61.0 |
| | Change in Median | 0.0 | 2.7 | 18.1 |
| | $25^{th}$ Percentile | 27.2 | 29.4 | 40.5 |
| | $75^{th}$ Percentile | 72.4 | 76.9 | 84.5 |
| | % Band 1 | 12.7 | 9.5 | 4.0 |
| | % Band 2 | 34.3 | 30.1 | 19.5 |
| | % Band 3 | 52.9 | 60.5 | 76.5 |

*Table 5.8: Summary table of performance statistics for changes to scaling rainfall linearly and in a stepwise fashion.*

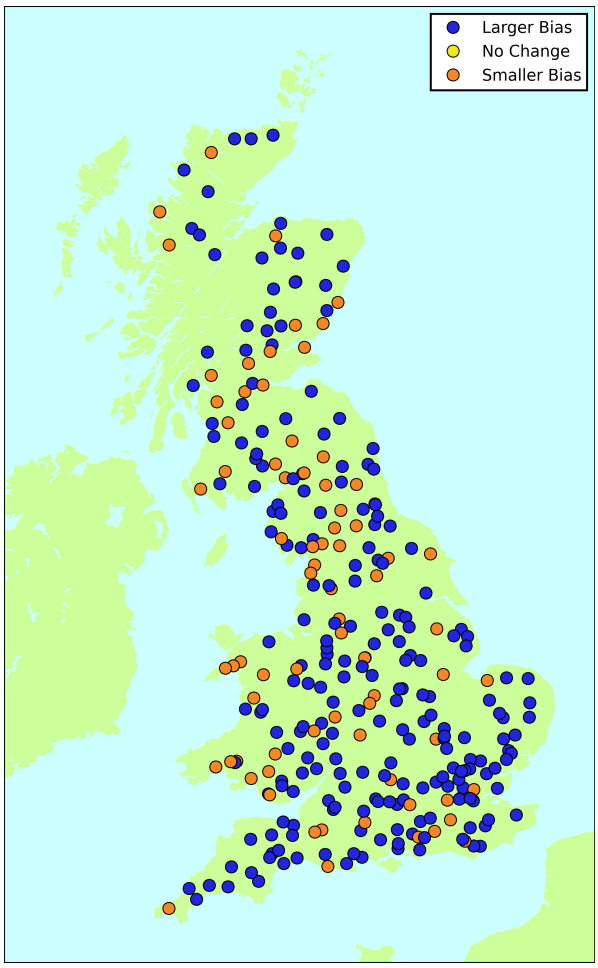



*Figure 5.24: Map showing changes to water balance bias due to increasing the rainfall linearly. Blue indicates a larger bias and orange a smaller bias. The bias increases almost everywhere.*

20% respectively) but more so with the threshold increase in rain, as the net amount of rainfall increase resulting from this change was greater. This was to be expected as the previous set of sensitivity tests suggested that increasing rainfall would not be likely to improve the simulation. Figure 5.24 shows that improvements in the bias are not particularly located in upland areas, and in general there seems to be no regional pattern to improvements. The mmfd statistics also increase across all measures in these tests, confirming deterioriation in the simulation of the flow regime results from adjusting rainfall in these ways.

Overall the results from this set of sensitivity tests show that spatial and altitudinal biases in input rainfall data are not likely to be simple or compensated for using some basic approaches. A fairer test may have been to enhance high altitude rainfall, balanced by reductions in low elevation rainfall to maintain the same average. However, these tests do demonstrate one of the many uses of the SHETRAN for GB system, namely the ability to test hypotheses at a national scale with evaluation based on a large number of diverse catchments. As studies of rain gauge undercatch are normally conducted on a small or catchment scale, this new framework for modelling could allow for testing of new theories and datasets in a systematic and comprehensive way.

### 5.5.5 Uniform temporal distribution of rainfall input

The default setup of SHETRAN is to distribute daily rainfall input equally over each day according to the size of its sub-daily time steps, in effect mimicking a constant drizzle. This is not representative of what happens in reality of course, where sub-daily precipitation distributions may vary considerably depending on underlying mechanisms. The possible consequence of distributing the rainfall in such a uniform way could be biases in simulated flows, as a result of unrealistic dynamics of soil moisture and evapotranspiration compared with more realistic rainfall intensities. A set of sensitivity tests were therefore conducted to investigate the impact of distributing rainfall uniformly over 12, 6 and 3 hours periods at the start of the day instead, in order to investigate the aggregate impacts on catchment response arising from different sub-daily rainfall intensities and durations.

None of the three runs show an improvement in NSE for the majority of catchments (Table 5.9). In fact, each of the tests shows only a relatively small percentage of catchments improving compared with the standard run, by 35%, 26% and 18% for 12, 6 and 3hrs respectively. The median, $25^{th}$ and $75^{th}$ percentile NSE values decrease in all of the tests, the 12hr test by the least and the 3hr test by the most. The percent-

| Metric | Statistic | Standard | 12hrs | 6hrs | 3hrs |
|---|---|---|---|---|---|
| NSE | Mean NSE Change | 0.00 | -0.25 | -1.06 | -1.76 |
| | % Improved | 0.0 | 35.0 | 25.7 | 18.3 |
| | Median | 0.69 | 0.65 | 0.6 | 0.56 |
| | Change in Median | 0.00 | -0.04 | -0.1 | -0.13 |
| | $25^{th}$ Percentile | 0.39 | 0.22 | 0.17 | 0.07 |
| | $75^{th}$ Percentile | 0.79 | 0.76 | 0.73 | 0.72 |
| | % Band 1 | 22.5 | 16.3 | 11.6 | 6.5 |
| | % Band 2 | 40.8 | 42.6 | 38.4 | 40.8 |
| | % Band 3 | 36.6 | 41.1 | 50.0 | 52.7 |
| Water balance | % Improved | 0.0 | 35.7 | 30.6 | 32.1 |
| | Median | 5.1 | 9.7 | 12.1 | 13.1 |
| | Change in Median | 0.0 | 4.6 | 7.0 | 8.0 |
| | $25^{th}$ Percentile | -2.2 | -2.8 | -1.2 | -0.4 |
| | $75^{th}$ Percentile | 18.3 | 24.1 | 27.3 | 29.0 |
| | % Band 1 | 52.3 | 38.9 | 34.4 | 35.5 |
| | % Band 2 | 22.5 | 28.5 | 30.0 | 25.3 |
| | % Band 3 | 25.2 | 32.6 | 35.6 | 39.2 |
| mmfd | % Improved | 0.0 | 30.4 | 29.0 | 26.3 |
| | Median | 42.9 | 46.4 | 47.1 | 51.0 |
| | Change in Median | 0.0 | 3.4 | 4.2 | 8.1 |
| | $25^{th}$ Percentile | 27.2 | 31.0 | 32.2 | 33.3 |
| | $75^{th}$ Percentile | 72.4 | 82.1 | 84.5 | 87.0 |
| | % Band 1 | 12.7 | 7.0 | 7.6 | 6.5 |
| | % Band 2 | 34.3 | 31.9 | 29.2 | 27.8 |
| | % Band 3 | 52.9 | 61.1 | 63.2 | 65.7 |

*Table 5.9: Summary table of performance statistics for 12, 6, and 3 hour rainfall durations.*

age of catchments with band 1 NSE also decrease and the percentage of catchments with band 3 NSE increase across all tests, reflecting a deterioriation in NSE across the sample overall. This pattern of declining performance is also seen in the water balance bias statistics, which show that the water balance bias only improves in around 30% of catchments. The median, $25^{th}$ and $75^{th}$ percentiles of the distribution of water balance biases all shift to higher values, indicating that the model is behaving as expected, i.e. increasing the intensity of the rainfall reduces the influence of the evaporation resulting in more water flowing throught the catchment and entering the channel network for a given event. Yet as the water balance bias was already adequate or slightly overestimating flows in the standard run, the increase in runoff arising from these tests does not improve model performance. The percentage of catchments with biases classified in band 1 decreases and the percentage of catchments with band 3 biases increase as a reflection of this, while all statistics for mmfd also worsen.

Whilst none of these sensitivity tests improve model performance, they do show that the (sub-daily) temporal distribution of rainfall input has an impact on the overall model performance as indicated by NSE and the water balance bias. This suggests that further investigation of this issue may be warranted using more realistic distributions. The issue of generally slightly overestimated runoff from the model over the longer-term indicates that this should perhaps be evaluated with respect to evapotranspiration patterns too.

### 5.5.6 Exponential Distribution of rainfall input

Further investigation of the implications of rainfall distribution below the level of daily model inputs was carried out by assessing another approach, namely exponentially distributing sub-daily rainfall. The reasoning behind this was to see if this approach could at all mimic in very general terms the pattern of a storm event, with the exponential distribution defined by:

$$P(X \leq x) = 1 - e^{-\lambda x}$$

where

$$\lambda = \frac{1}{\bar{x}}$$

Therefore for an hours worth of rainfall:

$$rainfallathourh = dailytotal(e^{\frac{-h}{\bar{x}}}) - e^{\frac{-(h+1)}{\bar{x}}}$$

The mean used depended on the daily total rainfall as in chapter 4.

| Metric | Statistic | Standard | Exp. rain |
|---|---|---|---|
| NSE | Mean NSE Change | 0.00 | -0.05 |
| | % Improved | 0.0 | 49.2 |
| | Median | 0.69 | 0.71 |
| | Change in Median | 0.00 | 0.02 |
| | $25^{th}$ Percentile | 0.39 | 0.38 |
| | $75^{th}$ Percentile | 0.79 | 0.80 |
| | % Band 1 | 22.5 | 25.5 |
| | % Band 2 | 40.8 | 40.1 |
| | % Band 3 | 36.6 | 34.4 |
| Water balance | % Improved | 0.0 | 32.7 |
| | Median | 5.1 | 5.7 |
| | Change in Median | 0.0 | 0.6 |
| | $25^{th}$ Percentile | -2.2 | -1.6 |
| | $75^{th}$ Percentile | 18.3 | 18.8 |
| | % Band 1 | 52.3 | 52.0 |
| | % Band 2 | 22.5 | 22.5 |
| | % Band 3 | 25.2 | 25.5 |
| mmfd | % Improved | 0.0 | 34.3 |
| | Median | 42.9 | 43.1 |
| | Change in Median | 0.0 | 0.1 |
| | $25^{th}$ Percentile | 27.2 | 27.4 |
| | $75^{th}$ Percentile | 72.4 | 73.5 |
| | % Band 1 | 12.7 | 12.6 |
| | % Band 2 | 34.3 | 34.1 |
| | % Band 3 | 52.9 | 53.3 |

*Table 5.10: Summary table of performance statistics for exponentially distributed rainfall.*

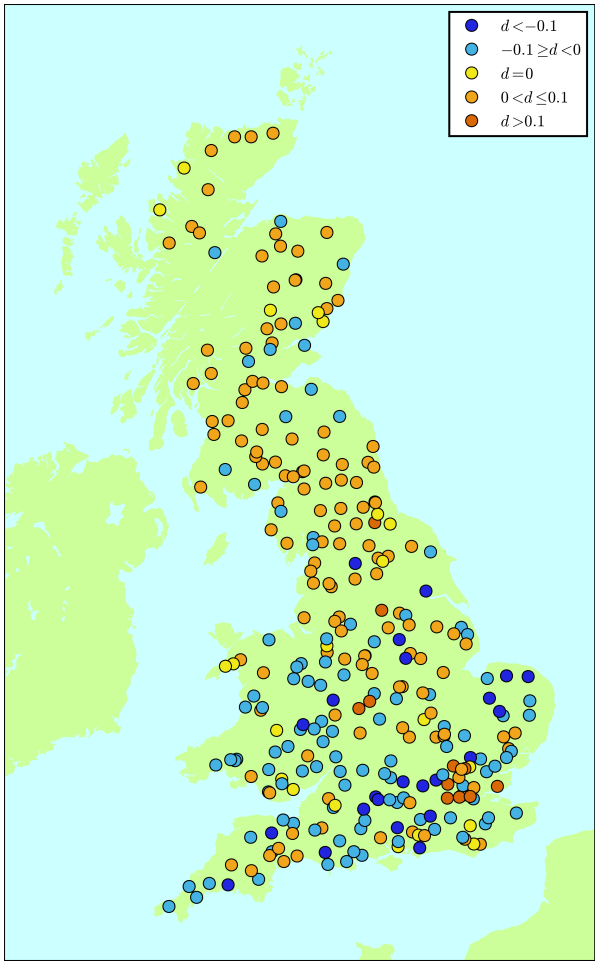



*Figure 5.25: Map showing changes to NSE due to distributing rainfall exponentially. Blue indicates a decrease in NSE and orange an increase. Model performance increases in Northern catchments.*

It can be seen from Table 5.10 that giving the rainfall an exponential distribution prevented the problems encountered with a uniform temporal distribution. NSE improved in 49% of catchments, showing that overall exponentially distributed rainfall had no real detrimental effect. The median NSE improved a small amount from 0.69 to 0.71, while the percentage of catchments in band 1 increased slightly from 23% to 26%. The interquartile range of NSE changed very little, which was also the case for water balance and mmfd. This is most likely because the tail of the distributed rainfall is likely to be available for evaporation in a similar way to the case of 24 hour uniform distribution. As it appears that exponentially distributing rainfall improves peak flows in some cases (reflected where NSE improves) while preserving the initial level of water balance performance, if a simulation requires hourly input of rainfall data without observed data being available, an exponential distribution of daily rainfall may be a good substitute given that it has no detrimental impact.

### 5.5.7 Hourly disaggregation of meteorological input

The natural progression from testing an exponential distribution was to attempt to use real hourly rainfall data as a basis for the sub-daily temporal distribution in SHETRAN. The process for hourly disaggregation of the UKCP09 gridded rainfall dataset was outlined in chapter 3, and it is the resultant hourly gridded dataset which has been applied in the model here. Two tests were then conducted with these hourly rainfall grids. The first test simply involved applying the hourly rainfall grids while retaining the rest of the model setup as per the standard run. A second test was then conducted using an alternative approach to deriving potential evapotranspiration at a sub-daily level.

As for rainfall, by default SHETRAN treats daily potential evapotranspiration as being spread evenly over the course of a day. As an hourly dataset has been created for input, uniformly distributed PET over a day does not represent a similar level of detail to the input rainfall data, which could limit the usefulness of applying a sub-daily rainfall product. The sensitivity of SHETRAN to the sub-daily distribution of PET has not been well studied, such that it was deemed worthy of some investigation here. However, as PET is only estimated, and as hourly wind speed and temperature records were not available to this project, the hourly distribution of PET had to be approximated. The basic approach to the approximation applied in this study was to distribute 50% of the daily PET total according to (the positive part of) a sinusoidal shape beginning at sunrise and ending at sunset, sunrise and sunset times calculated for each grid

square using the Python package "PyEphem" (Rhodes, 2008). The remaining 50% of the daily PET total was distributed uniformly throughout the day. This method is therefore based on the typical diurnal pattern of variation in temperature - albeit without accounting for variation in the diurnal cycle, due to lack of data - but selecting a proportion of total daily PET for this disaggregation is of course subjective and only an initial test at this stage.

Clearly there are several limitations with this method:

- The relative influence of PET from wind vs. T are unknown
- Temperature lags behind radiation
- Other factors affecting PET are completely ignored.

However, this approach is conceptually more realistic than a uniform distribution, providing a more intense PET peak for example. Further work involving the collection of additional datasets or testing the sensitivity of different disaggregation assumptions could reveal a more suitable approach to deriving sub-daily PET inputs, but the rationale of this test is really to see if there is much sensitivity within SHETRAN to the approach adopted to PET variation.

The first test of moving to hourly rainfall based on observations while retaining the rest of the inputs and settings from the standard run does not change the median NSE. Applying hourly rainfall and PET disaggregated using the method described above very slightly increases median NSE from 0.69 to 0.72. However, the $75^{th}$ percentile of the distribution of NSE from all catchments increases from 0.79 to 0.82 in the hourly rainfall test and to 0.84 in the hourly rainfall and PET test, indicating that giving SHETRAN hourly meteorological data improves the simulated hydrographs of those catchments that are already perform well in the standard run. This is reflected in the percentage of band 1 NSE catchments that increases from 23% in the standard run to 32% and 35% in the tests. These changes to the performance statistics seem physically realistic, as giving the rainfall a more accurate distribution should lead to more realistic partitioning of incident precipitation to infiltration and runoff generation processes. This is likely to particularly affect flashier, more responsive catchments, which tend to be well simulated in the standard run and then indeed improve further in terms of NSE. Hourly rainfall distributions generated based on observations may possibly also improve the timing and magnitude of peak flows, which would also be likely to contribute to higher NSE scores. Notably, in the first test just applying hourly rainfall improvements tend to be confined to upland catchments (see Figure 5.26, but

| Metric | Statistic | Standard | Hourly rain | Hourly rain and PET |
|---|---|---|---|---|
| NSE | Mean NSE Change | 0.00 | -0.83 | -0.29 |
| | % Improved | 0.0 | 48.3 | 61.0 |
| | Median | 0.69 | 0.68 | 0.72 |
| | Change in Median | 0.00 | -0.01 | 0.03 |
| | $25^{th}$ Percentile | 0.39 | 0.29 | 0.43 |
| | $75^{th}$ Percentile | 0.79 | 0.82 | 0.84 |
| | % Band 1 | 22.5 | 31.6 | 35.1 |
| | % Band 2 | 40.8 | 31.6 | 30.8 |
| | % Band 3 | 36.6 | 36.8 | 34.1 |
| Water balance | % Improved | 0.0 | 26.8 | 57.6 |
| | Median | 5.1 | 10.4 | 4.3 |
| | Change in Median | 0.0 | 5.3 | -0.8 |
| | $25^{th}$ Percentile | -2.2 | 0.7 | -2.8 |
| | $75^{th}$ Percentile | 18.3 | 26.1 | 18.2 |
| | % Band 1 | 52.3 | 44.0 | 52.9 |
| | % Band 2 | 22.5 | 22.2 | 21.0 |
| | % Band 3 | 25.2 | 33.8 | 26.1 |
| mmfd | % Improved | 0.0 | 20.7 | 34.2 |
| | Median | 42.9 | 46.4 | 42.1 |
| | Change in Median | 0.0 | 3.5 | -0.9 |
| | $25^{th}$ Percentile | 27.2 | 31.0 | 28.8 |
| | $75^{th}$ Percentile | 72.4 | 80.3 | 74.0 |
| | % Band 1 | 12.7 | 11.3 | 12.0 |
| | % Band 2 | 34.3 | 29.7 | 35.1 |
| | % Band 3 | 52.9 | 59.0 | 52.9 |

*Table 5.11: Summary table of performance statistics for disaggregated rainfall and disaggregated PET.*

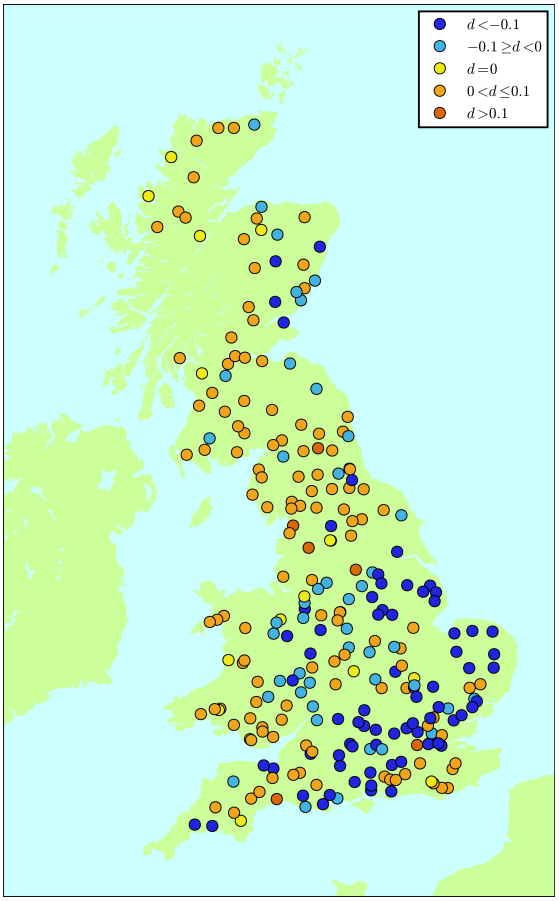



*Figure 5.26: Map showing changes to NSE due to adding hourly rainfall. Blue indicates a decrease in NSE and orange an increase. Model performance increases in most catchments apart from the groundwater dominated ones.*

when applied in conjunction with disaggregated PET there is improvement across the country. Several catchments in regions of chalk do show worse NSE scores as a result of applying hourly rainfall and PET in this test, but as already discussed there are a number of reasons why these areas are not yet adequately simulated in the model.

It is also important to note that, despite general improvements in NSE in the first test when applying hourly rainfall only, increases in water balance bias appear to result from this configuration (the median water balance bias moves from 5% in the standard run to 10% with the incorporation of hourly rainfall). This arises from less evapotranspiration and faster runoff from the more intense rainfall. However, this problem is mitigated by the introduction of hourly PE, where more intense rainfall is compensated for by the more intense evapotranspiration arising from the diurnal cycle (the water balance bias decreases to 4%). Therefore, it is advisable that hourly PET data should be used in SHETRAN if hourly rainfall data are being used, in order to maintain physical consistency between the two inputs. Disaggregating daily rainfall and PET values to an hourly time step appears to be a promising direction that could form part of future work refining model inputs. The timing of precipitation events relative to the presumed diurnal cycle of evapotranspiration could also be a consideration for further research.

## 5.6 Nested catchments

The sensitivity analysis described above was undertaken with respect to all available and suitable gauged catchments in Great Britain, in order to use as broad a range of catchments in the analysis as possible. A significant number of these gauged catchments are in fact sub-catchments of larger basins, which opens up the possibility of investigating the consistency of model performance between catchments and their nested sub-catchment(s). Of the 306 NRFA catchments with freely available flow data that have been modelled in this study, 51% are independent catchments (neither containing a smaller catchment or nested within a larger catchment), 35% are nested catchments (i.e. within another catchment) and 13% are large catchments that contain subcatchments but are not subcatchments themselves. This section therefore evaluates how model performance in terms of NSE for some of the catchments containing the largest number of nested sub-catchments. This provides some insights into variations in process representation between different parts of large catchments that can show significant internal spatial variation in properties and hydrological regimes.

From Table 5.12 it is apparent that some some nested catchments perform as well as

| Station | Catchment NSE | Sub1 | Sub2 | Sub3 | Sub4 | Sub5 | Sub6 | Sub7 | Sub8 |
|---|---|---|---|---|---|---|---|---|---|
| Wye at Redbrook | 0.93 | 0.7 | 0.67 | 0.21 | 0.88 | 0.75 | 0.46 | 0.36 | 0.81 |
| Tyne at Bywell | 0.71 | 0.7 | 0.68 | 0.7 | 0.57 | | | | |
| Kennet at Theale | -2.91 | -11.18 | 0.53 | -14.16 | -12.36 | | | | |
| Lee at Feildes Weir | -0.52 | -22.34 | 0.46 | 0.79 | | | | | |
| Dove at Marston on Dove | 0.82 | 0.79 | 0.78 | 0.76 | | | | | |
| Nith at Friars Carse | 0.84 | 0.83 | 0.78 | 0.84 | | | | | |
| Tees at Broken Scar | 0.58 | 0.7 | 0.59 | | | | | | |
| Don at Doncaster | 0.72 | 0.64 | 0.72 | | | | | | |
| Weaver at Ashbrook | 0.77 | 0.51 | 0.77 | | | | | | |
| Avon at Amesbury | -0.45 | -3.35 | 0.42 | | | | | | |
| Dee at Manley Hall | 0.9 | 0.79 | 0.77 | | | | | | |
| Bedford Ouse at Roxton | 0.82 | 0.86 | 0.8 | | | | | | |

*Table 5.12: Shows that subcatchments can perform consistently or inconsistently with the main catchment.*

other sub-catchments or the overall catchment. This pattern seems to apply to both good and poor NSE values, i.e. all of the sub-catchments tend to either perform well (typically for the upland catchments, e.g. Nith at Friars Carse) or they all perform equally poorly (typical for groundwater dominated catchments, e.g. Kennet at Theale, with one exception discussed below). In these cases a common level of performance stems from the fact that the catchment and its sub-catchments share quite closely their catchment characteristics (as found in the National River Flow Archive), for example, the Nith at Friars Carse and its subcatchments are all underlain by Silurian shales and Boulder Clay and the land use is arable. However, when there is spatial variation in catchment properties, for example in the Lee at Feildes Weir, there is also large variability in model performance judged through NSE scores. In the case of the Lee catchment, one of the smaller subcatchments, Pincey Brook, is predominantly impervious despite being underlain by a chalk aquifer , which therefore has properties allowing for it to be fairly well modelled by SHETRAN (see earlier discussion regarding the types of catchments best simulated by the modelling system). This results in it having a much higher NSE (0.79) compared with the other sub-catchments and the basin overall. Similarly, the Kennet at Theale (Figure 5.27), a mainly pervious catchment of chalk, contains the Enborne at Brimpton sub-catchment (Figure 5.28) in the lower reaches of the Kennet catchment, which are mainly impervious as the chalk is overlain by Tertiary clays. The impervious sub-catchment therefore performs significantly better than the larger groundwater-dominated catchment of the Kennet overall.

The Wye at Redbrook (Figure 5.29) appears to show some variation in model per-

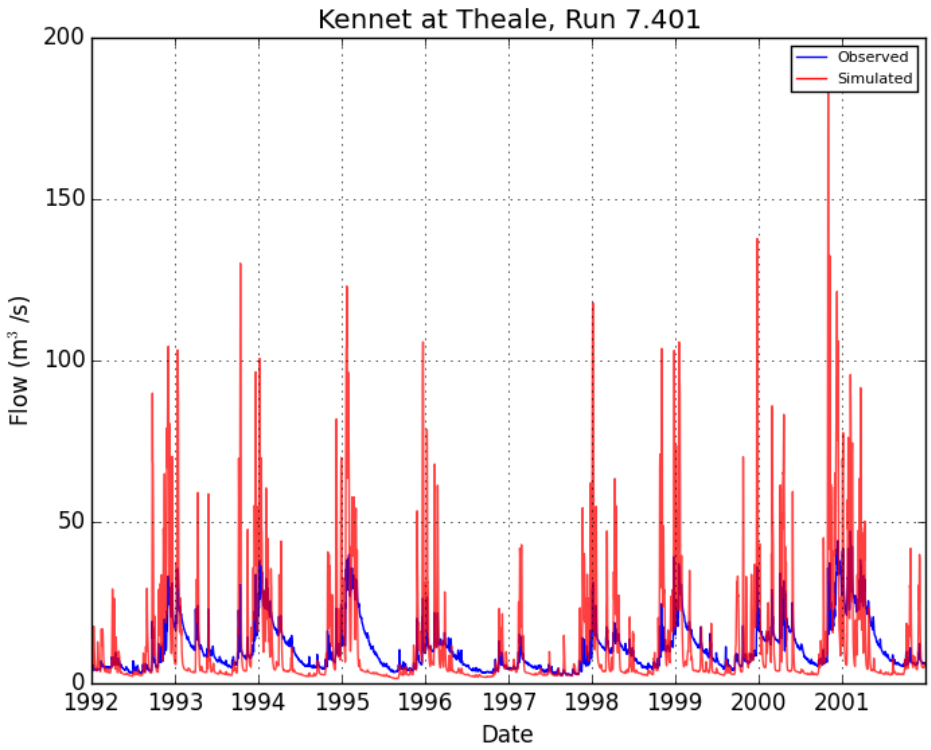



*Figure 5.27: A large groundwater dominated catchment performs poorly.*

formance between its 8 sub-catchments. As it is a large catchment underlain by impermeable geology, it tends to perform very well as a SHETRAN model at its most downstream gauging point at Redbrook. However, its subcatchment the Lugg at Byton (Figure 5.30) is covered by extensive alluvial gravel deposits in the valleys, which act as an aquifer and provide significant baseflow to this tributary. This superficial deposit aquifer is not fully taken into account in the SHETRAN model and so the hydrograph is poorly simulated 5.30, hence a lower NSE score. In conjunction with the other examples, this confirms that location alone is not a good predictor of model performance; catchment characteristics are a more powerful indicator. As discussed earlier, this highlights the issue that certain types of catchment - largely responsive catchments with a predominantly surface water regime - are more consistently well simulated by the SHETRAN modelling system than catchments containing significant aquifers and groundwater-surface water interactions.

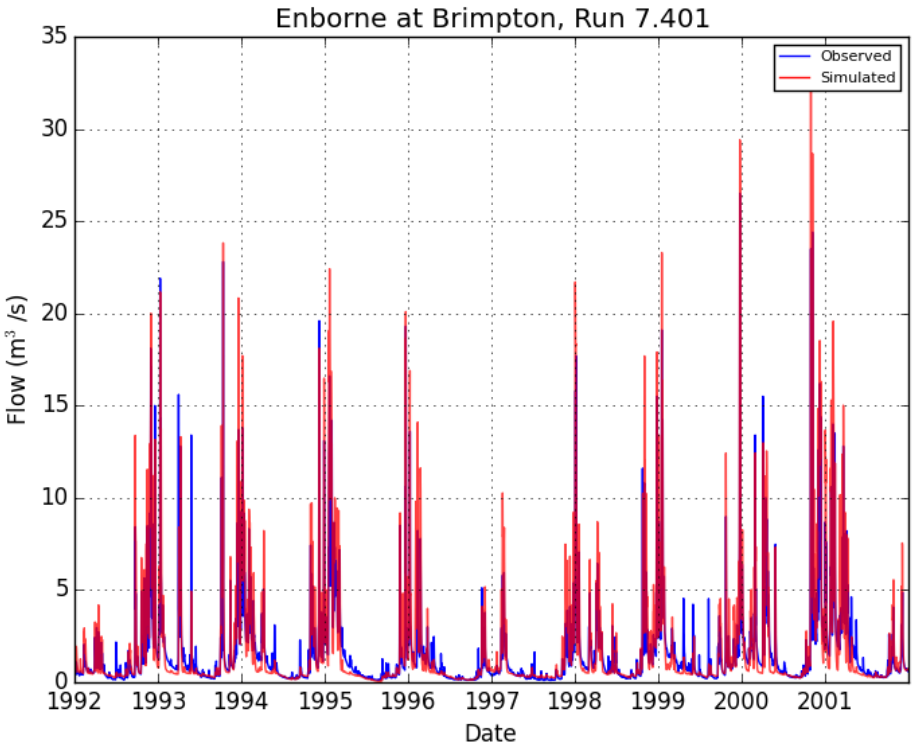



*Figure 5.28: A sub catchment of the Kennet with better model performance because the chalk is overlain with impermeable clays.*

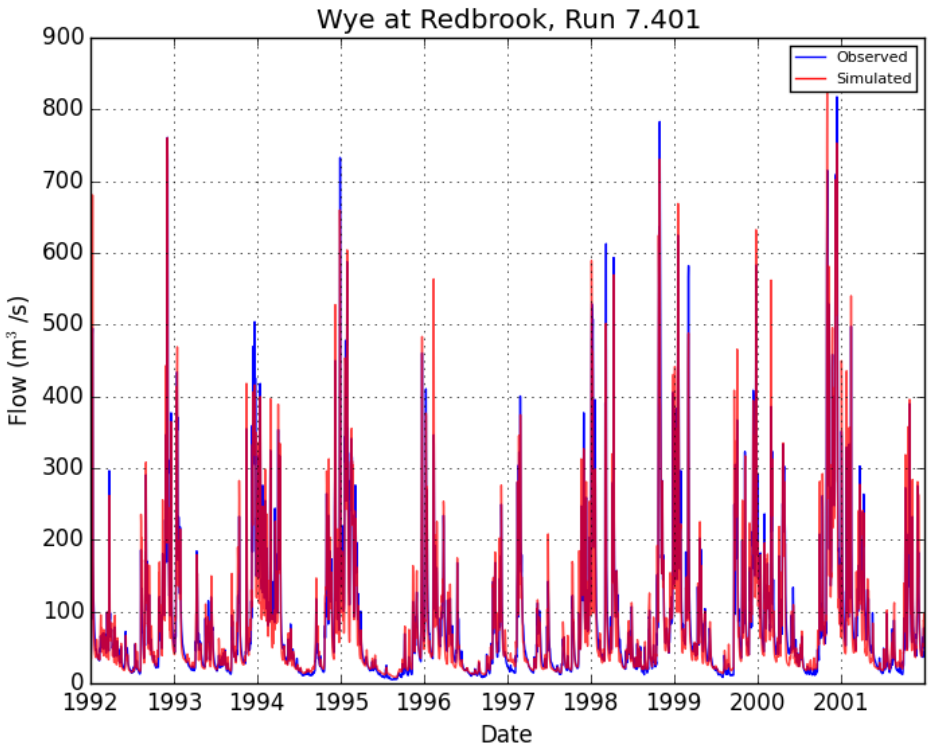



*Figure 5.29: An impermeable large catchment with good model performance.*

## 5.7 The relationship between catchment characteristics and model performance

In order to further characterise the controls on variation in model performance across the simulated catchments, this section analyses some of the ways in which the accuracy of simulated flows may be related to catchment properties using a statistical approach. There is of course a significant body of published work on relationships between runoff and catchment characteristics (also known as descriptors) (Institute of Hydrology, 1999; Singh et al., 2014; Ali et al., 2012), which shows that a large number of possible indices of catchment properties can be defined (area, base flow index, steepness, annual average rainfall etc.). However, many of these catchment characteristics may be correlated with each other, as is certainly the case in the UK owing to spatial patterns and correlations of catchment properties and different hydrological regimes. For example, catchments underlain by chalk in the south-east of England are typically flat, permeable, relatively warm and dry, conditions which may contrasted with those found in steep, impermeable, cool and wet upland catchments in areas of Wales, northern England and Scotland. Commonly applied catchment descriptors such as indices of slope, permeability and

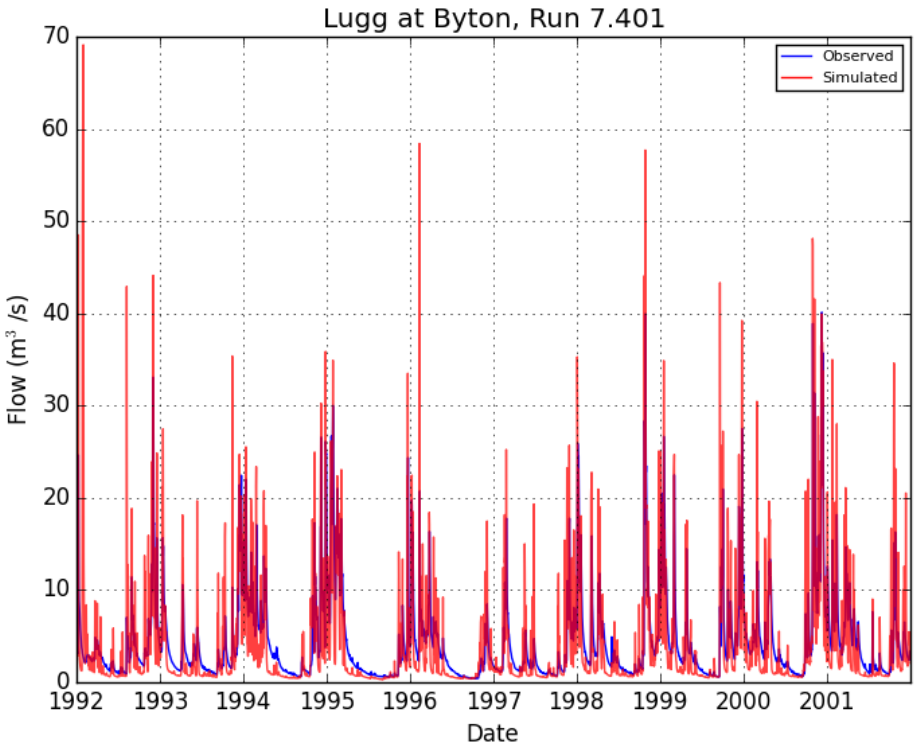



*Figure 5.30: A sub catchment of the Kennet, which performs poorly because the impervious bedrock is overlain with gravels which act as an aquifer.*

| Cluster | Annual Average Rainfall (mm) | Annual Average PET (mm) | BFI | DPSBAR | Mean Flow (mm/day) |
|---|---|---|---|---|---|
| 1 | 657 | 607 | 0.40 | 13.3 | 0.8 |
| 2 | 700 | 590 | 0.66 | 19.6 | 0.6 |
| 3 | 1181 | 507 | 0.37 | 53.4 | 2.1 |
| 4 | 2313 | 514 | 0.3 | 132 | 5.5 |
| 5 | 1749 | 515 | 0.41 | 85.1 | 3.8 |
| 6 | 1152 | 565 | 0.54 | 55.6 | 1.6 |
| 7 | 823 | 598 | 0.89 | 32.6 | 0.8 |

*Table 5.13: Properties of the centre points of the clusters.*

annual mean rainfall will therefore be related to each other in these examples.

The relationships between catchment characteristics and their resultant lack of independence from each other complicate interpretation of links between catchment characteristics and flow, as well as the extent to which model performance can be predicted based on catchment characteristics. This is particularly the case for certain statistical methods, such as multiple regression. As such, instead of trying to predict model performance from all characteristics as a means of more quantitatively characterising variation in simulation accuracy, a clustering algorithm was applied to group catchments based on a small number of basic hydrological characteristics. This was conducted to see whether catchments with similar features could be seen to perform similarly in a statistical test, as suggested by the geographical patterns and interpretations presented earlier in the chapter. The characteristics chosen were the basic properties of annual average rainfall, annual average PE, baseflow index (BFI), DPSBAR (a measure of catchment steepness) and mean flow (normalised by area from cumecs to mm/day) (National River Flow Archive, 2014b; Institute of Hydrology, 1999).

The affinity propagation (AP) algorithm (Frey and Dueck, 2007) implemented in the Python module scikit (Pedregosa et al., 2011) was used to cluster the catchments based on the characteristics outlined above. AP is a fast and efficient way for identifying clusters in data that works by treating each data point as a possible exemplar (cluster centre/archetype) and recursively calculating the similarity (squared error) of this exemplar to the other points (Frey and Dueck, 2007). The scikit AP function requires two parameters, 'damping' and 'preference', which were set to 0.5 and -35 respectively, chosen by trial and error to give a small set of clusters. Applying the algorithm resulted in definition of 7 clusters. Table 5.13 outlines the properties of each cluster exemplar, and their relationships to catchment descriptors may be seen in Figure 5.31. The characteristics of each of the simulated catchments are plotted against each other in this figure, with each catchment coloured according to the cluster to which it belongs in the

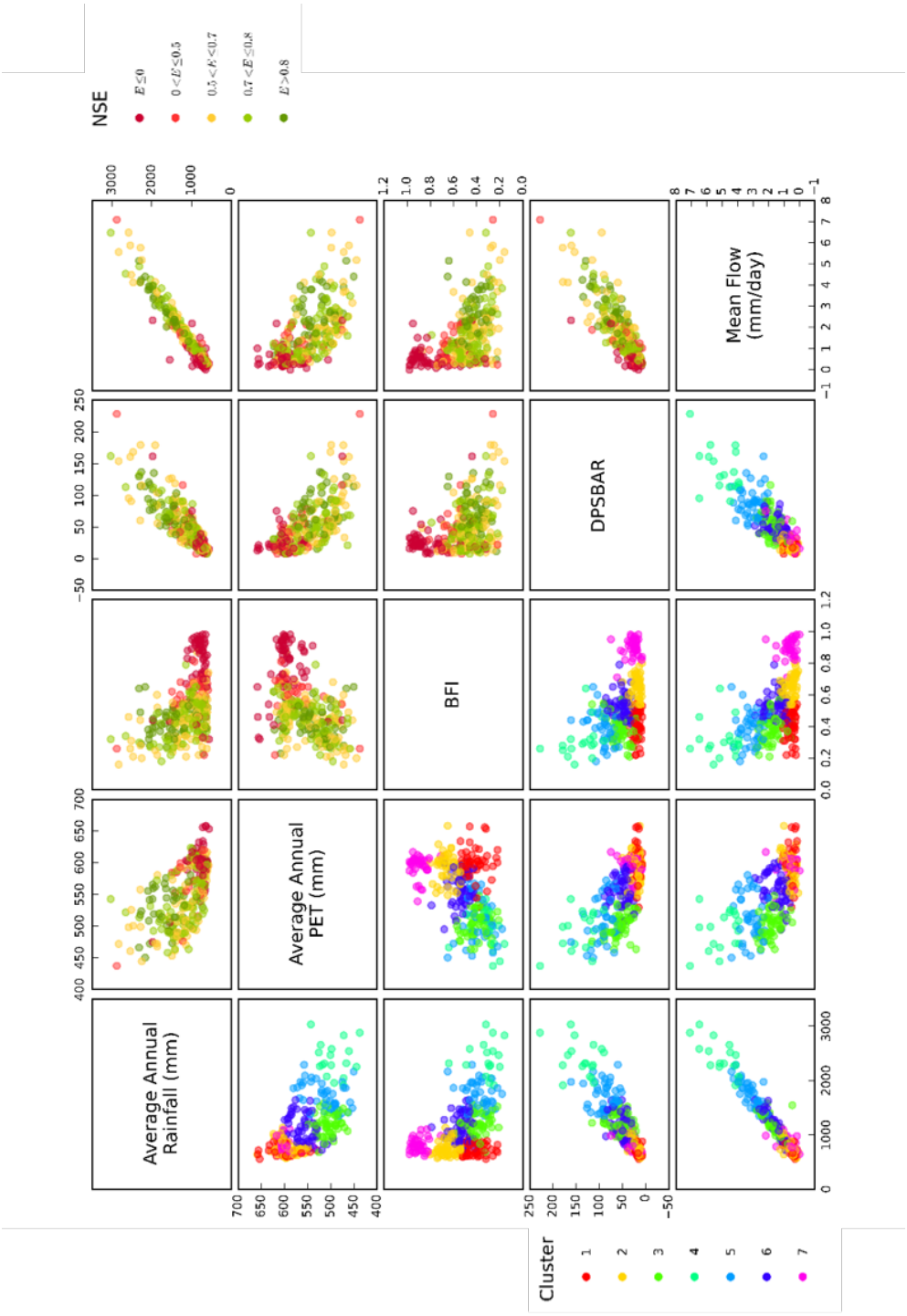



*Figure 5.31: A matrix of catchment characteristics. Each dot represents a modelled catchment. In the lower triangle the catchments are coloured according to the cluster they belong to. In the upper triangle the catchments are coloured by NSE. cluster 7 is composed of entirely of catchments with low NSE.*

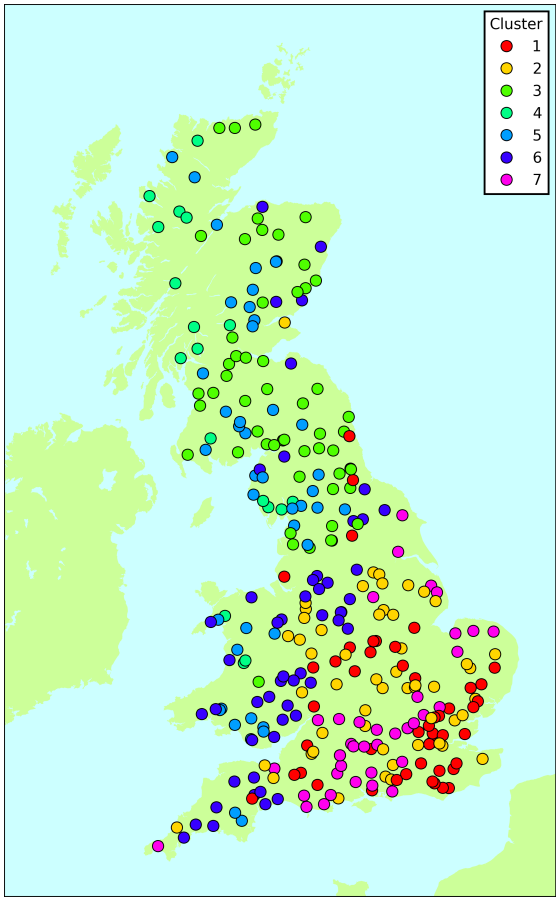



*Figure 5.32: The clusters have been plotted by gauge location. Given the layout of catchments in GB, i.e. that catchments of similar properties are located close together, the clusters also mainly follow a spatial pattern.*

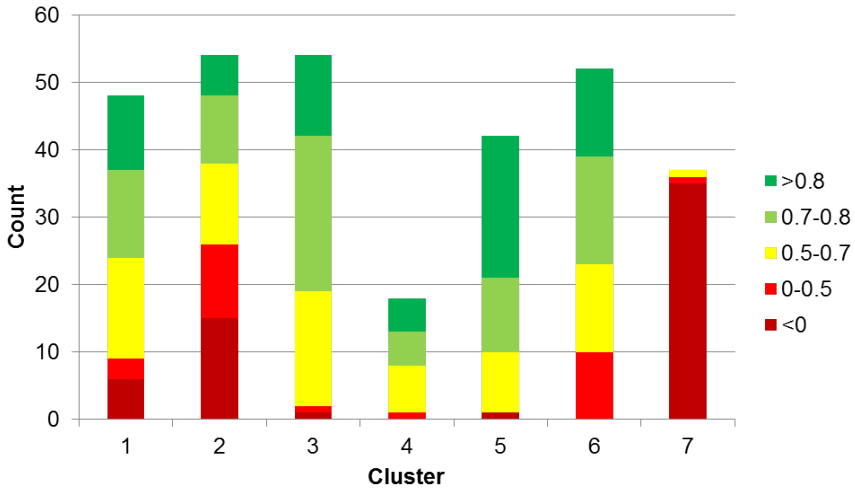



*Figure 5.33: Bar chart showing the composition of each cluster by model performance. Clusters 7 and 2 are mainly composed of catchments with low NSE.*

lower half of the figure (i.e. below the main diagonal of the scatterplot matrix). The upper half of the figure also shows the relationships between characteristics for each catchment, but in this case the points are coloured to NSE, which is discussed further below.

From Table 5.13 and Figure 5.31 it can be seen that cluster 1 is representative of dry, high PET, flat catchments with typically low flows. Cluster 2 contains catchments subject to low rainfall, high PET, moderate BFI and low flows, whereas cluster 3 has more moderate rainfall, fairly low PET, BFI and DPSBAR but moderate flows. Containing steep catchments with the highest mean annual flows, cluster 4 is not unexpectedly also characterised by having the highest rainfall and some of the lowest PET values, as well as the lowest BFI. Cluster 5 is generally similar to clusters 3 and 4, but it does have lower rainfall and contains typically shallower catchments, resulting in lower flows compared with those two clusters. Cluster 6 has moderate rainfall, PET, BFI, slope and flow, whereas cluster 7 is set apart by its very high BFI, which are combined with low rainfall, high PET and shallow gradients to result in low flows.

Plotting the geographical distribution of the clusters in Figure 5.32 shows that clear spatial patterns are associated with the results of the clustering algorithm. Clusters 7 and 2 can be seen to be largely composed of groundwater-fed catchments, particularly in chalk regions in the south of England that are comparatively flat and dry. Figure 5.32 demonstrates that clusters 3, 4 and 5 include steep, impermeable upland catchments, while cluster 6 is mainly composed of catchments in Wales and the south-west of England, but also a few northern and Scottish catchments (particularly the drier ones on the east coast). Cluster 1 is formed mainly of catchments in the very south-east of England, but also contains some catchments located in the Midlands.

In conjunction with earlier figures showing maps of catchment performance, Figure 5.32 indicates that there is a notable north-west/south-east divide across Great Britain in terms of both catchment properties and model performance. This is additionally reflected in Figure 5.33, which shows the composition of each cluster according to the number of catchments in different NSE bands. As expected, cluster 7 is composed almost entirely of catchments with low NSE scores, due to the fact that it is dominated by catchments in poorly simulated groundwater-dominated regions of England. Cluster 2 shows a roughly even split between low and high NSE. Most of the rest of the clusters consist of catchments with generally good model performance, although none of the clusters are associated with exclusively good NSE scores. This reflects the point that the modelling system does not necessarily work well for every catchment exhibiting

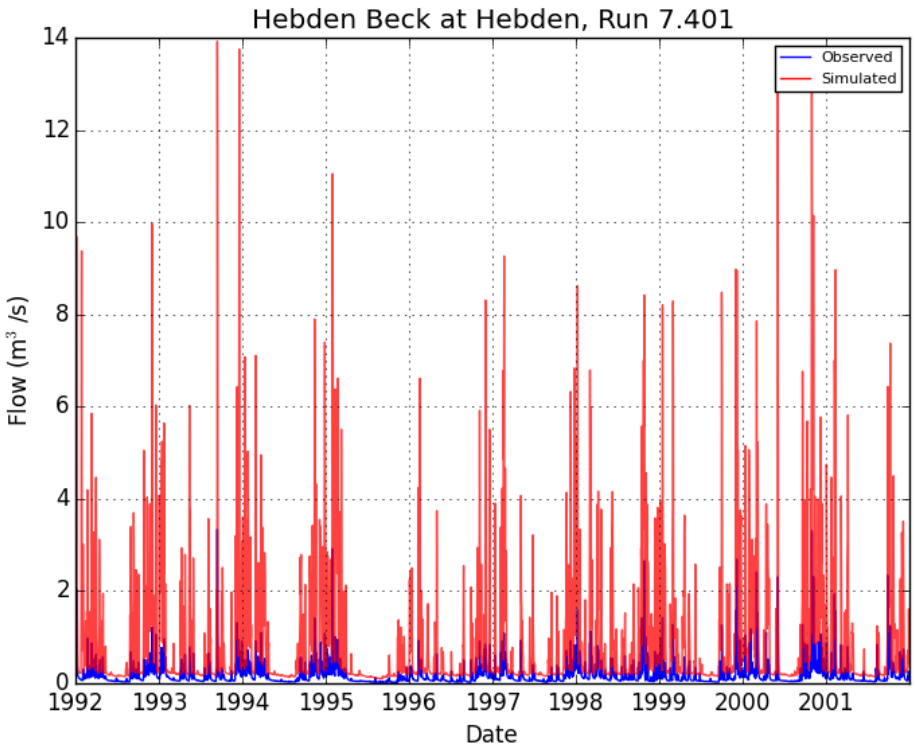



*Figure 5.34: Hydrograph for Hebden Beck at Hebden (NSE -34,4). Simulated flow is over-predicted as the karstic limestone of the catchment is not properly represented in the model.*

properties suggesting that it should be adequately simulated. There can always be some issue specific to a catchment complicating its representation that is not fully captured in the broad-scale characteristics used to describe it in this clustering approach. This is particularly the case when working with relatively coarse resolution national datasets. For example, there are a number of important features of catchment hydrology that cannot yet be readily incorporated into the modelling system due to data limitations, such as detailed descriptions of superficial deposits, abstractions and discharges, flow regulation, river diversions and limitations of flow measurements amongst other factors (Coxon et al., 2015).

Examples of the effects of some of these omissions can be seen in clusters 3, 4, 5 and 6, which are all comparatively 'good' clusters that each contain some poorly performing catchments. Taking group 3 as an example, the worst performing model appears to be for Hebden Beck at Hebden, which has an NSE of -34.4 (see Figure 5.34). Investigating this catchment further by reviewing the NRFA metadata reveals that the problems in simulating this catchment are likely to stem from the fact that it is actually a karstic limestone catchment with significant sink holes and complicated subsurface flow pat-

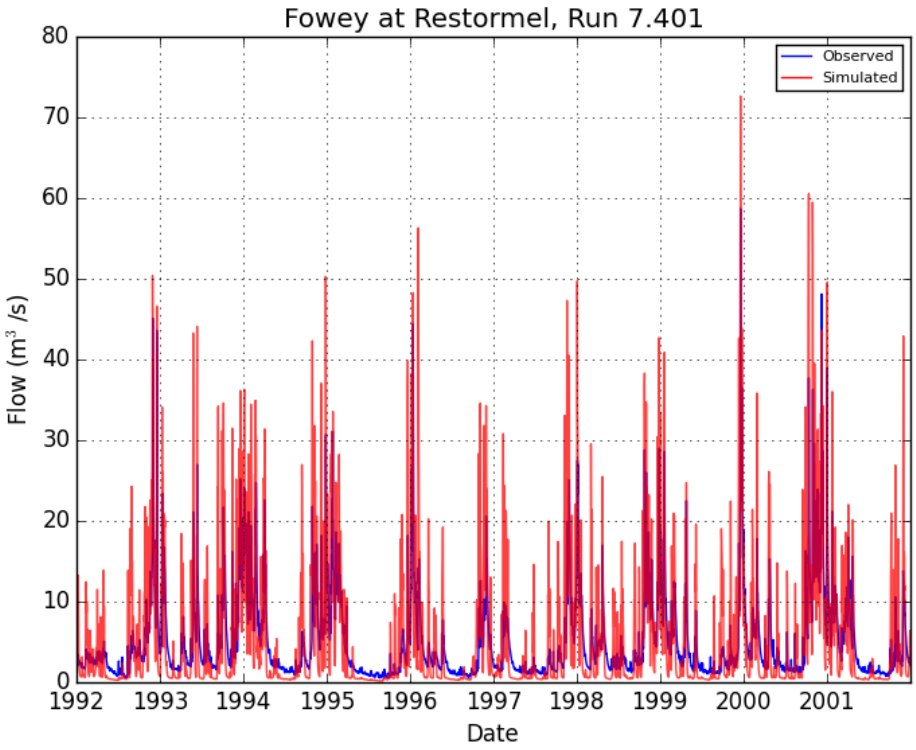



*Figure 5.35: Hydrograph for Fowey at Restormel (NSE 0.08). The baseflow nature of the observed flows is not recreated in the simulation as the gravels in the lower reaches of the valley are not modelled.*

terns. The BGS hydrogeology map used as input for the study marks this catchment as overlying a moderately productive aquifer. Clearly this level of information is insufficient to accurately approximate the physical processes occurring in this context. Together with the issues in simulating catchments underlain by chalk aquifers, examples such as this add weight to the need to incorporate more accurate geological information into the modelling system.

In group 4 the worst performing catchment with respect to NSE is the Nevis at Claggan (NSE of 0.47), which is a relatively well-performing catchment to be bottom of a cluster. Worse performance is found in group 5 for the Lyon at Comrie Bridge, which has an NSE of -0.06 (see Figure 5.19). As discussed above, it is likely that the poor performance in this case is due to diversions in the flow for hydropower generation, something for which adequate data were unavailable for this project. The worst catchment in group 6 is the Fowey at Restormel, with an NSE of 0.08 (see Figure 5.35). The hydrograph for this catchment in Cornwall appears to be more baseflow-dominated in reality than in the model simulations. This is because there is storage from gravel deposits in the lower reaches of the catchment, which again is not included in the current model set up and highlights the potential usefulness of more accurate data for superficial and solid geology. There is also flow regulation within the catchment, while abstraction from groundwater and surface water may also affect the catchment water balance, although these effects cannot be quantified without additional incorporation of data on abstractions and discharges. These cases all demonstrate the existence of peculiarities that can confound the basic setup of the SHETRAN system, but in each case there appears to be clear potential for improving simulations if more and/or better quality data can be included in the model.

The cluster analysis therefore demonstrates that there are some clear relationships between fundamental catchment properties and flow regimes can be defined. However, the variation in NSE scores within clusters - which is more of an issue for some clusters than others - indicates that the relationships are not necessarily simple or universal. Indeed, analysing some specific examples shows that various local factors complicate the nature of the relationships between catchment descriptors and model performance using the SHETRAN system. These factors can come from different sources, such as currently inadequate descriptions of superficial deposits or insufficient data on abstractions and diversion. It appears that many of these issues could be rectified with more and/or better quality data. These results would therefore suggest that attempting to calibrate parameters of the modelling system across all catchments on the basis of catchment descriptors could be difficult at this stage at least, due to the potential for parameter

estimation to be confounded or biased by significant local factors. Careful selection of a smaller sample of catchments to use in such a process could be possible, but the analysis above has demonstrated that a relatively high degree of performance can be obtained by using universal parameters. The main catchments for which this approach appears to fall down are more likely to be simulated poorly for other reasons, such as insufficient data on geological structure and properties in chalk regions. Parameters could of course be tuned further and local influences accounted for in specific applications in individual (or groups of) catchments, but the cluster analysis demonstrates that the modelling system generally provides a good baseline level of performance for the majority of catchment types in Great Britain.

## 5.8 Comparison with other models

To put the performance of SHETRAN for GB into the wider context of other national modelling studies for the UK, the results of the standard run were assessed against two calibrated conceptual model studies: Deckers et al. (2010) using HBV and Bell et al. (2009) using Grid-To-Grid (G2G) (see Chapter 2 for discussion of these models). The first of these studies involved calibration of HBV for 48 catchments across England and Wales over the period September 1983 to December 1990 using Monte Carlo simulation. This is a slightly different model evaluation period compared with that used in this chapter, but of course an comparison indicating overall similarities and differences can still be undertaken. The objective functions used in the calibration by Deckers et al. (2010) were relative volume error and NSE, the latter being calculated separately for all flows, high flows (Q5) and low flows (Q90).

| Catchment | HBV NSE | SHETRAN NSE |
|---|---|---|
| Aire at Kildwick Bridge | 0.82 | 0.89 |
| Arrow at Titley Mill | 0.84 | 0.67 |
| Bedburn Beck at Bedburn | 0.61 | 0.69 |
| Blyth at Hartford Bridge | 0.63 | 0.83 |
| Box at Polstead | 0.69 | 0.03 |
| Boyd at Bitton | 0.8 | 0.83 |
| Brue at Lovington | 0.78 | 0.84 |
| Chater at Fosters Bridge | 0.75 | 0.81 |
| Cheriton Stream at Sewards Bridge | 0.47 | -16.22 |
| Coquet at Morwick | 0.65 | 0.79 |
| Dee at New Inn | 0.73 | 0.77 |
| Dove at Kirkby Mills | 0.67 | 0.72 |
| Dove at Rocester Weir | 0.75 | 0.79 |
| Fal at Tregony | 0.87 | 0.16 |
| Greta at Rutherford Bridge | 0.71 | 0.7 |
| Gwash South Arm at Manton | 0.67 | 0.52 |
| Gwili at Glangwili | 0.85 | 0.81 |
| Hayle at St Erth | 0.86 | -0.96 |
| Ise Brook at Harrowden Old Mill | 0.65 | 0.76 |
| Leven at Leven Bridge | 0.69 | 0.86 |
| Lod at Halfway Bridge | 0.76 | 0.74 |
| Lugg at Byton | 0.89 | 0.21 |
| Nadder at Wilton | 0.8 | 0.28 |
| Otter at Dotton | 0.78 | 0.64 |
| Roden at Rodington | 0.79 | 0.12 |
| South Tyne at Featherstone | 0.72 | 0.68 |
| Stringside at Whitebridge | 0.7 | -4.82 |
| Teme at Knightsford Bridge | 0.82 | 0.8 |
| Thrushel at Tinhay | 0.85 | 0.57 |
| Tywi at Nantgaredig | 0.76 | 0.85 |
| Ure at Kilgram Bridge | 0.75 | 0.81 |
| Wellow Brook at Wellow | 0.88 | 0.78 |

*Table 5.14: Comparison of HBV model performance with SHETRAN for GB.*

Table 5.14 shows the NSE values from the HBV study compared with those from SHETRAN for GB. Out of the 32 catchments that are simulated in both studies, SHETRAN performs better with respect to NSE for 14 catchments. A clear pattern is seen in Figure 5.36, in which catchments associated with good NSE values (>0.5) in SHETRAN for GB also tend to be well simulated in Deckers et al. (2010)'s HBV study. However, SHETRAN performs much worse than HBV for catchments with poor NSE values (<0.5) in both studies. From the discussion above regarding catchment characteristics, this is not particularly unexpected. This stems from the issue that the poorly performing catchments in the SHETRAN system are often those that are

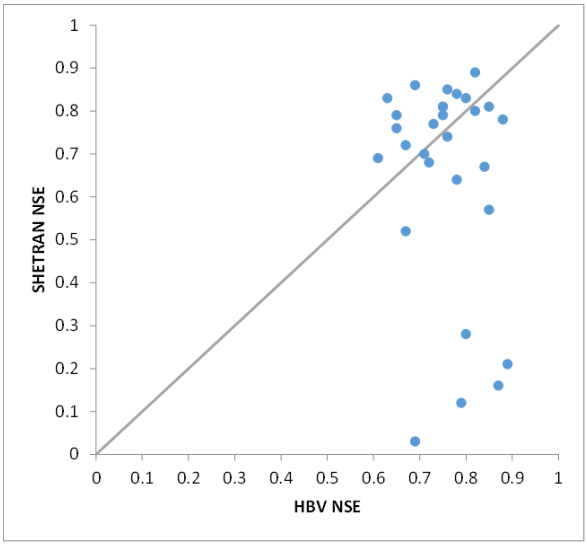



*Figure 5.36: Comparison of NSE values from the calibrated HBV model and SHETRAN for GB.*

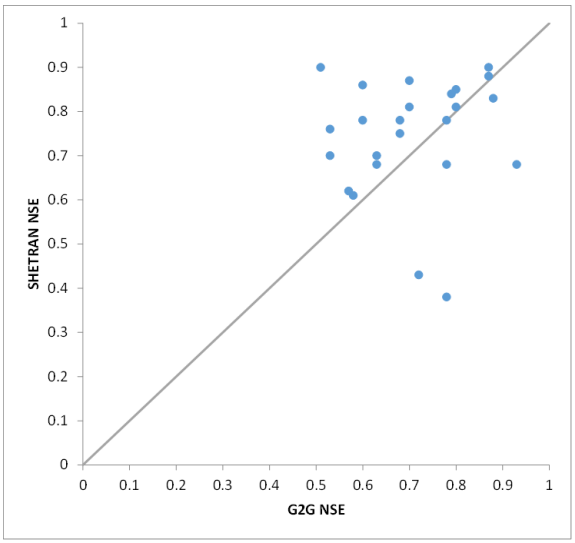



*Figure 5.37: Comparison of approximate NSE values from the calibrated G2G model and SHETRAN for GB.*

| Name | G2G NSE (approx) | SHETRAN NSE |
|---|---|---|
| Beult at Stile Bridge | 0.7 | 0.87 |
| Greta at Rutherford Bridge | 0.63 | 0.7 |
| Crimple at Burn Bridge | 0.58 | 0.61 |
| Lune at Caton | 0.78 | 0.78 |
| Ewe at Poolewe | 0.51 | 0.9 |
| Tawe at Ynystanglws | 0.8 | 0.81 |
| East Dart at Bellever | 0.78 | 0.68 |
| Wye at Cefn Brwyn | 0.53 | 0.7 |
| Falloch at Glen Falloch | 0.63 | 0.68 |
| Mole at Kinnersley Manor | 0.68 | 0.78 |
| Leet Water at Coldstream | 0.6 | 0.78 |
| Cynon at Abercynon | 0.88 | 0.83 |
| Ruchill Water at Cultybraggan | 0.68 | 0.75 |
| Leven at Leven Bridge | 0.6 | 0.86 |
| Nith at Friars Carse | 0.79 | 0.84 |
| Yscir at Pontaryscir | 0.8 | 0.85 |
| Taff at Pontypridd | 0.87 | 0.88 |
| Exe at Thorverton | 0.93 | 0.68 |
| Colne at Lexden | 0.7 | 0.81 |
| Dee at Manley Hall | 0.87 | 0.90 |
| Dove at Izaak Walton | 0.53 | 0.76 |
| Blackwater at Swallowfield | 0.78 | 0.38 |
| Great Stour at Horton | 0.72 | 0.43 |
| Dun at Hungerford | 0.7 | -14.16 |
| Little Ouse at Abbey Heath | 0.56 | -0.68 |
| Frome at Ebley Mill | 0.57 | 0.62 |
| Mimram at Panshanger Park | -0.5 | -22.34 |
| Lambourn at Shaw | 0.69 | -11.18 |

*Table 5.15: Comparison of G2G model performance with SHETRAN for GB.*

groundwater-dominated. This is largely due to the fact that the parameters associated with the highly productive aquifer bedrock type in SHETRAN for GB are not likely to be suitable. The Stringside at Whitebridge, Hayle at St Erth, Cheriton Stream at Sewards Bridge and Box at Polstead all perform poorly in SHETRAN as a result, whereas the catchment models in Deckers et al. (2010)'s study have been explicitly calibrated to give good model performance.

A similar pattern is seen when comparing SHETRAN for GB to the national G2G model (see Table 5.15). G2G was calibrated on the period 28 November 1980 to 18 December 1982 and the results above are from simulations of 1 January 1985 to 31 December 1993. SHETRAN outperforms G2G on 19 out of 28 catchments that are common to both studies. In addition, Figure 5.37 shows that SHETRAN generally gives higher NSE values than G2G for catchments with good performance (NSE>0.5) in both studies. However, SHETRAN for GB tends to perform worse than G2G for comparatively poorly simulated catchments in both studies for exactly the same reasons given above - a calibrated conceptual model simulates groundwater catchments better than an incorrectly parameterised, physically-based one.

Results from other national studies are not directly comparable, as different catchments have been modelled, other measures of performance have been used or individual catchment results have not been reported. Henriksen et al. (2003) give NSE values for 28 gauged catchments in their national MIKE SHE model for Denmark, with 75% of these catchments showing NSE>0.5. This is quite a high percentage, but of course the sample of catchments is much smaller relative to those modelled in SHETRAN for GB. Furthermore, explicit calibration was undertaken by Henriksen et al. (2003), which is likely to be aided by the comparative hydrological homogeneity of their study area. In a separate study, Crooks et al. (2014) present results for 54 catchments in Great Britain modelled using the calibrated conceptual model CLASSIC-GB. Only 3 of these catchments have an NSE value of less than 0.5, which again reflects the substantial benefits of calibration for faithfully reproducing observed hydrographs. The French national model (Habets et al., 2008) showed good results for 610 gauged catchments over a 10-year simulation, with 66% of catchments showing NSE greater than 0.55 and 36% showing NSE greater than 0.65. These results are similar but a little lower than those from SHETRAN for GB, and it is interesting to note that neither modelling system utilised parameter calibration during their respective development processes.

The standard configuration of SHETRAN for GB produces acceptable results across much of the country, satisfactorily (NSE>0.5) simulating 72% of catchments and pro-

viding good simulations (NSE>0.7) for 48%. The standard configuration of SHETRAN for GB therefore performs at a comparable - and in a number of cases favourable - level with respect to results from other national modelling systems. The comparisons confirm that SHETRAN for GB simulates a large proportion of catchments well, indeed better than calibrated conceptual models for many of the best modelled catchments. However, as recognised earlier in the chapter, limitations in subsurface representation lead to comparatively poor simulations in catchments with significant groundwater systems. Calibration in other modelling studies reveals the clear potential to circumvent the issue of particularly poorly simulated catchments through refining parameter estimates. However, it should also be noted that this evaluation is based primarily on NSE, due to its common use in other national modelling studies, but NSE is an imperfect descriptor of model performance, as discussed above.

## 5.9 Conclusion

Extensive analysis of the performance of the national modelling system and its sensitivity to several of the most important input datasets and parameters has been conducted. The initial run undertaken revealed acceptable model performance (NSE>0.5) in the majority of catchments, with some exceptions located generally in parts of Scotland and the south-east of England. Structural improvements were then made to the SHETRAN system to include snow melt processes relevant particularly in upland regions, as well as better representation of lakes and handling of sinks in the input DEM. These modifications resulted in clear improvements leading to definition of the standard configuration of the national SHETRAN modelling system for Great Britain, against which subsequent sensitivity tests and scenario runs are analysed. The standard configuration produces acceptable results across much of the country. 72% of catchments are satisfactorily modelled (NSE>0.5), with 48% well simulated (NSE>0.7). This is comparable to other national modelling systems and UK studies using calibrated conceptual models. The major exception to this level of performance appears to be found in some catchments in the south of England, particularly those underlain by significant aquifers and so where hydrological regimes are highly influenced by subsurface processes and their interaction with surface hydrology. Some of the limitations of currently available national datasets with respect to capturing important features of local hydrological cycles are also apparent from the sensitivity analysis, important examples of which include insufficient data on superficial deposits and artificial influences, such as abstractions or flow regulation.

The relationship between catchment characteristics and model performance was examined using the affinity propagation clustering technique. Catchments were clustered based on 5 key characteristics: average rainfall, average PET, base flow index, steepness and mean flow. The resulting 7 clusters contained catchments that are typically spatially concentrated, which is due to the distinct meteorological, topographical and geological patterns in GB. Model performance of each catchment was shown to be related to its cluster membership and thus its characteristics. However, local variation of factors affecting runoff combined with missing information in the national datasets means that model performance is not homogeneous in any cluster.

Sensitivity to the Strickler coefficient representing surface roughness was examined, as were sensitivities to rainfall and potential evapotranspiration (PET) quantities and distributions. In tests where relatively small perturbations were applied to the standard datasets, such as increasing PET by 5% or using an exponential distribution for sub-daily rainfall, mixed changes in model performance were typically obtained, with a fairly even split between catchments showing overall improvements and declines in performance. These tests appear to indicate a degree of robustness in terms of system setup and inputs, with variation in catchment performance responses to changes suggesting little systematic bias and good overall simulation if not necessarily always locally optimal. This is reflected when larger changes are applied, such as a 20% increase in rainfall or a Strickler coefficient of 5, which leads to large overall decreases in model performance. In conjunction with the higher variability of responses arising from smaller perturbations, this is taken as an indication that generally appropriate parameters and input datasets are being used in the system. Furthermore, where the changes made had a physical basis - particularly using realistic hourly meteorological as inputs - model performance tended to increase slightly at the national scale. The largest percentage of improved catchment simulations resulted from incorporating hourly rainfall based on the new dataset created during this project, as well as PET disaggregated to hourly intervals. Improvements were also seen when the AE/PE ratio was increased by 0.1, PET increased by 5% or rainfall decreased by 10%, which reflects the possibility that calculated PET values may be too low or rainfall slightly too high, although the latter may be less congruent with the expectation of undercatch for a number of raingauges. This is also supported by Henriksen et al. (2003) who find a systematic deviation in water balances, which they attribute to either an overestimation of precipitation or an underestimation in PET.

# Chapter 6

# Climate change impact assessment

The SHETRAN for GB modelling system described in previous chapters has been used to conduct a preliminary assessment of potential climate change impacts on water resources for selected catchments across the UK. At this stage the assessment is intended to demonstrate the fidelity of the modelling system for conducting broad scale analyses of climate change impacts. For this reason the analysis focusses on a single emissions scenario and a given time slice. Therefore this study should be treated as a proof of concept rather than a conclusive evaluation of changes across the ensemble of change.

Many studies have been conducted on the impact of climate change on river flows in the UK, particularly focusing on specific catchments and areas. Kay and Jones (2010) used 3 transient climate projections from HadRM3 for 2 catchments to investigate annual maxima timeseries for changes in flood frequency using the PDM model introduced in Chapter 2. Kay and Jones showed that using time slices as the basis for analysis of climate change impacts can be unreliable, as small changes in the definition of the simulation window can have a large impact on results. The authors then proceeded to examine the national context by running G2G for the UK, which highlighted the possibility of higher flood risk across much of the country in the future. Ledbetter et al. (2012) took a different approach by developing a new resampling methodology in order to create an ensemble of change factors to meteorological input data. Ledbetter et al. used this approach to analyse impacts on flows in the Eden catchment in Scotland and showed that projected changes in flows are sensitive to the techniques used to create input meteorological data. Lopez et al. (2009) used a large perturbed physics ensemble with a water resource model for an area of south-west England and demonstrated that value is added by using a larger set of climate inputs, which provide better understanding of the range of plausible futures. Limbrick et al. (2000) applied

several climate scenarios to model one catchment (Kennet at Theale), while Wilby and Harris (2006) developed an uncertainty framework for climate change impact assessment accounting for four GCMs, two emissions scenarios, two statistical downscaling techniques, two hydrological model structures, and two sets of hydrological model parameters for the Thames catchment. Applying this uncertainty framework revealed that GCMs and scenario selection are the largest sources of uncertainty in a climate change impact assessment. These studies are typically small scale, examining only one or two catchments, using a limited selection of driving climate data or examining only one aspect of the flow regime.

Christierson et al. (2012) go further and model 70 catchments across the UK with the UK Climate Projections 2009 for the 2020s under the A1B scenario. The Future Flows project (Prudhomme et al., 2013) goes beyond this and used a consistent methodology to assess changes to flow across the whole country. The meteorological inputs to the hydrological models were derived from Future Flows Climate (Prudhomme et al., 2012), a national 11-member ensemble of projections derived from the Hadley Centre's ensemble projection HadRM3-PPE at the 2050s time slice.

A commonality between these studies is that they largely depend on conceptually based model formulations rather than physically based approaches. As discussed in Chapter 2, conceptual models may not necessarily form a sound basis for hydrological predictions under a changing climate, as a result of their high dependency on calibration to prevailing climatic conditions. In theory, physically based models provide a more robust method of simulating catchment hydrology for the future. Application of such models in national scale projections has been limited to date because of their high complexity and computational cost. Therefore this analysis provides a novel insight into the applicability of such models for broad scale impact assessment as well as the opportunity for model intercomparison.

## 6.1 Climate change projections and the UKCP09 Weather Generator

Current understanding suggests that climate change could have a profound impact on the hydrological cycle with possible increases in the frequency of extreme events such as floods (Murphy et al. (2009); Kendon et al. (2014)). The potential magnitude of hydrological changes is such that government agencies at various levels and other organizations need to consider adaptation strategies. Therefore, these institutions require

indications of the likely direction of change in key hydrological quantities in order to plan effectively for extreme events and changes in catchment regimes.

A climate impact assessment is not a simple process. Analysis begins with global climate models (GCMs) which provide indications of change for the world under various emissions scenarios. Their coarse resolution is inadequate for regional studies and so their outputs are used as boundary conditions for Regional Climate Models (RCMs). The finest RCM resolution to date for the UK is 1.5km (Chan et al., 2014) which is just fine enough to model convective cells for intense rainfall. UKCP09 (Murphy et al., 2009) provides probabilistic change factors for climate variables from an 11 member RCM ensemble at a 25km resolution and are the standard set of climate change projections for the UK. However, change factors cannot be used directly in a hydrological model. Rather, time series of all model input variables are required. For SHETRAN these variables are rainfall, PET and maximum and minimum temperature. As SHETRAN is a spatially distributed model, the variables should be consistent internally, spatially and with each other. The UKCP09 point weather generator (Kilsby et al., 2007) has been further developed to accommodate these needs.

The spatial weather generator (unpublished) produces internally consistent series of meteorological variables: precipitation, temperature, vapour pressure, wind and sunshine, as well as a number of derived variables calculated from the meteorological series, for example PET. Observed daily rainfall totals and values of other weather variables are used to calibrate the weather generator for a baseline climate 1961 to 1990. Change factors at the monthly time scale for each grid square are taken from the UKCP09 probabilistic projections to define the range of possible climate change futures. The weather generator uses a stochastic rainfall model (Neyman-Scott Rectangular Pulse) and the other weather variables are then generated conditioned on the rainfall series on a 5km grid. This procedure provides an internally and spatially coherent time series of the input variables required in SHETRAN. Kay and Jones (2012) show that using the weather generator for impact assessments provides benefits over simply using a range of change factors, as it more fully represents natural variability.

## 6.2 Uncertainty

Climate change is an important and complex global problem. The associated ongoing and future impacts are predicted to be wide-ranging and profound, as reported in the IPCC Fifth Assessment Report (WGII AR5) (Pachauri et al., 2014) such that timely

and effective decision-making are critical for mitigation and adaptation. However, decision-making processes are confounded by a range of issues, not least the scientific uncertainty associated with appraisal of possible climatic responses and trajectories, impacts and mitigation/adaptation options. For example, computer models are an essential tool for assessing future scenarios, but they are inevitably reductionist and subject to myriad limitations, their accuracy limited by our partial understanding of the highly complex systems we attempt to simulate. This means that model results contain uncertainty that needs to be robustly factored into human responses to climate change.

The uncertainties associated with the kind of study outlined in this chapter are numerous and can be summarised in a 'Cascade of Uncertainty' (Wilby and Dessai, 2010) which shows the cumulative effect of uncertainties through the modelling process; beginning with those associated with scenario generation, which rely on an understanding of future societies, flowing through to climate models, impact models, cities, vulnerability, adaptation strategies and decision-making processes. Each additional step not only propagates the previous set of uncertainties but also introduces its own uncertainties as a function of the research questions asked, methods and data used, and the spatial and temporal scales considered inter alia.

Many attempts have been made to quantify uncertainty, and each study takes a different approach. Christierson et al. (2012) use the GLUE methodology (Beven and Binley, 1992) to assess the structural uncertainty of their hydrological model and use Latin Hypercube sampling to select 20 change factors from the UKCP09 projections. Prudhomme et al. (2013) compare projections from 11 RCM scenarios. Ledbetter et al. (2012) resample both the baseline and future precipitation (from 13 climate models) to incorporate an estimate of climate variability. Regardless of method, the uncertainty associated with climate change impact projections is still large, and so these studies should be used to indicate likely directions of change for the decision making process. Clearly it is very difficult to accurately project into the future and so all studies should be evaluated with reference to their simplifying assumptions and limitations.

## 6.3 Scenario selection

The SRES A1B Medium emissions scenario has been selected in conjunction with the 2050s time slice as the primary focus for this analysis. This selection is consistent with the scenarios investigated in the future flows project. For a full investigation, multiple

emissions scenarios and time slices would be modelled, however as the main purpose of this exercise is to scope the feasibility of assessing broad scale climate change impacts using SHETRAN, a single scenario will suffice.

## 6.4 Catchment selection

Of the 350 catchments discussed in previous chapters, 20 have been selected for this analysis. These catchments have been chosen to provide a good geographical coverage of the UK. The selected catchments are also ones that are well simulated by SHETRAN, with all but two catchments exhibiting NSE of greater than 0.7 (the Witham at Claypole Mill was chosen for its location in the south east of England). The flow duration curves for these catchments are provided in Appendix B. The range of base flow index (BFI) is 0.21 to 0.69, although the majority of the selected catchments are largely surface water dominated. Further refinements to the representation of some of hydrological processes in some regions of the UK are required before climate change analysis can be conducted, e.g. the chalk (catchments with a significant groundwater component are still required). Catchment areas range from 69.4km$^2$ to 2175.6km$^2$. Full details of their properties are given in Table 6.1. The locations of the catchments can be found in Figure 6.1.

## 6.5 Automated set up and technical issues

The weather generator and SHETRAN are not directly compatible as they do not share a common data format. In order to simplify the process of coupling weather generator output to SHETRAN for GB several tools have been developed. The first tool automates the identification of weather generator cell IDs, the second converts weather generator output into SHETRAN input and the third runs the SHETRAN models in parallel to reduce computing time.

To identify the weather generator grid cell ids:

- Select a catchment from SHETRAN for GB
- Use the existing rainfall map and mask identify grid squares within the catchment
- Look up the weather generator ID codes from a file of pre calculated IDs and coordinates
- Output a list of IDs as a text file.

| Name | Number | Area ($km^2$) | BFI | Mean Flow ($m^3$/d) | NSE | Geology | Factors affecting Runoff |
|---|---|---|---|---|---|---|---|
| Aire at Kildwick Bridge | 27035 | 282.3 | 0.37 | 6.46 | 0.89 | Limestone with extensive boulder clay cover | Reservoirs |
| Almond at Craigiehall | 19001 | 369.0 | 0.39 | 6.16 | 0.79 | Carboniferous rocks overlain by superficial deposits | Abstraction and effluent returns |
| Bervie at Inverbervie | 13001 | 123.0 | 0.56 | 2.15 | 0.69 | Bedrock of mixed permeability; almost 100% covered with superficial deposits | Natural |
| Braan at Hermitage | 15023 | 210.0 | 0.43 | 6.99 | 0.79 | Metamorphic bedrock geology with approx. 60% overlain by superficial deposits | Natural |
| Ceiriog at Brynkinalt Weir | 67005 | 113.7 | 0.54 | 3.11 | 0.79 | 90% of catchment is impermeable | Natural |
| Dart at Austins Bridge | 46003 | 247.6 | 0.52 | 11.30 | 0.70 | Granite | Reservoirs |
| Don at Doncaster | 27021 | 1256.2 | 0.56 | 15.96 | 0.72 | Grit, limestones and sandstones" | Reservoirs, abstraction and effluent returns |
| Eden at Kirkby Stephen | 76014 | 69.4 | 0.26 | 2.62 | 0.79 | Carboniferous Limestone and Permian Sandstone. variable Boulder Clay cover | Natural |
| Enrick at Mill of Tore | 6008 | 105.9 | 0.30 | 3.31 | 0.83 | Impermeable with approx. 20% superficial deposits. | Natural |
| Frome (Somerset) at Tellisford | 53007 | 261.6 | 0.52 | 3.77 | 0.80 | limestone | Abstraction and groundwater abstraction |
| Frome at Yarkhill | 55018 | 144.0 | 0.52 | 1.15 | 0.75 | Bedrock of very low permeability. | Effluent returns |
| Greta at Rutherford Bridge | 25006 | 86.1 | 0.21 | 2.28 | 0.70 | Millstone Grit with ¿ 90% superficial deposits of peat and boulder clay | Natural |
| Leet Water at Coldstream | 21023 | 113.0 | 0.33 | 0.98 | 0.78 | Boulder Clay overlying calciferous sandstone | Natural |
| Lossie at Sheriffmills | 7003 | 216.0 | 0.53 | 2.74 | 0.8 | Bedrock Schists. Extensive superficial deposits. | Abstraction |
| Stour at Throop | 43007 | 1073.0 | 0.65 | 13.71 | 0.73 | Predominantly Chalk ( 50%); some clay ( 30%); limestone and Upper Greensand. | Abstraction, groundwater abstraction and effluent returns |
| Teme at Tenbury | 54008 | 1134.4 | 0.55 | 14.40 | 0.74 | Mainly Palaeozoic sediments with Pre-Cambrian crystalline rocks | Effluent returns |
| Thurso at Halkirk | 97002 | 412.8 | 0.45 | 9.09 | 0.76 | Mixed bedrock permeability and approx. 95% overlain by superficial deposits | Regulation and abstraction |
| Tyne at Bywell | 23001 | 2175.6 | 0.38 | 46.23 | 0.71 | Carboniferous Limestone with extensive superficial deposits | Reservoirs |
| Urr at Dalbeattie | 80001 | 199.0 | 0.36 | 5.96 | 0.84 | Silurian shales extensively covered by Boulder Clay. | Natural |
| Witham at Claypole Mill | 30001 | 297.9 | 0.69 | 1.86 | 0.64 | Clay with limestone | Regulation, abstraction and effluent returns |

*Table 6.1: Summary table of catchment characteristics.*


*Figure 6.1: Locations of the 20 catchments studied.*

The grid IDs can now be input to the weather generator. To run the spatial weather generator:

- Open the Weather Generator software
- Copy and paste the grid IDs into the program
- Select 10 control runs and 100 2050 medium emissions scenario runs, each 30 years long
- Select no change in urban fraction (a parameter of the weather generator)
- Select a daily time step (running the spatial weather generator at an hourly time step is still very slow and for 20 catchments it was estimated that it would take several months to generate input for SHETRAN).

The output from the weather generator gives one file for each grid square which contains 100 x 30 year time series written directly after each other. To convert this to SHETRAN input:

- Reformat the data so that each variable has its own .csv file and so that there is 1 .csv file for each of the 100 (or 10) runs. The data for every grid ID should be written to one file
- Exponentially distribute the rainfall using the same process outlined in Chapter 4.

The meteorological variables time series are now ready for use with SHETRAN. To run these simulations efficiently:

- Set up the other SHETRAN files required to run a simulation, including a new library file
- Run the simulations in parallel using a service such as Condor or The Cloud
- Or use a Python script to parallelize processing over a series of multi processor servers
- Retrieve and analyse the results.

This process can be finished in 24 hours for an average catchment (using all available computing resources). For larger catchments the process may take 2-3 days with no technical difficulties.

Two main technical issues were identified when using weather generator inputs with SHETRAN:

– Due to increased rainfall intensity in future scenarios, the maximum allowed rainfall in a time step (under section :fr20 in the SHETRAN input files) had to be increased from 0.5 to 0.7

– Numerical instabilities occurred caused by backwater effects at channel junctions in the flatter parts of large catchments. This was rectified by increasing the 'Minimum Drop Between Channels' in the Library file.

Both of these changes have a negligible impact on flow calculations but were necessary for simulations to complete. For a full national climate impact assessment, it would be necessary to further automate some of these steps. The weather generator would have to be run via a script and not via the user interface as described above. Also, the identification of causes of model instability and subsequent alterations would need to be automated. Both of these steps would be trivial to instigate when they are required.

## 6.6 Validation of baseline data

From carrying out 10 baseline (or control) runs from each catchment, it was found that in some cases the rainfall differed between the observed and control simulations. Further investigation of this issue revealed that this was due to the difference in monthly mean rainfall between the UKCP09 daily gridded datasets and the statistics coded into the weather generator These statistics were based on an earlier national gridded rainfall dataset provided by the Met Office several years ago. This is shown in Figure 6.2 which indicates that in most areas the difference in monthly mean rainfall can be around +=10%, but for a very few grid squares the differences can be as large as 92%. However, there is no coherent pattern in the discrepancies between the UKCP09 data and weather generator statistics, with the percentage difference varying both spatially and temporally. Work is ongoing to update the weather generator to be wholly consistent with the UKCP09 daily gridded dataset. This update should appear in the next version of the weather generator.

The difference between the observed and control runs is not considered to significantly affect this analysis as the main purpose is to automate a system for using the spatial weather generator in conjunction with SHETRAN. When the new version of the weather generator is available, it will link very easily to the system described in this chapter. Furthermore, the differences between the control and future runs are still considered to be valid differences although caution is required in interpreting in changes in hydrology

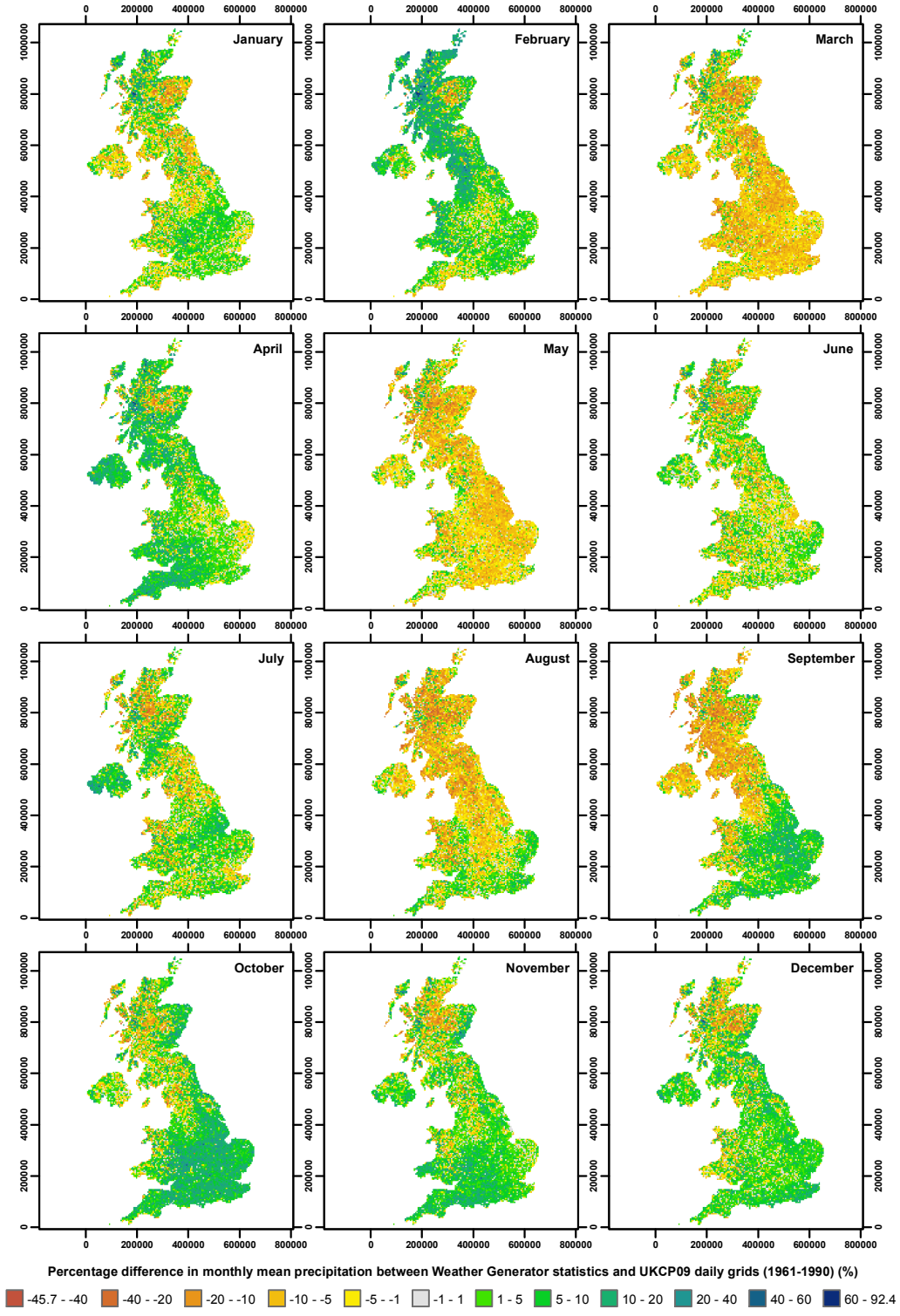



*Figure 6.2: Graphs showing the percentage differences between the mean monthly rainfall values that the weather generator uses and the mean monthly rainfall values calculated from the 5km gridded dataset.*

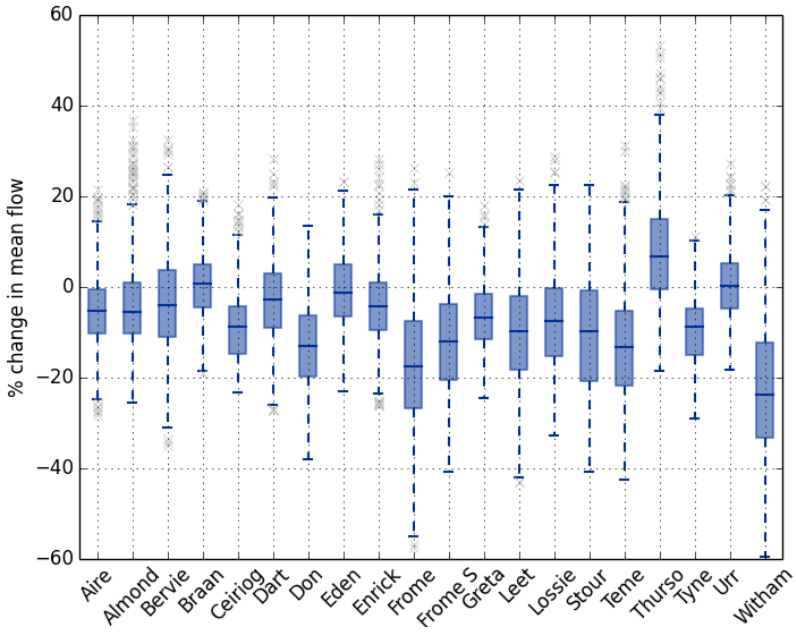



*Figure 6.3: % change in mean annual flows between baseline and 2050s runs.*

in absolute terms. To avoid confusion, results from simulations using observed data are not included in the analysis in subsequent sections of this chapter.

## 6.7 Future scenarios

The analysis in this section now focusses on changes in flows, rainfall and PET between the control period (1961-1990) and future scenarios (2050s). Both control and future scenarios have been generated by the weather generator. 10 control scenarios and 100 future scenarios were generated for each catchment. This is in line with the minimum number of runs recommended in the weather generator guidance to give a representative sample of the distribution of possibilities.

### 6.7.1 Annual changes

The discussion of results begins with changes in annual mean flows and then moves to look at the results on a seasonal basis. The annual results are summarised in Table 6.2, which shows the percentage change in the mean of the distribution of annual average flows (cumecs), rainfall (mm/day), pet (mm/day) from control to future.

Figure 6.3 shows the changes in distributions of annual average flow between control and

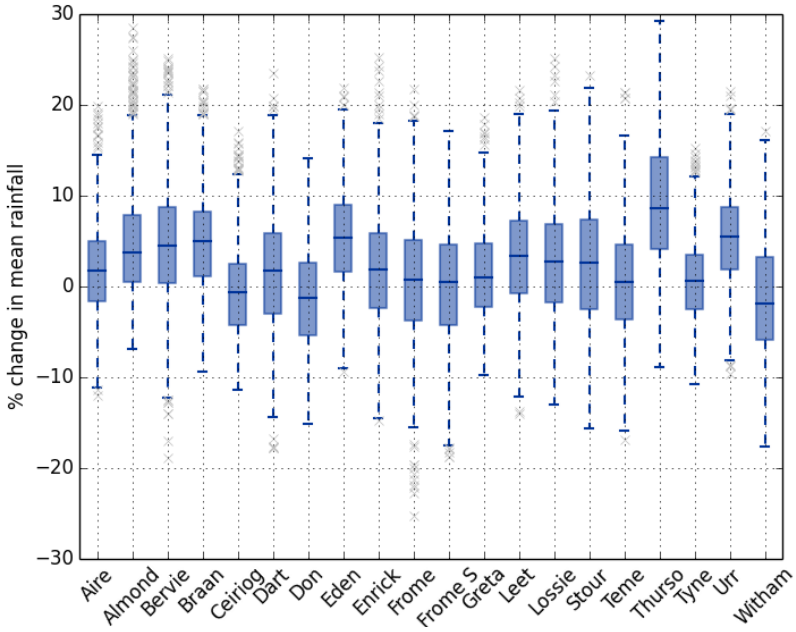



*Figure 6.4: % change in mean annual rainfall between baseline and 2050s runs.*

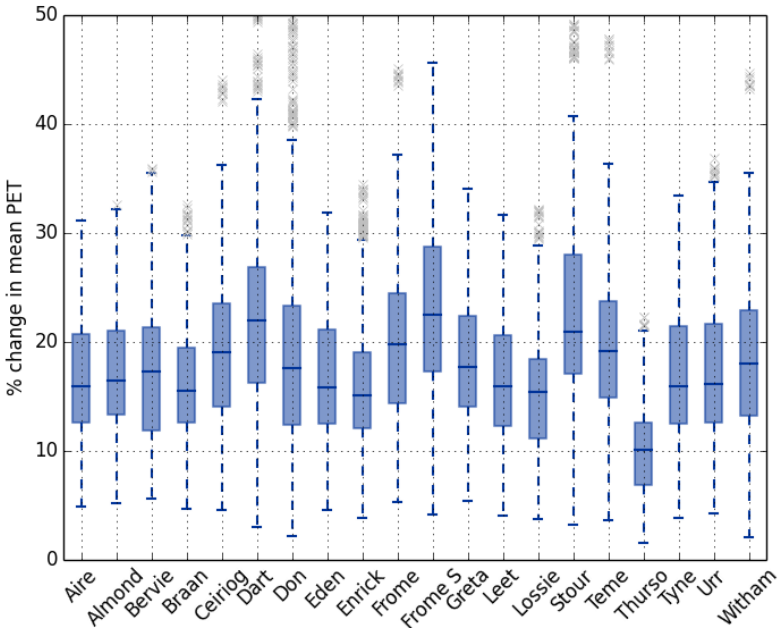



*Figure 6.5: % Change in mean annual PET between baseline and 2050s runs.*

| Name | Flow | Rain | PE |
|---|---|---|---|
| Aire at Kildwick Bridge | -5.2 | 1.8 | 16.0 |
| Almond at Craigiehall | -5.5 | 3.8 | 16.5 |
| Bervie at Inverbervie | -3.9 | 4.5 | 17.3 |
| Braan at Hermitage | 0.7 | 5.0 | 15.5 |
| Ceiriog at Brynkinalt Weir | -8.7 | -0.7 | 19.0 |
| Dart at Austins Bridge | -2.7 | 1.8 | 21.9 |
| Don at Doncaster | -13.1 | -1.3 | 17.6 |
| Eden at Temple Sowerby | -1.2 | 5.4 | 15.8 |
| Enrick at Mill of Tore | -4.1 | 1.9 | 15.1 |
| Frome at Yarkhill | -17.4 | 0.7 | 19.8 |
| Frome (Somerset) at Tellisford | -11.9 | 0.5 | 22.5 |
| Greta at Rutherford Bridge | -6.9 | 1.0 | 17.7 |
| Leet Water at Coldstream | -9.9 | 3.3 | 15.9 |
| Lossie at Sheriffmills | -7.6 | 2.8 | 15.4 |
| Stour at Throop | -9.9 | 2.6 | 21.0 |
| Teme at Tenbury | -13.3 | 0.5 | 19.1 |
| Thurso at Halkirk | 6.7 | 8.6 | 10.1 |
| Tyne at Bywell | -8.8 | 0.7 | 15.9 |
| Urr at Dalbeattie | 0.3 | 5.5 | 16.1 |
| Witham at Claypole Mill | -23.7 | -1.8 | 18.0 |

*Table 6.2: Summary table of the median % change in mean annual Flow, Rainfall and PET between baseline and future scenarios.*

future climates based on the multiple realisations of each case. In general terms, annual average flows typically decrease in the modelled catchments, with greater decreases in the south of GB compared with the north. Looking at the differences between the mean of average annual flows for the control and future scenarios, 12 catchments exhibit decreases in flow of up to 10% and 5 show decreases of up to 20%. The Witham at Claypole Mill shows by far the greatest reduction in flow of 23.7%, which reflects large projected increases in evapotranspiration in a small catchment (where changes in intrinsically small flows may be large in relative terms). Only 3 catchments situated in the north of GB show increases in flow, but in these cases the magnitudes of change are relatively modest. For example, the maximum increase out of the modelled catchments occurs for the Thurso at Halkirk, which shows a change in the mean of annual flows of 6.7% . The increase in flow here is a function of the fact that this catchment is subject to the largest relative increases in rainfall and the smallest increases in PET out of the selected catchments due to its high latitude.

Examining the distributions of average annual flows for the realisations of the control and future climates in Figure 6.3 reveals a general pattern across the catchments located in England and Wales particularly. In these catchments, the upper tail of the

distribution is typically similar for both control and future scenarios, indicating that changes in annual flows in wetter years are simulated as being fairly minor. However, the lower tail of the future distribution is typically lower than that of the control for the majority of catchments modelled. This results in a skew towards lower average annual flows in the future scenario relative to the control period, which corresponds with the reduction in the mean of average annual flows for the future climate relative to the control climate discussed above. This pattern is less apparent for the catchments located in Scotland, which tend to show less change in the distribution of annual average flows.

Analysis of the distribution of changes in rainfall shown in Figure 6.4 shows that average annual rainfall is very similar between control and future climates. The only appreciable change is seen at the Thurso at Halkirk, which exhibits a shift towards higher precipitation (increase of 8.6%), in turn explaining the more significant increases in annual flows simulated here relative to the other catchments. In contrast, a simpler and more consistent trend is apparent for changes to PET in the future, which are shown in Figure 6.5. This figure shows a simple trend of PET increasing more the further south a catchment is, although all selected catchments show large increases in PET relative to the control period. From this figure it can also be seen that the distribution of changes in future climate is wider than that of rainfall, which suggests that a broader range of PET futures is possible. The lower tail of the distribution is always greater than 0 and the upper tail of the distribution can be greater than 40% in some cases. It is therefore apparent that the pattern of changes in PET is more consistent than that of rainfall, which tends to demonstrate more complex variability.

The analysis thus suggests that, at the annual level, it appears that the main driver of the overall reduction in flows across GB is an increase in PET rather than a reduction in rainfall. As discussed above, the main exception to this is the increase in rainfall in the Thurso at Halkirk, which counteracts the comparatively small increase in PET, resulting in small increases in annual flows observed. This preliminary analysis of average annual flows across GB suggests a reasonably consistent picture in terms of the direction of change, though further research could elucidate the geographical locations and characteristics of catchments showing different behaviour. There is also clearly some variation in the magnitude of changes in average annual flows between catchments, which reflects both the influence of projected future scenarios applied through the weather generator and the modulating influence of catchments characteristics.

### 6.7.2 Seasonal hydrological means

This section outlines the relative changes in seasonal mean flows. Relative changes in flows are analysed in line with the approach taken in several water management applications, including abstraction licensing in line with EU Water Framework Directive legislation (2000). Spring, summer, autumn and winter are defined as MAM, JJA, SON, and DJF respectively. Table 6.3 summarises changes in the average daily rainfall, PET and flow between the 10 control runs and 100 future runs.

In general, Figures 6.6 to 6.11 show:

- Slight decreases in flow in spring due to moderate increases in PET and negligible variation in rainfall
- Large decreases in flow in summer due to decreased rain and increased PET
- Large decreases in flow in autumn due to increases in PET but little change to rainfall
- Small increases in flow in winter due to increases in rainfall but diminished by increases in PET.

The general patterns of seasonal changes in rainfall, PET and flow between control and future scenarios may be interpreted to form an overall picture of the potential seasonal effects of climate change in this scenario. Table 6.3 indicates that spring mean changes in rainfall between control and future periods are generally fairly limited in magnitude, with relative changes typically less than around 7%. Spatial variation in the direction of change in spring rainfall is apparent, however. Catchments situated in the northern part of GB show modest increases in spring rainfall of up to 7.5%, whereas some of the catchments in the southern part of the study area show slight decreases (up to -3.9%). PET increases range from 6.3% in the north of GB to 16.3% in the south. The combined effects of these changes in rainfall and PET result in the changes in flows shown in Figures 6.6 and 6.7. These figures show reductions in flow in the south of GB, with the distribution of flows being slightly skewed towards lower flows. However, some catchments appear to exhibit more complex responses. For example, the Tyne at Bywell shows a particularly large reduction in flows relative to overall patterns, given its location in the north of England. In addition, the Dart at Austins Bridge shows a relatively small reduction in flows despite high increases in PET due to it being one of the wettest catchments.

In summer, this analysis suggests a ubiquitous reduction in rainfall across GB. This is complemented by PET increases, with greater increases apparent in the south, a

| | Flow | | | | Rain | | | | PE | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Name | Spring | Summer | Autumn | Winter | Spring | Summer | Autumn | Winter | Spring | Summer | Autumn | Winter |
| Aire at Kildwick Bridge | -8.8 | -42.0 | -10.7 | 6.8 | 1.9 | -20.4 | 7.5 | 11.1 | 13.0 | 20.5 | 15.1 | 13.0 |
| Almond at Craigiehall | -2.2 | -21.6 | -11.6 | 2.0 | 6.9 | -15.5 | 11.8 | 10.5 | 14.8 | 18.8 | 18.3 | 14.3 |
| Bervie at Inverbervie | -7.6 | -23.4 | -8.2 | 3.2 | 2.7 | -10.8 | 9.8 | 10.4 | 13.7 | 18.0 | 18.4 | 12.5 |
| Braan at Hermitage | -8.2 | -28.5 | -0.3 | 13.5 | 2.8 | -13.8 | 8.5 | 14.5 | 12.1 | 19.2 | 15.1 | 10.9 |
| Ceiriog at Brynkinalt Weir | -12.3 | -26.3 | -17.0 | 1.7 | -3.2 | -20.2 | 2.0 | 7.9 | 16.3 | 22.3 | 17.5 | 12.8 |
| Dart at Austins Bridge | -7.5 | -41.1 | -19.7 | 16.6 | -2.3 | -25.2 | -2.3 | 18.3 | 15.4 | 28.9 | 18.4 | 11.8 |
| Don at Doncaster | -13.5 | -23.9 | -29.7 | -2.7 | -3.2 | -20.2 | -1.9 | 10.0 | 14.4 | 21.6 | 17.9 | 10.5 |
| Eden at Temple Sowerby | -7.1 | -25.9 | -10.2 | 12.8 | 5.3 | -14.4 | 5.0 | 17.1 | 13.5 | 19.8 | 14.9 | 13.7 |
| Enrick at Mill of Tore | -9.8 | -17.6 | -2.0 | 0.4 | 0.7 | -12.5 | 8.7 | 2.2 | 12.3 | 16.3 | 16.1 | 11.6 |
| Frome at Yarkhill | -15.6 | -22.7 | -39.7 | -8.4 | -3.9 | -21.3 | 3.4 | 18.5 | 15.7 | 23.9 | 17.8 | 11.5 |
| Frome (Somerset) at Tellisford | -4.7 | -31.7 | -38.3 | 1.1 | -0.9 | -23.2 | 0.0 | 16.0 | 15.1 | 28.3 | 21.8 | 13.0 |
| Greta at Rutherford Bridge | -13.4 | -25.0 | -19.9 | 8.1 | 3.4 | -14.5 | -0.3 | 11.9 | 15.2 | 19.8 | 16.1 | 15.2 |
| Leet Water at Coldstream | -6.8 | -36.0 | -22.1 | -1.5 | 2.7 | -15.6 | 10.3 | 14.5 | 14.7 | 17.1 | 17.0 | 10.6 |
| Lossie at Sheriffmills | -5.4 | -18.0 | -13.9 | -0.9 | 4.0 | -12.2 | 10.6 | 8.3 | 11.6 | 17.3 | 18.4 | 9.4 |
| Stour at Throop | -6.9 | -24.0 | -41.7 | 0.6 | 1.4 | -17.3 | -0.9 | 14.6 | 16.0 | 25.4 | 22.3 | 15.8 |
| Teme at Tenbury | -12.1 | -28.8 | -34.0 | -0.8 | 0.5 | -16.9 | 0.0 | 11.2 | 14.4 | 21.8 | 19.8 | 13.7 |
| Thurso at Halkirk | 0.9 | -12.5 | 5.2 | 13.8 | 7.5 | -8.5 | 12.8 | 16.5 | 6.3 | 13.9 | 10.3 | 4.1 |
| Tyne at Bywell | -15.5 | -31.1 | -17.0 | 3.3 | -0.8 | -15.3 | 3.8 | 9.5 | 13.6 | 18.6 | 15.9 | 11.4 |
| Urr at Dalbeattie | -3.6 | -39.7 | -4.7 | 13.2 | 3.5 | -15.9 | 7.5 | 16.5 | 13.5 | 18.7 | 15.9 | 15.6 |
| Witham at Claypole Mill | -12.2 | -26.9 | -42.8 | -22.2 | 0.7 | -16.3 | -2.6 | 10.6 | 13.2 | 21.0 | 17.1 | 9.8 |

*Table 6.3: Table showing median % change in means of flow, rainfall and PET from the control to future scenarios.*

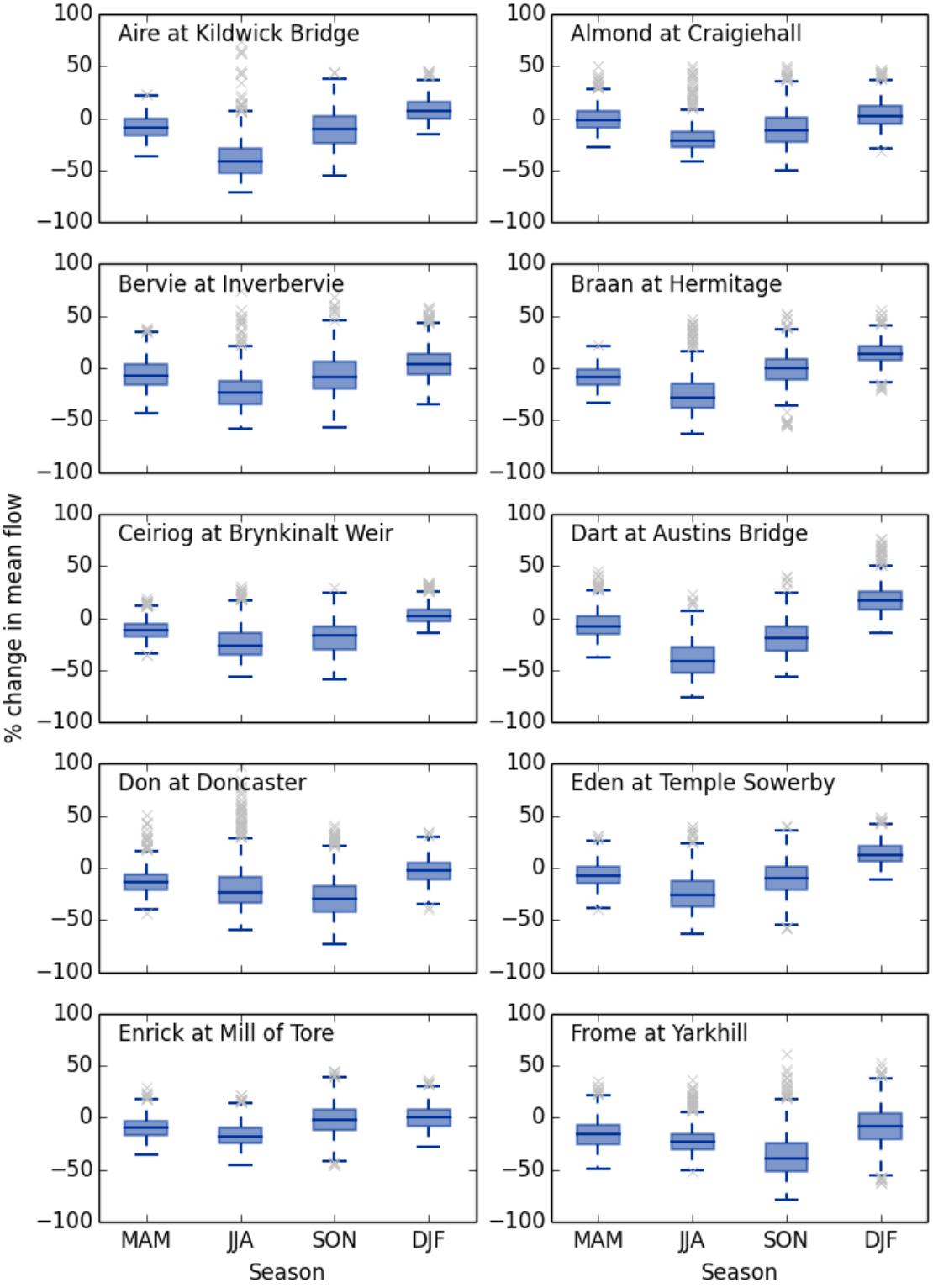



*Figure 6.6: % change in mean seasonal flows between baseline and 2050s scenarios.*

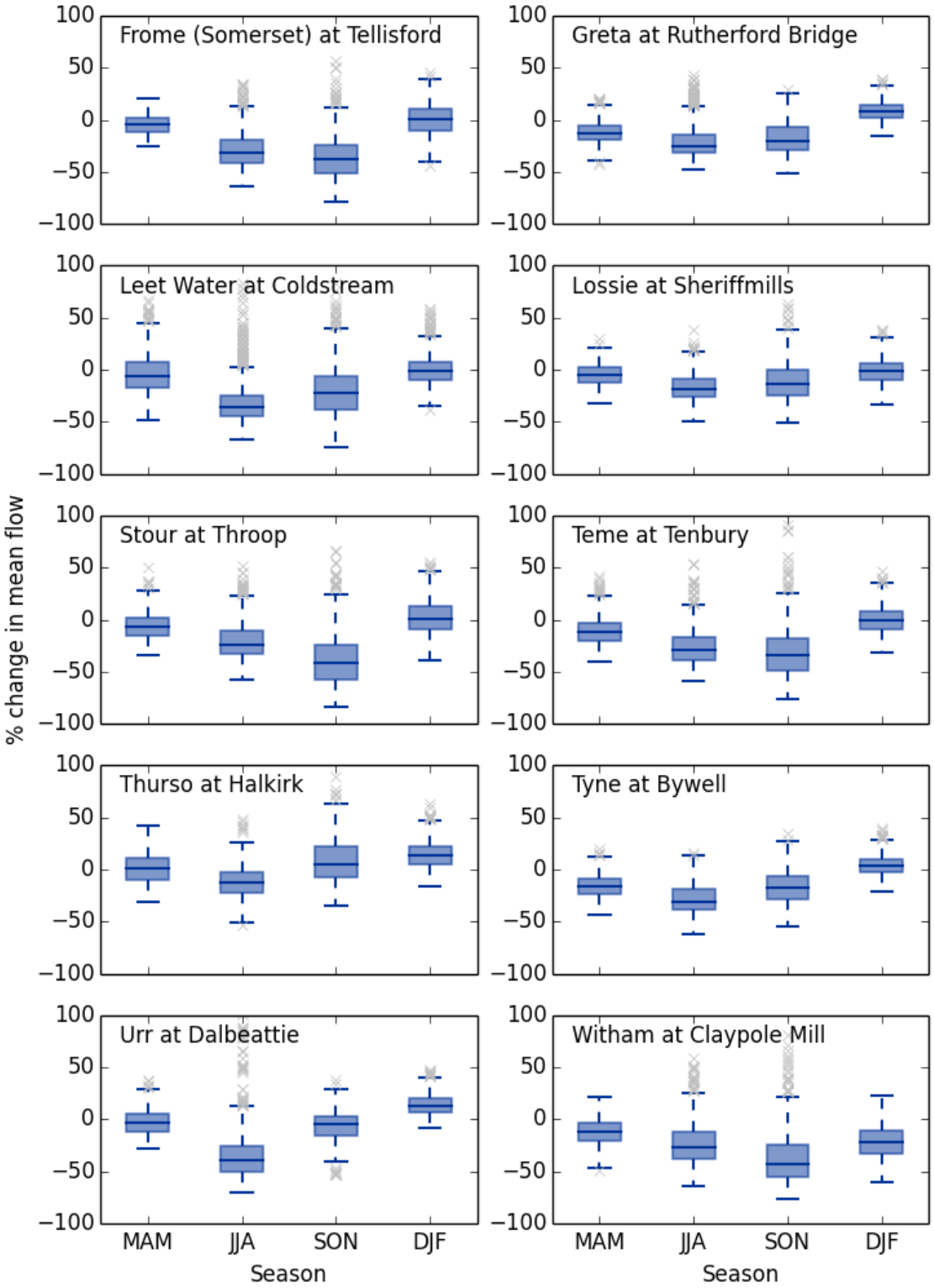



*Figure 6.7: % change in mean seasonal flows between baseline and 2050s scenarios.*

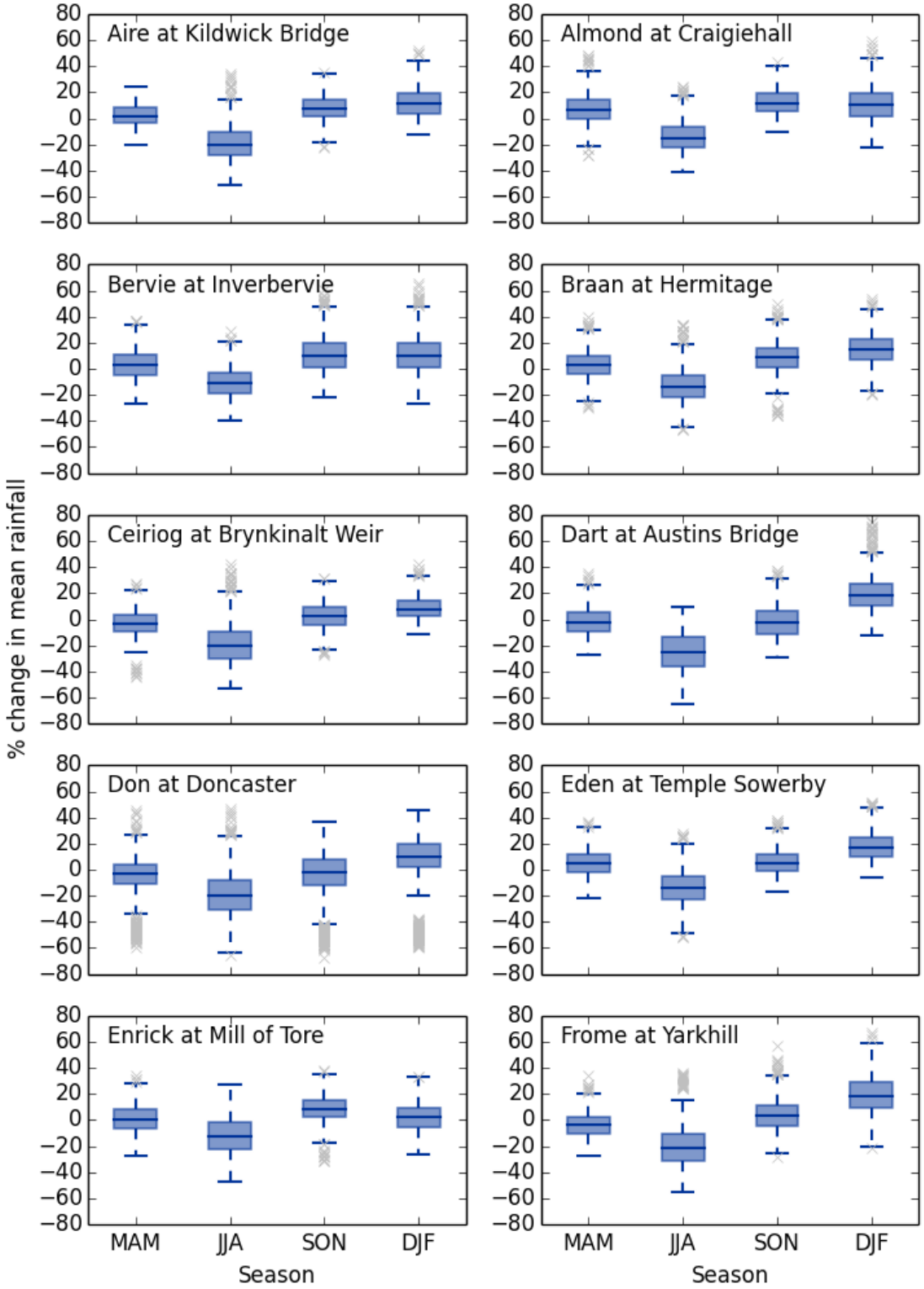



*Figure 6.8: % change in mean seasonal rainfall between baseline and 2050s scenarios.*

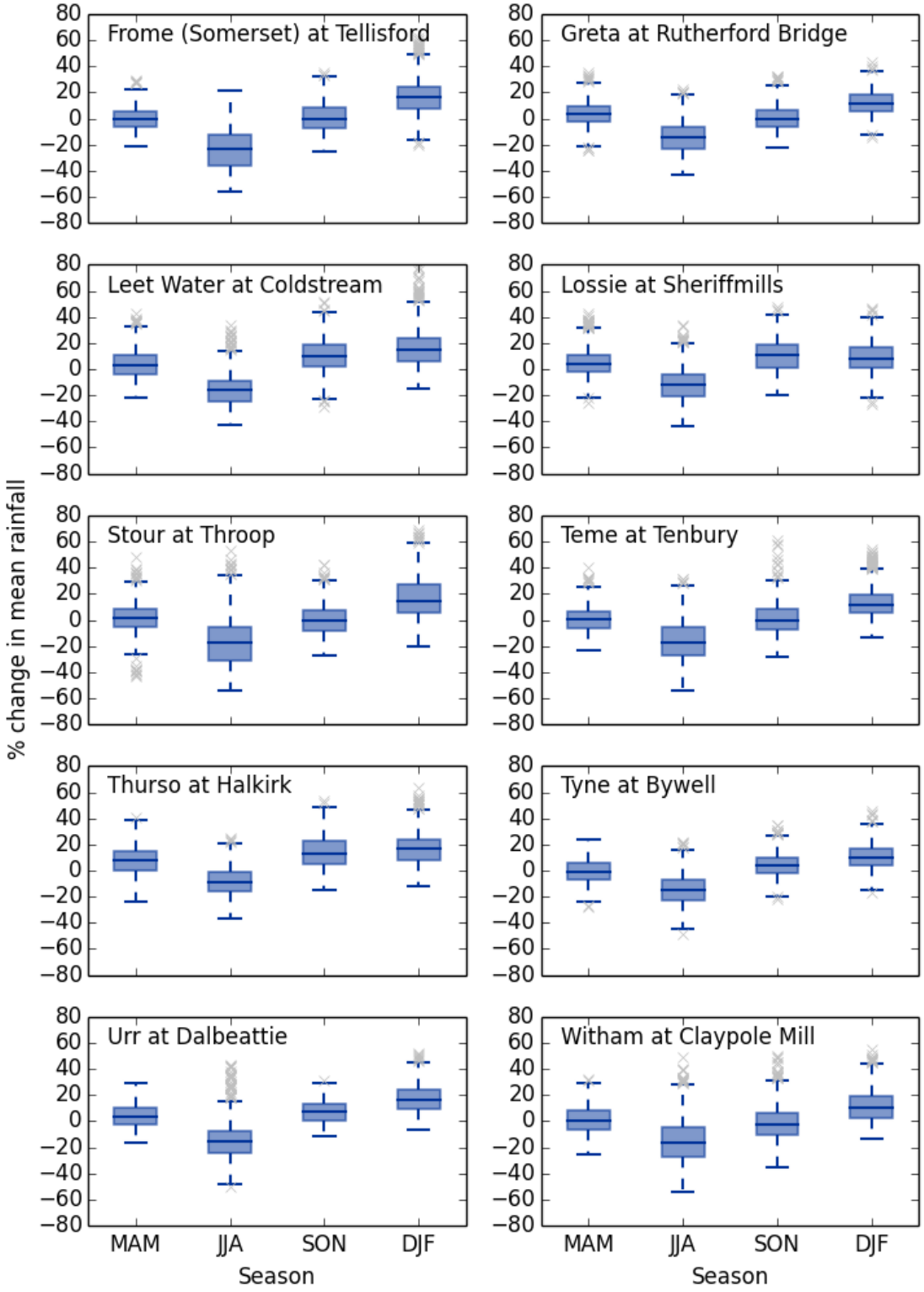



*Figure 6.9: % change in mean seasonal rainfall between baseline and 2050s scenarios.*

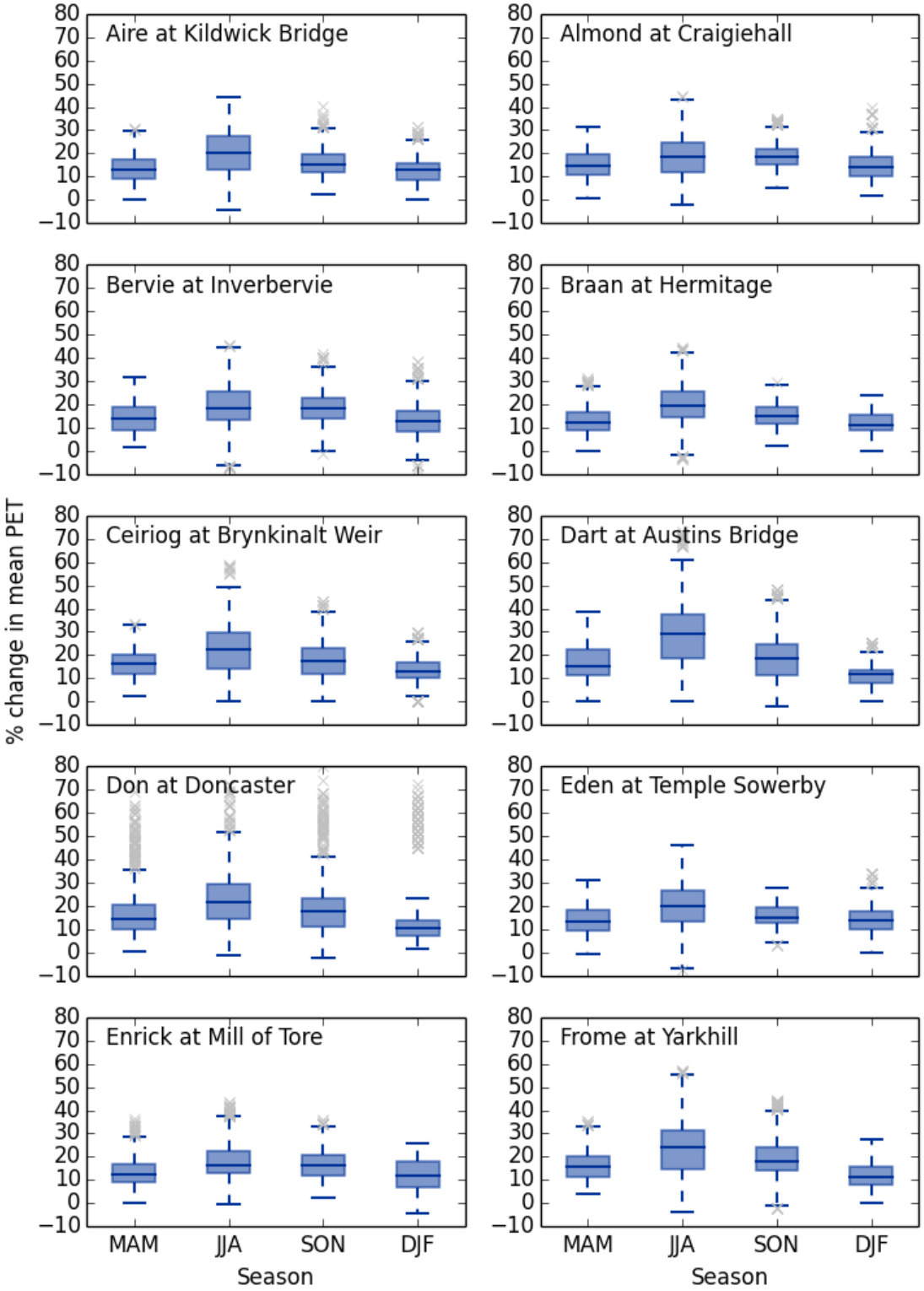



*Figure 6.10: % change in mean seasonal PET between baseline and 2050s scenarios.*

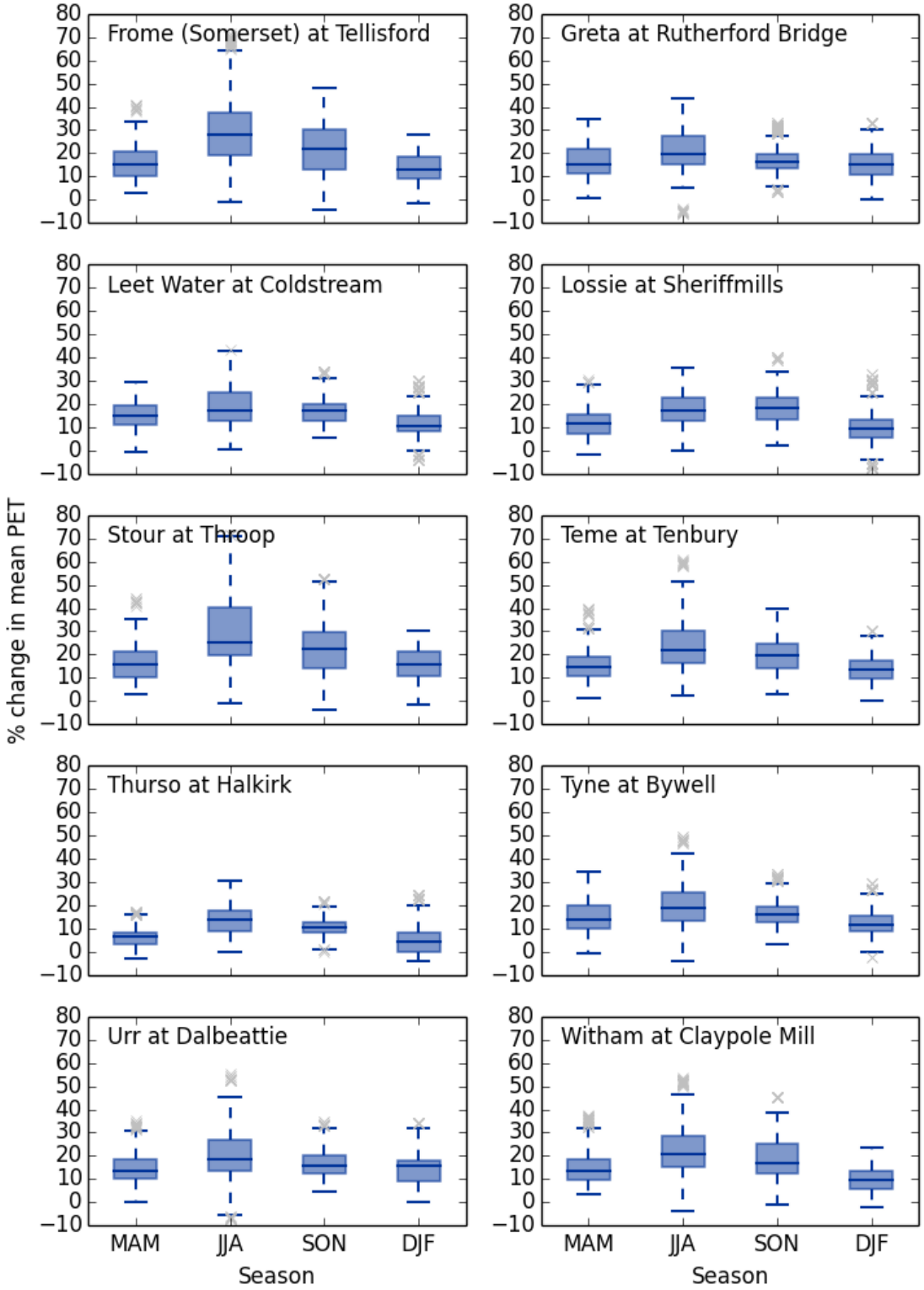



*Figure 6.11: % change in mean seasonal PET between baseline and 2050s scenarios.*

pattern which is consistent across the seasons. Lower rainfall and higher PET incur large reductions in flows across GB (see Figures 6.6 and 6.7). The magnitude of these reductions in flows is notable; the range of decreases in summer flow for the modelled catchments is from -12.5% to -42.0%. The Thurso at Halkirk experiences the smallest decrease in rainfall, as well as the smallest increase in PET, and so shows the smallest reduction in flow. Likewise, the Dart at Austins Bridge exhibits one of the largest reductions in summer flow and the largest increase in PET, which results in a very large decrease in flows of -41.1%. In terms of the distribution of flows, long upper tails combined with some skew towards the lower end of the range of summer flows are apparent. This reflects the continued possibility of observing lower probability occurrences of relatively high flows for the season in both control and future cases, although these could be considered to perhaps constitute atypical summer flows.

Moving to autumn, it can be seen from Figures 6.8 and 6.9 that rainfall tends to exhibit relatively small deviations from the control to future periods. Catchments in the northern half of GB tend to show increases in rainfall, whereas southern catchments generally exhibit small reductions. The magnitude of rainfall increases in the northern half of GB is typically less than 13%, whilst reductions in rainfall in the catchments to the south are -1% or smaller. PET is consistently higher in the future period relative to the control; all but one of the increases are in the range of 14.9% to 22.3%, with the Thurso at Halkirk being lower at 10.3%. The distributions of changes in PET are typically wider in autumn (and summer), suggesting a wider range of possible PET futures despite the general transition to higher PET. These patterns of rainfall and PET translate to decreases in autumn flow in the future period across the country. This is due largely to the significant soil moisture deficit created in the summer months. These reductions in flow are much higher in the south of GB than in the north, ranging from -2% for the Enrick at Mill of Tore in Scotland to -41.7% for the Stour at Throop in southern England. The exception to this pattern is the small increase in flows for the Thurso at Halkirk (5.2%). It is also apparent that the Dart at Austins Bridge tends to show lesser reductions in flow compared with the other catchments in the southern part of England and Wales, which is considered to reflect the persistence of this area's generally higher rainfall totals through into the future period. In addition, it should be noted when interpreting these changes in flow that relative (percentage) changes are considered here, such that absolute reductions in flow in autumn are notably larger than those in summer.

The deficit in flows in autumn is recovered to some degree in winter, principally as a result of increases in winter rainfall in the future period combined with more modest

increases in PET relative to those apparent in summer and autumn. Changes in winter flows range from -22.2 to 16.6%, with notable variability between catchments on the magnitude of change. Spatial patterns of the direction of change in flows between control and future periods are less clear in winter compared with summer and autumn. For example, of the catchments modelled in southern England, some experience relatively little change whilst one (Witham) experiences a -22.2% decrease in flows and another (Dart) shows a 16.6% increase in flow. This is likely to be related to the specific nature of geographical variations in rainfall and PET changes in winter and preceding seasons, as well as the influence of catchment properties affecting the soil moisture deficit particularly. Unusually, the Dart at Austins Bridge shows the largest increase in rainfall and smallest increase in PET (in other seasons this behaviour is normally seen in the Thurso at Halkirk, at the other end of the country) and as a result shows the largest increase in flow. On the other hand, the Witham at Claypole Mill shows the largest relative decrease in flows, which reflects a large soil moisture deficit developed during summer and persisting through autumn, as well as the influence of its small size (i.e. relative differences between small flows are large).

In terms of meteorological drivers, the differences between control and future periods in spring show ostensibly similar overall patterns to those apparent in autumn. In both cases, rainfall tends to exhibit relatively small deviations from the control to future periods. Catchments in the northern half of GB tend to have rainfall increases, in contrast to the southern catchments, which are generally associated with reductions in rainfall. In spring, the majority of selected catchments show changes in rainfall of around 5% or less, yet in autumn the magnitude of rainfall increases in the northern half of GB increases is generally higher. PET is consistently higher in the future period relative to the control in spring, by a similar magnitude to the overall increase in autumn, whilst the distributions of PET are again typically wider in the future period in spring. Flows in the future period in spring are lower everywhere except the Thurso at Halkirk. This catchment sees a very modest increase in flow of 0.9%, whereas the other modelled catchments show decreases in flow of up to -15.6% . The magnitude of flow reduction varies by catchment across GB; it could be suggested that the largest reductions tend to occur in England and Wales rather than Scotland, but latitude does not appear to explain the patterns as well as in autumn for example.

It should also be noted that, as annual flows have been shown in the previous section to be lower in the future period compared with the control period, it is apparent that the magnitude of increases in winter flows does not fully compensate for the typical reductions in spring, summer and autumn flows.

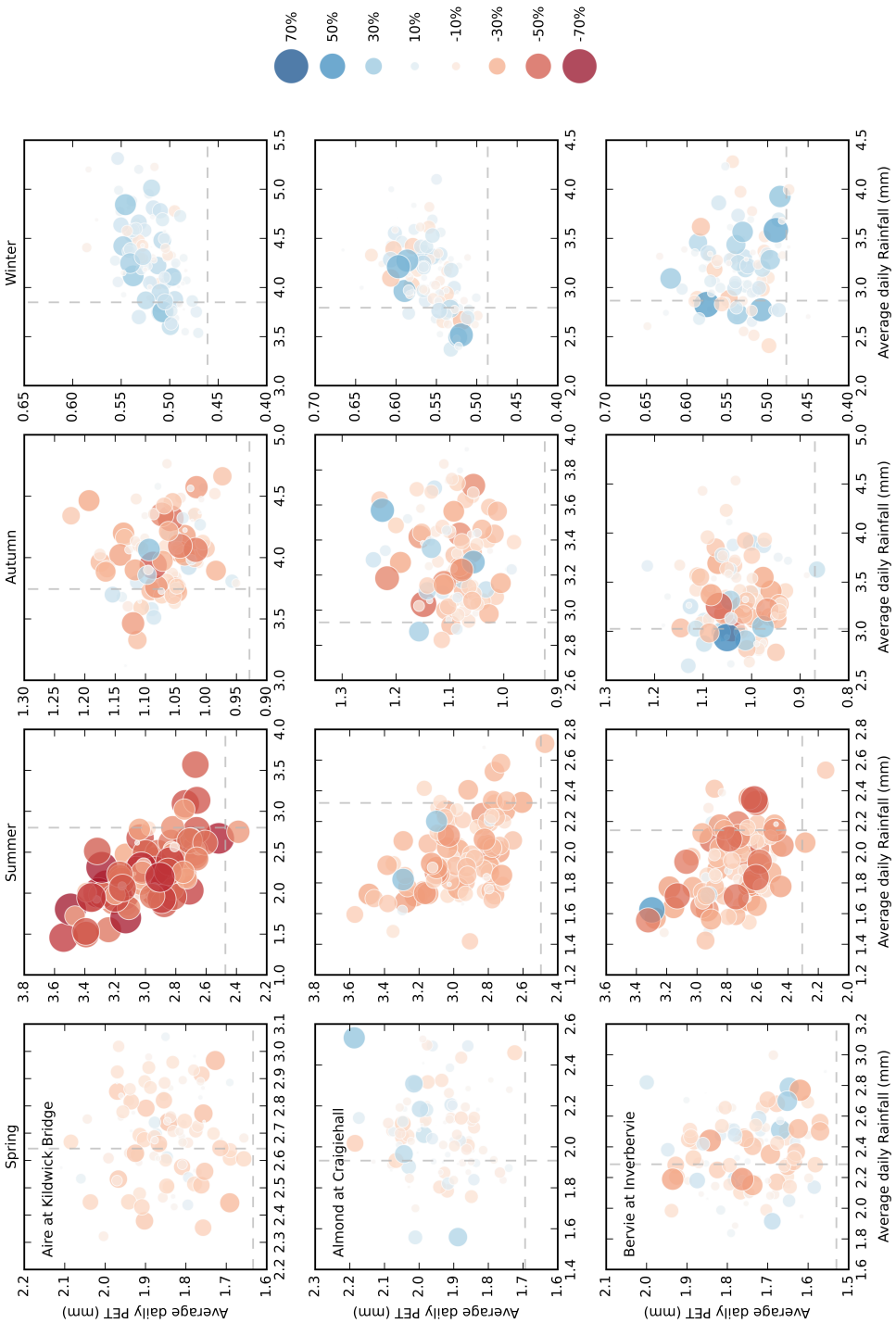



*Figure 6.12: Change in mean flows for each future run from the mean control flow, for each catchment and season, plotted against average daily rainfall and average daily PET. The size and colour of the dot indicates the direction and magnitude of the percent change in mean flow. The grey dashed lines indicate the mean rainfall and PET of the control runs.*

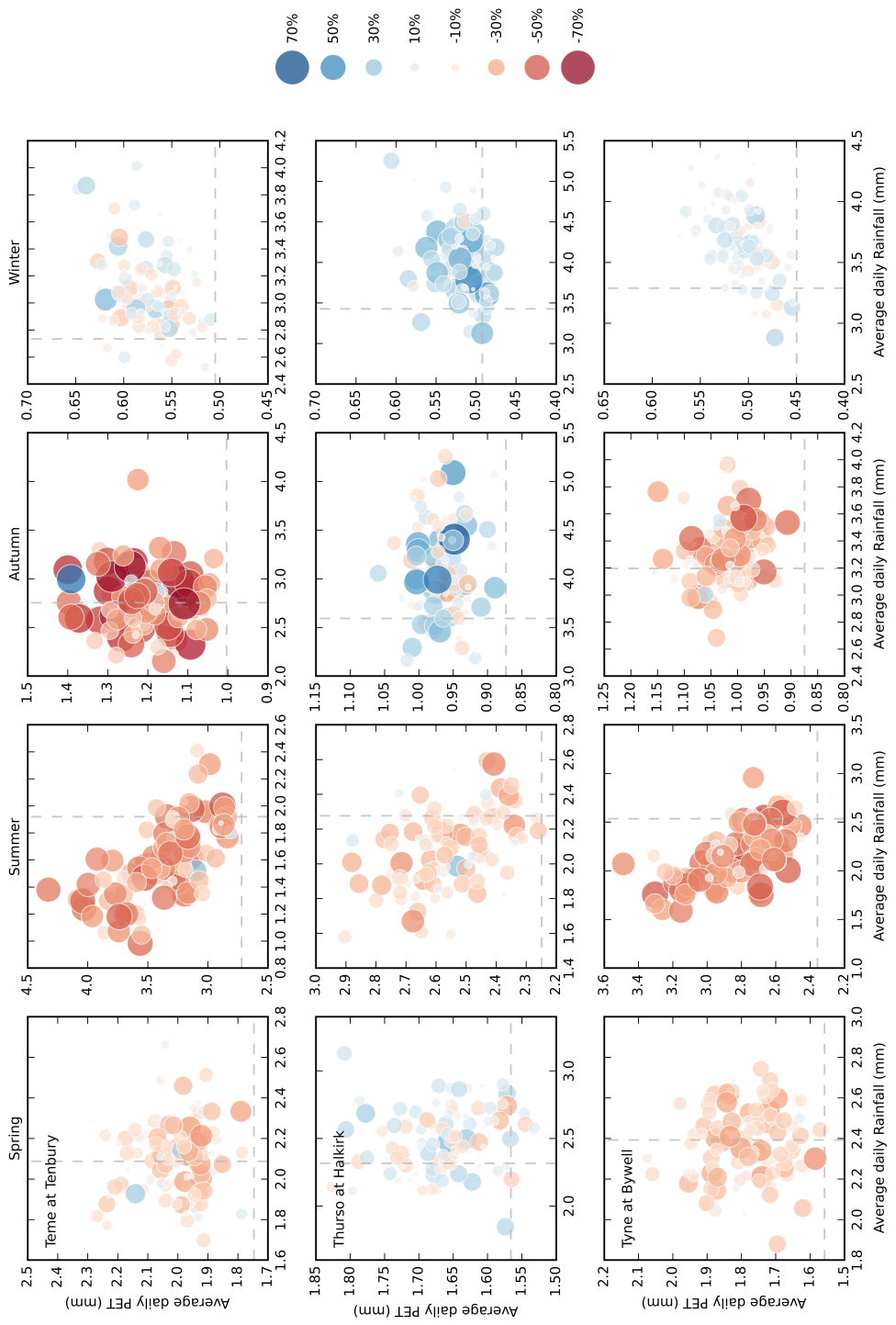



*Figure 6.13: Change in mean flows for each future run from the mean control flow, for each catchment and season, plotted against average daily rainfall and average daily PET. The size and colour of the dot indicates the direction and magnitude of the percent change in mean flow. The grey dashed lines indicate the mean rainfall and PET of the control runs.*

Figures 6.12 and 6.13 provide further detail on the simulated seasonal changes in future flows. For each catchment and season, the percentage change from the mean of the control runs to each of the 100 means of the future period runs is plotted against the mean seasonal rainfall and PET for each future period run (figures for all catchments can be found in Appendix B). The magnitude and direction of change in flow are represented by the size and colour of the dots, with larger dots signifying greater changes. Red dots indicate reductions in flows in the future period relative to the control, whereas blue dots denote increases in flow in the future period. The grey dashed lines indicate the mean rainfall and PET of the control runs. Note that the axes of the graphs are not a fixed scale across catchments or seasons, as this would make the results illegible. These graphs therefore allow for more detailed appraisal of the seasonal differences between control and future periods, as it is possible to see the changes in flow for each realisation of the future period and how these changes relate to input rainfall and PET overall and for each simulation.

These figures confirm the key patterns evident from Figures 6.6 to 6.11, namely large relative reductions in summer, autumn and spring flows, with generally modest increases in winter flows. The plots highlight that the largest relative reductions in flow tend to occur in summer, with lower percentage declines in autumn flows that in turn tend to be larger than the decreases simulated in spring flows. The closest relationships between average rainfall, PET and flows in the future period are observed in summer, which exhibits a strong negative correlation of rainfall and PET, such that combinations of lower rainfall and higher PET result in lower flows. Less clear relationships between these variables are apparent for spring and autumn. These seasons also exhibit slightly more variation in the direction of change in some catchments compared with summer, which indicates that some realisations of the future period result in higher average flows. However, these instances are only a minority in all catchments except the Thurso at Halkirk. Similarly, in winter there are also some realisations of the future period that show reductions in flow, particularly in the Teme at Tenbury, although in the majority of instances modest increases in flow are apparent. Interestingly, for some catchments there appears to be a positive correlation between higher rainfall and higher PET in winter, as is apparent for the Aire at Kildwick Bridge, the Almond at Craigiehall and the Tyne at Bywell for example.

Figures 6.12 and 6.13 also reflect some important points regarding the output of the weather generator. In each season and for all of the catchments simulated, average future PET for all realisations is greater than the mean control PET. The situation for rainfall varies with season. In winter, mean rainfall in the future period is higher than

in the control period in the majority of cases across all of the catchments displayed. However, in autumn a more mixed picture emerges. Most of of the catchments in Figures 6.12 and 6.13 show generally higher autumn rainfall - for example the Aire at Kildwick Bridge and the Almond at Craigiehall - whereas others show limited changes and scatter around the control period mean rainfall. This is particularly the case for the Teme at Tenbury, which additionally shows comparatively large relative reductions in autumn flow as a consequence. Spring rainfall in the realisations of the future period does not appear to show consistent deviation from the control period mean, whilst most realisations of future summers clearly exhibit less rainfall than the control period means for all catchments.

Given the strong relationship between PET and rainfall in summer, it would be expected that simulations with the greatest PET and lowest rainfall would show the largest decrease in flows i.e. the darkest red, largest circles would be seen in the upper left corner of the plots, with the dots becoming paler and smaller towards the bottom right corner. Equally, in winter it would be expected that the largest increases in flows would coincide with the wettest runs. However, it is interesting to note that this is not the case. Instead, the change in flow appears to bear little relationship to its PET and rainfall with respect to the other runs. This reflects the fact that the change in flow is influenced by more than just the ambient PET and rainfall. The magnitude (and in some cases direction) of change in a given realisation of the future period is highly dependent on catchment antecedent conditions, which are in turn a function of the specific spatial and temporal patterns calculated by the weather generator. The precise distribution, duration and intensity of rainfall events is likely to be a key part of this.

The implication of this is that similar values of mean rainfall and PET in a season can result in very different changes in flow. To demonstrate this, two individual runs for the future period for the Bervie at Inverbervie have been analysed. These runs have similar summer mean rainfall of 1.6mm and mean PET of 3.3mm. However, the change in flows in the two runs are +45.7% and -41.8%, i.e. both very large changes but opposite in sign. These runs can be identified on the graph for the Bervie at Inverbervie as the large blue dot overlapping a large red dot. When the context of these runs is considered, the causes of the differences becomes apparent. The run with increased summer flows received some of the highest rainfall in the preceding spring (point is an outlier relative to the main cloud of points), whereas the run with reduced summer flows was subject to spring rainfall totals more consistent with the majority of realisations. The run with higher summer flows also had higher spring flows, whilst the run with

| Station Name | Q99 | Q95 | Q90 | Q75 | Q50 | Q25 | Q10 | Q5 | Q1 |
|---|---|---|---|---|---|---|---|---|---|
| Aire at Kildwick Bridge | -21.68 | -19.3 | -19.28 | -17.43 | -23.83 | -10.85 | -2.2 | 1.34 | 8.48 |
| Almond at Craigiehall | -10.85 | -18.48 | -17.45 | -15.69 | -10.51 | -9.9 | -4.58 | -0.37 | 3.34 |
| Bervie at Inverbervie | -17.32 | -14.42 | -13.84 | -12.83 | -15.64 | -10.14 | -1.73 | 1.2 | 8.59 |
| Braan at Hermitage | -5.06 | -15.95 | -15.16 | -14.79 | -10.59 | -1.34 | 4.73 | 6.65 | 8.59 |
| Ceiriog at Brynkinalt Weir | -11.84 | -20.35 | -20.33 | -20.63 | -15.19 | -12.16 | -6.77 | -2.92 | 1.99 |
| Dart at Austins Bridge | -15.33 | -23.18 | -22 | -18.44 | -28.28 | -11.06 | 2.61 | 7.26 | 11.93 |
| Don at Doncaster | -23.13 | -18.73 | -17.22 | -14.87 | -16.9 | -22.08 | -13.69 | -8.73 | -3.65 |
| Eden at Temple Sowerby | -14.9 | -17.3 | -17.51 | -14.99 | -14.8 | -5.82 | 2.42 | 5.61 | 9.25 |
| Enrick at Mill of Tore | -6.36 | -15.93 | -15.96 | -12.04 | -7.59 | -3.87 | 0.49 | 0.82 | -0.38 |
| Frome at Yarkhill | -21.63 | -20.24 | -18.56 | -16.21 | -15.15 | -30.64 | -17.98 | -14.05 | -7.04 |
| Frome (Somerset) at Tellisford | -52.7 | -44.7 | -40.01 | -32.14 | -21.17 | -13.13 | -5.28 | -0.78 | 7.87 |
| Greta at Rutherford Bridge | -6.99 | -10.18 | -10.22 | -10.59 | -12.91 | -11.6 | -4.94 | -3.14 | -3.3 |
| Leet Water at Coldstream | -25.59 | -19.38 | -17.02 | -13.11 | -15.5 | -26.53 | -8.73 | -4.71 | -1.8 |
| Lossie at Sheriffmills | -16.78 | -17.12 | -16.92 | -15.31 | -10.06 | -7.65 | -4.13 | -3.13 | -1.12 |
| Stour at Throop | -31.47 | -33 | -32.47 | -28.88 | -23.42 | -19.52 | -6.88 | -1.61 | 6.83 |
| Teme at Tenbury | -22.1 | -19.6 | -17.63 | -17.23 | -26.76 | -23.25 | -12.45 | -6.24 | 0.69 |
| Thurso at Halkirk | -4.12 | -8.03 | -6.95 | -4.77 | -5.87 | 7.08 | 9.87 | 11.45 | 9.72 |
| Tyne at Bywell | -13.74 | -15.81 | -16.53 | -15.16 | -21.02 | -13.24 | -8.45 | -5.02 | -1.67 |
| Urr at Dalbeattie | -11.69 | -15.34 | -14.68 | -14.08 | -22.62 | -3.25 | 4.76 | 9.02 | 14.88 |
| Witham at Claypole Mill | -39.54 | -34.92 | -33.01 | -27.73 | -22.19 | -27.33 | -27.54 | -24.09 | -11.49 |

*Table 6.4: Summary table of % change in mean flow percentiles from control to future scenarios.*

lower summer flows had lower spring flows. This reflects how antecedent conditions in a catchment play a large role in determining current flows - in particular how the relationship between spatial and temporal patterns of rainfall, PET and soil moisture deficit crucially affects runoff generation. This can be seen again in autumn in the Teme at Tenbury. Two future runs give very different changes in flow of -58.9% and +57.9% in autumn (see blue dot overlapping red dot), despite both receiving average daily rainfall of 3mm and PET of 1.4mm. On examination of the behaviour of the two runs in other seasons it can be seen that the run showing large decreases in flow experienced some of the highest PET in spring and summer, whilst the run with increased flows in autumn experiences moderate PET and higher rainfalls in spring and winter.

### 6.7.3 Flow duration curves

Flow duration curves usefully characterise the flow regime of a catchment across the full range of flows. As climate change is considered likely to alter the flow distribution, with possible implications for numerous water management applications, it is therefore of interest to assess potential changes using flow duration curves. Figures 6.14 and 6.15

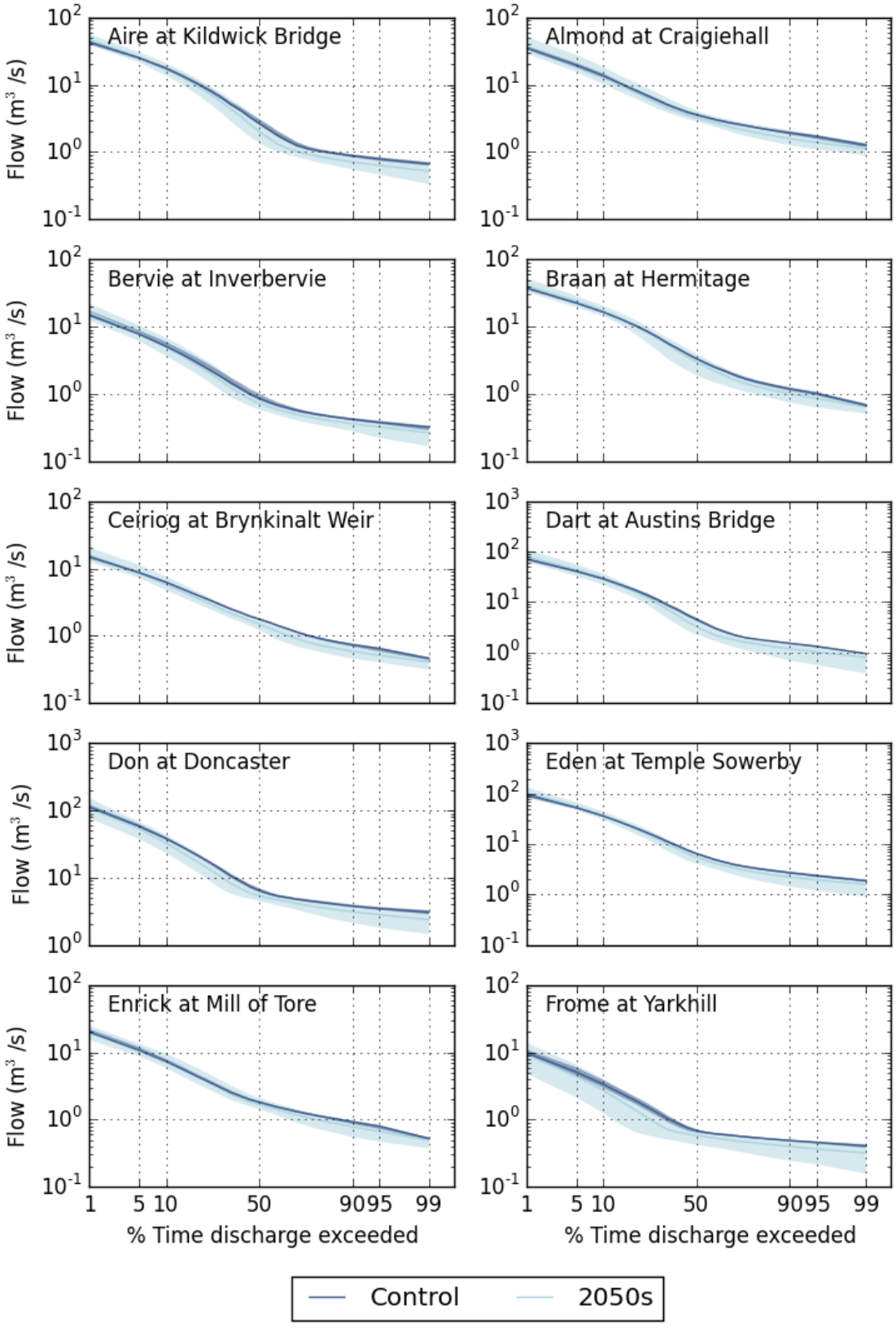



*Figure 6.14: Range of flow duration curves for control and 2050s scenarios. The solid lines represent the means of the distributions.*

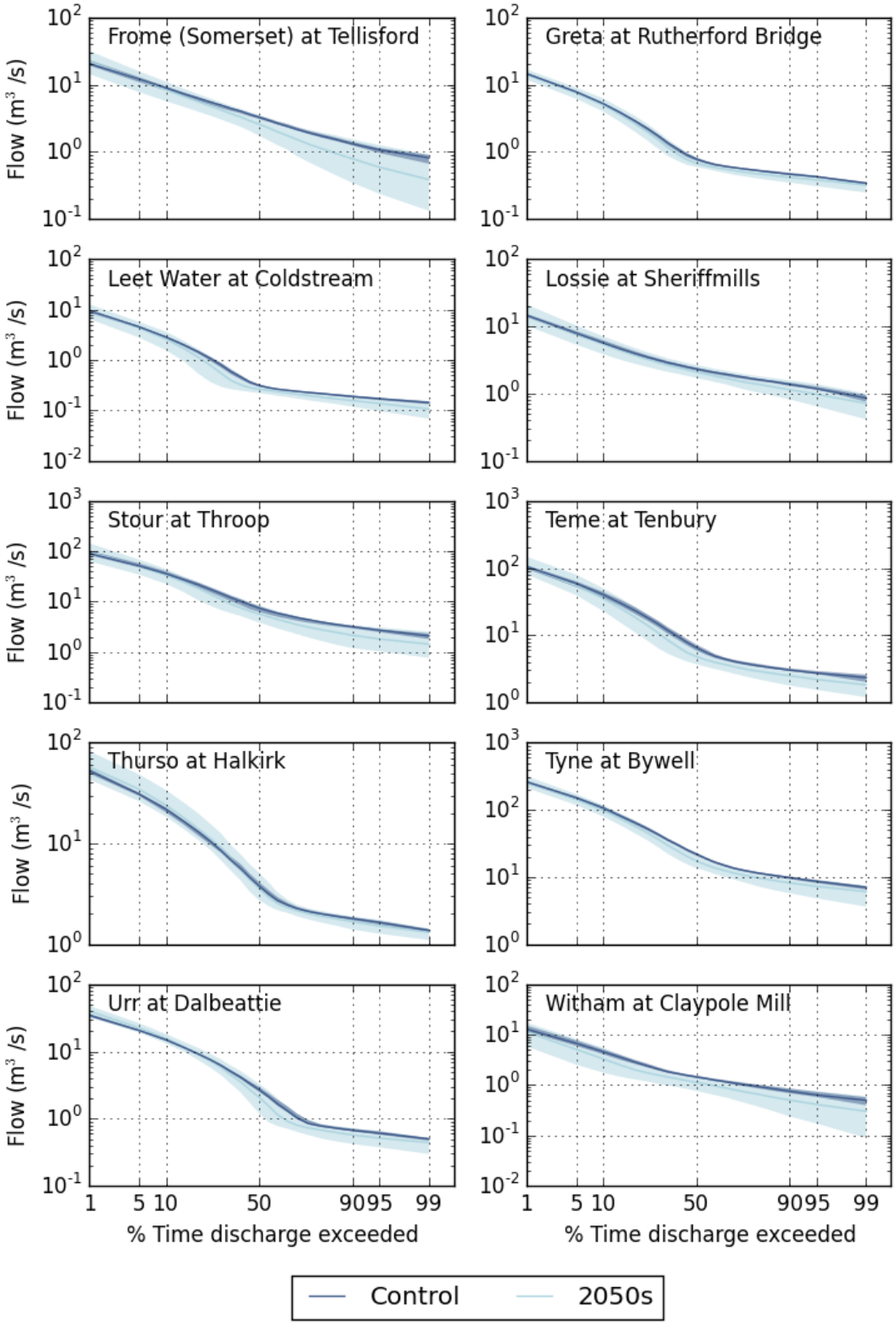



*Figure 6.15: Range of flow duration curves for control and 2050s scenarios. The solid lines represent the means of the distributions.*

show flow duration curves for the set of control and future scenarios for each catchment. Table 6.4 summarises the percentage change in flow at various flow percentiles from mean control to mean future. For flows lower than Q50, all of the catchments simulated experience reductions in flow relative to the control period. These reductions can sometimes be as large as 40% in the catchments of the Witham at Claypole Mill and the Frome (Somerset) at Tellisford. On the other hand, it is apparent that increases in flow occur towards the higher end of the flow range in a number of catchments, albeit to varying degrees. On average, flows increase the most in the Thurso at Halkirk, which is consistent with results from the previous sections describing its projected increases in spring, autumn and winter rainfall. Similarly, the Dart at Austins Bridge, despite being at the other end of the country, is the wettest catchment and so large projected increases in winter rainfall create greater high flows (increases in Q10, 5 and 1). In addition, the Bervie, Braan, Eden, Enrick and Urr all show slight projected increases in Q10 to Q1 as a function of increased winter rainfall. However, the remaining catchments show less changes or experience reductions in flow at almost every percentile on average. Yet even in these cases there tends to be a shift in the range of projected changes in flow at each percentile derived from individual realisations. For medium and lower flow percentiles, the maximum flows for the future runs tend to approximately equal the flows simulated in the control period, but at higher flow percentiles the maximum flows in the future period do generally exceed the control period flows to varying degrees. The projected differential changes in flows at the higher end of the flow range reflects how the properties of different characteristics modulate higher (winter) rainfall in the future period.

It is also of note that most catchments show a fairly narrow set of possible futures, with the exception of the Frome at Yarkhill, the Frome (Somerset) at Tellisford and the Witham at Claypole Mill. The Frome at Yarkhill and the Witham at Claypole Mill have similar patterns, where future flows decrease across the flow range but with quite a wide range based on the realisations conducted. The Witham at Claypole Mill is a small, dry catchment in the south east of England that is sensitive to increases in PET and decreases in rainfall. Both the Frome at Yarkhill and the Frome (Somerset) at Tellisford exhibit particularly large decreases in flow in autumn, a feature which is manifest in the flow duration curves through reduced medium and low flows. In addition, the Frome at Yarkhill and the Witham at Claypole Mill also exhibit decreases in flow in winter, which results in a wider range of flows in the upper part of the flow duration curve.

### 6.7.4 Peak flows

Assessing potential changes in peak flows is highly important given the extensive damage that can be caused by floods. A useful method for attempting to analyse possible changes at the high end of the flow range is the Gumbel plot (or 'extreme value plot'). The first step in constructing these plots is to analyse the annual maximum flow series. For each 30 year run, the 30 annual maxima are extracted and sorted in order of decreasing magnitude. The annual maximum flows are then ranked from 1 (highest flow) to 30 (lowest flow). The probability of exceedence for each flow, P(X), is then calculated using the Gringorten formula:

$$P(X) = \frac{(r - 0.44)}{(N + 0.12)}$$

where r is the rank and N is the total number of data values. The reduced variate for each flow value is then calculated:

$$F(X) = 1 - P(X)$$

$$ReducedVariate = -log(-log(F))$$

(Shaw et al., 2010) A Gumbel distribution is then fitted using L-moments, in this case using a Python package. The Gumbel reduced variate is therefore related to the return period of a flow through the probability of exceedence.

Figures 6.16 and 6.17 show extreme value plots for each of the modelled catchments. Table 6.5 summarises the change in the mean of the control to future distribution at each Gumbel reduced variate/return period.

Figures 6.16 and 6.17 show that half of the selected catchments (the Aire, Almond, Bervie, Braan, Ceiriog, Dart, Eden, Leet Water, and Urr) are subject to increases of between 8 to 23% from the mean of the control scenarios to the mean of the future scenarios at the 5 year return period (reduced variate of 1.5, see Table 6.5). However, considering the range of changes apparent from each of the 100 runs in the future scenario, the largest changes can be as great as 66% for the Eden at Temple Sowerby or 100% for the Bervie at Inverbervie. Mean increases in high flows in the future period calculated from all of the 100 realisations of the future period may therefore be significantly less than some of the specific increases in more extreme scenarios of rainfall increase produced by the weather generator.

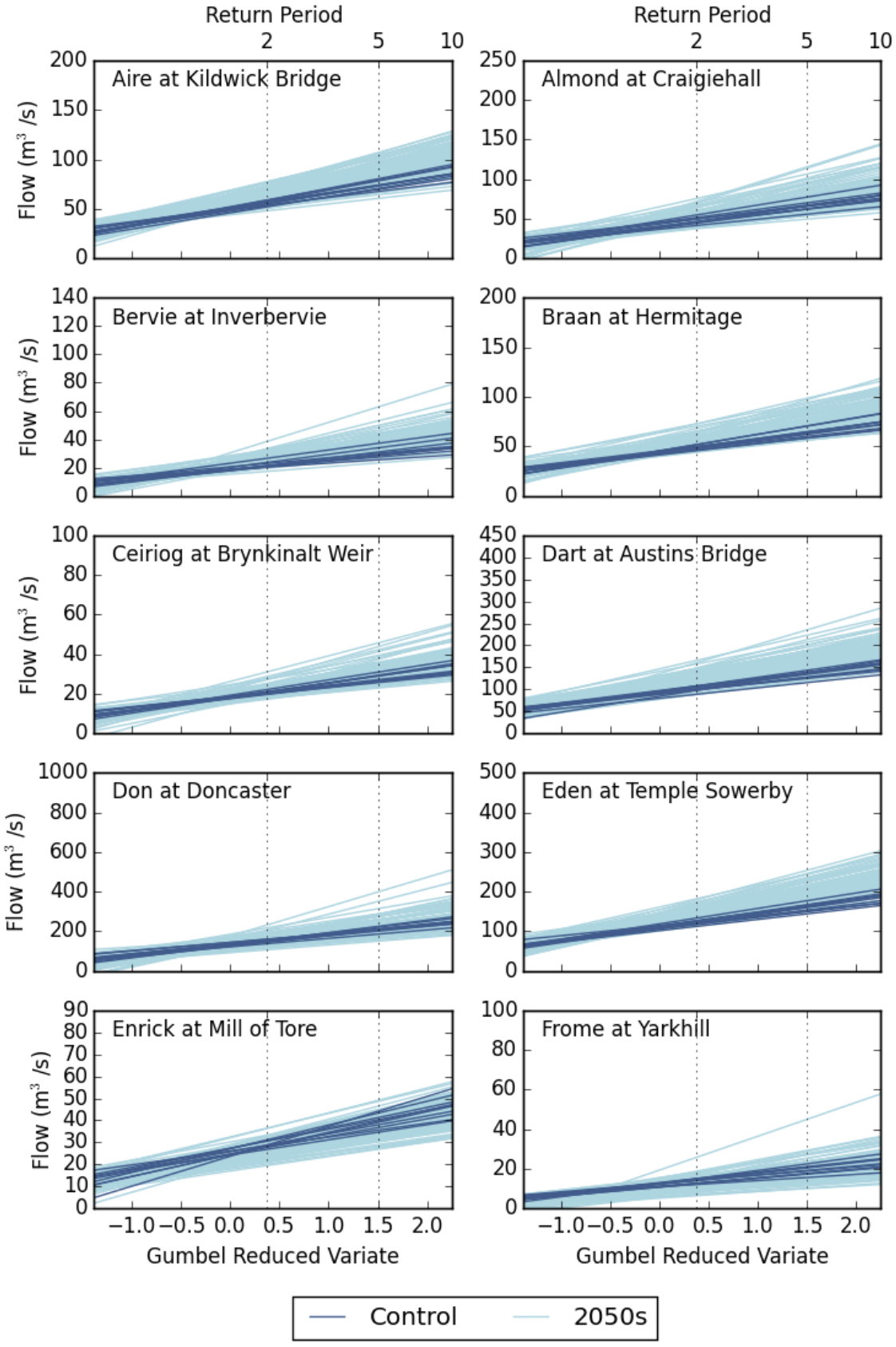



*Figure 6.16: Gumbel plots for control and 2050s scenarios. Each line represents a separate simulation.*

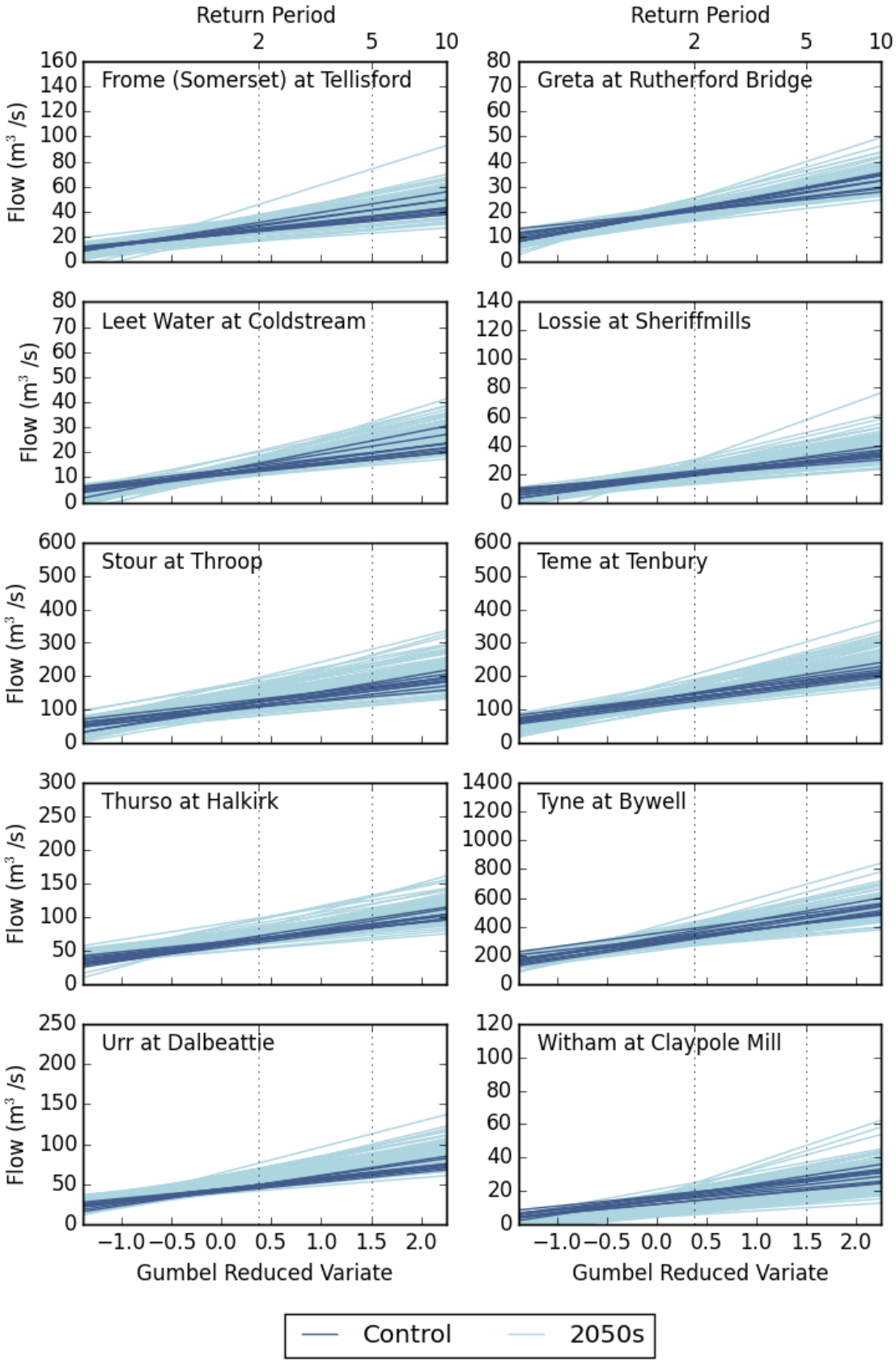



*Figure 6.17: Gumbel plots for control and 2050s scenarios. Each line represents a separate simulation.*

| Station Name | -1.53 | -0.83 | -0.48 | -0.19 | 0.09 | 0.37 | 0.67 | 1.03 | 1.50 | 2.25 |
|---|---|---|---|---|---|---|---|---|---|---|
| Aire at Kildwick Bridge | 3.07 | 9.4 | 11.36 | 12.59 | 13.54 | 14.35 | 15.11 | 15.85 | 16.65 | 17.63 |
| Almond at Craigiehall | -12.79 | 0.51 | 3.96 | 6 | 7.52 | 8.78 | 9.91 | 11 | 12.13 | 13.49 |
| Bervie at Inverbervie | -29.3 | -2.58 | 4.62 | 8.94 | 12.17 | 14.88 | 17.32 | 19.68 | 22.15 | 25.12 |
| Braan at Hermitage | 1.39 | 7.94 | 10.07 | 11.44 | 12.51 | 13.44 | 14.31 | 15.18 | 16.11 | 17.28 |
| Ceiriog at Brynkinalt Weir | -10.51 | -1.53 | 1.12 | 2.76 | 4.02 | 5.09 | 6.07 | 7.03 | 8.05 | 9.3 |
| Dart at Austins Bridge | 5.41 | 12.73 | 15.08 | 16.59 | 17.78 | 18.8 | 19.76 | 20.71 | 21.73 | 23.01 |
| Don at Doncaster | -29.54 | -9.84 | -4.6 | -1.47 | 0.86 | 2.81 | 4.57 | 6.26 | 8.03 | 10.15 |
| Eden at Temple Sowerby | -6.06 | 6.98 | 11.25 | 14 | 16.16 | 18.04 | 19.8 | 21.55 | 23.45 | 25.83 |
| Enrick at Mill of Tore | -8.27 | -6.79 | -6.38 | -6.13 | -5.94 | -5.78 | -5.63 | -5.49 | -5.34 | -5.16 |
| Frome at Yarkhill | -54.08 | -24.16 | -16.82 | -12.56 | -9.43 | -6.85 | -4.55 | -2.36 | -0.1 | 2.59 |
| Frome (Somerset) at Tellisford | -15.14 | -1.92 | 1.4 | 3.34 | 4.78 | 5.97 | 7.03 | 8.04 | 9.09 | 10.35 |
| Greta at Rutherford Bridge | -15.26 | -7.74 | -5.49 | -4.09 | -3.02 | -2.1 | -1.26 | -0.43 | 0.45 | 1.52 |
| Leet Water at Coldstream | -53.11 | -13.61 | -4.77 | 0.23 | 3.83 | 6.76 | 9.34 | 11.78 | 14.27 | 17.2 |
| Lossie at Sheriffmills | -57.43 | -17.1 | -8.59 | -3.86 | -0.48 | 2.24 | 4.63 | 6.87 | 9.15 | 11.8 |
| Stour at Throop | -3.6 | 4.21 | 6.38 | 7.7 | 8.7 | 9.54 | 10.3 | 11.04 | 11.82 | 12.76 |
| Teme at Tenbury | -25.9 | -9.31 | -4.12 | -0.84 | 1.71 | 3.91 | 5.94 | 7.95 | 10.11 | 12.78 |
| Thurso at Halkirk | 15.95 | 13.1 | 12.22 | 11.66 | 11.23 | 10.86 | 10.52 | 10.18 | 9.81 | 9.37 |
| Tyne at Bywell | -7.41 | -2.46 | -0.94 | 0.02 | 0.75 | 1.39 | 1.97 | 2.55 | 3.16 | 3.92 |
| Urr at Dalbeattie | 9.52 | 13.42 | 14.61 | 15.35 | 15.91 | 16.4 | 16.85 | 17.29 | 17.76 | 18.34 |
| Witham at Claypole Mill | -124.77 | -46.71 | -31.92 | -23.92 | -18.3 | -13.82 | -9.94 | -6.34 | -2.7 | 1.49 |

*Table 6.5: Summary table of % change in mean flow for a given reduced variate between control and future scenarios.*

For a number of the other catchments, the control runs generally lie in the middle of the range of flows in the future scenarios. This is the case for the Don, Frome (Somerset), Greta, Lossie, Stour, Teme, Thurso and Tyne. The changes in mean at the 5 year return period for these catchments range from 0-12%. Whilst the future high flows in these catchments tend to show overall increases in the future period, some catchments appear to exhibit overall decreases in the magnitude of flows at the return periods investigated. For example, the Enrick, Frome, and Witham show decreases in the means of the scenarios. In the most extreme future scenarios in these catchments, annual maximum flows can decrease by up to 50% relative to the control period.

The majority of annual maxima occur in the winter months and so the changes observed here are directly linked to the changes in winter flows, as examined in previous sections. It makes sense therefore that the Frome at Yarkhill and Witham at Claypole Mill show decreases in maximum flows, as the mean winter flow in these cases experience large decreases that are not apparent in other catchments. The Enrick at Mill of Tore shows a smaller decrease in mean winter flow. In addition, it should be noted that the timing of annual maximum flows used to construct the Gumbel plots does not change between control and future scenarios. Whilst high resolution regional climate models predict

increased intensity and frequency of summer extreme rainfall caused by convective cells, these recent developments are not incorporated into the weather generator statistics currently and are therefore not observed in the results presented here.

## 6.8 Discussion

This chapter has demonstrated a methodology for conducting a climate change impact assessment for Great Britain using the SHETRAN for GB system and provided preliminary results using the UKCP09 weather generator for a selection of catchments (for a given emissions scenario and time slice). The methodology offers a standard, easy to follow and largely automated process that is repeatable and amenable to application by others involved in climate change research, impact assessment or mitigation/adaptation work. The UKCP09 climate projections remain the standard set of scenarios for the UK and so their use is appropriate here. However, it is recognised that these projections are now dated and more recent, detailed RCM outputs have become available. As yet these updates have not been incorporated into a standard product for use in applications such as climate change impact studies. Fortunately, the weather generator has been designed with future updates in mind, such that no changes in overall methodology would be required to include updates in climate projections. Using 100 realisations of future climate goes part of the way to addressing the uncertainty surrounding climate change impact assessments; however, more detailed uncertainty analysis is definitely an important avenue for future research.

### 6.8.1 Technical issues

Whilst scripts have been created to automate the extraction of weather generator data, conversion to SHETRAN input and multiprocessing of model runs, it is not appropriate for use by non-experts who have stakes in the results, such as government agencies and local authorities for example. Indeed, the computing resources required to run the simulations in a reasonable time frame could by itself form a significant obstacle to direct use of the system by such parties. Clearly it is also necessary to be familiar with SHETRAN in case technical issues arise and to verify the correct behaviour of the system. For example, SHETRAN has generally only been tested and verified against historical climates and hydrological regimes to date; as such, its internal parameters (i.e. those relating to model convergence and stability as opposed to those directly controlling hydrological responses) are set accordingly. The implication of this is that,

when future rainfall values that exceed observed maxima are run through SHETRAN, the model can sometimes become unstable and fail. Some knowledge of SHETRAN is therefore required to set the relevant parameters appropriately. It follows of course that re-running failed models adds time to an already lengthy process in terms of both user intervention and additional computation, such that it would be useful to explore the potential for developing automated procedures to diagnose issues around run failure and attempt logical corrections. It is also worth noting that the main bottleneck in this methodology was the use of the weather generator, which was not originally designed for batch processing to support hydrological modelling of large numbers of catchments. However, it is anticipated that processing time will decrease with the next version of the weather generator, which will be designed to run on the cloud and so make use of parallel processing.

If stakeholders such as government agencies and local authorities did require a climate impact assessment tool, one could easily be developed by processing outputs into a form allowing non-experts to view results through different visualisations and key statistics. This chapter has presented some possible ways of analysing changes in flows that could be included in such a piece of software. The analysis has focused on changes across the flow range, looking at both the means and ranges of change apparent from multiple realisations of the future period. This forms a very large amount of information, and an interactive tool may be a particularly useful means for users to appreciate average changes and the range of possible futures according to the multiple realisations simulated. Such a tool could more easily contain outputs from large numbers of catchments than written reports, thereby providing flexibility for users to examine different results and develop their own understanding of the model outputs.

An assessment tool could also benefit from the fact that, in addition to simulated stream flows, SHETRAN produces a number of different outputs that are highly relevant for various water management applications. Amongst other variables, soil moisture, surface runoff, groundwater levels, actual evapotranspiration and snow cover are all calculated and may be written to output files at specified time steps during a simulation. The analysis of these changes is beyond the scope of this work, but it is ultimately one of the reasons why a physically based model could be particularly valuable in climate change impact studies. Enhanced process understanding is possible through analysis of the range of model outputs, whilst specific assessments of changes in both surface and subsurface water resources may be investigated in an integrated manner. Expanding this study to include more catchments would additionally give more scope for comparison of different catchment behaviours and their influence on water availability, something

which is more difficult to do with the 20 catchments examined so far, as the smaller sample size makes it more difficult to disentangle the effects of catchment properties from evaporation or rainfall changes for example.

### 6.8.2 Future changes

Although this study is a proof of concept for use of SHETRAN for GB rather than a complete climate change impact analysis, a number of interesting patterns emerge from the results. For example, it seems likely that GB could consistently experience a shift to higher PET rates compared with historical conditions. Particularly in summer and autumn, and to a lesser degree in spring, higher PET incurs reductions in flow relative to the control period (reflecting historical conditions). These declines in flow are not fully compensated for by the typically modest increase in winter flows apparent in most of the catchments simulated. As a result there is a generally decrease in annual flows in the future period for the mid-range emissions scenario and 2050s time slice selected. This pattern is of course based on simulation for a limited number of catchments. Although catchments with a range of geographical locations, climatological regimes, catchment characteristics and hydrological behaviours have been selected, changes in simulated flows in most cases tend to strongly reflect spatial patterns of meteorological inputs and their changes between the control and future periods.

### 6.8.3 Comparison to the Future Flows Project

| | Future Flows | SHETRAN for GB |
|---|---|---|
| Annual flow | +/-20% | -24% to +7% |
| Spring flow | -40% | -16% to +1% |
| Summer flow | -80% to +20% | -42% to -12% |
| Autumn flow | -80% to +60% | -43% to +5% |
| Winter flow | -20% to +40% | -22% to +17% |

*Table 6.6: Summary table comparing the change in flows from the Future Flows Project and this study.*

So far, national scale climate change impact analyses have not used the weather generator to derive meteorological inputs for hydrological modelling. As such, direct comparisons with other studies are not possible at present. However, some elementary and general comparison with other national scale studies can be made. Prudhomme et al. (2013) and Christierson et al. (2012) both conduct national scale impact assessments for

different time slices using different methods. Christierson et al. (2012) examine change in flows for the 2020s using the UKCP09 probabilistic changes in conjunction with the PDM model and Catchmod. The results of this study show a trend of decreased annual flows of up to 30% in the south of England. The Future Flows project Prudhomme et al. (2013) uses 11 ensemble members of the Future Flows Climate projections to derive change factors for meteorological inputs to the CERF generalised rainfall runoff model for the 2050s time slice. The main results of the Future Flows project and this work have been sumarised in Table 6.6. Mean annual flows were seen to change in a range of around +/-20 %. For about half of the scenarios, small increases in flow are observed, mainly in the south and east, although the majority of the scenarios show a reduction of mean annual flow of up to -40% in the west. In Scotland, simulations suggest limited changes in flow or reductions in mean annual flow. The results presented in this chapter agree to some extent with these conclusions. Reductions in mean annual flow in the west are certainly apparent in the results using the SHETRAN for GB system, but the increases in the south east are not simulated (although clearly only one catchment in this area has been modelled so far). Scottish catchments also show comparatively limited overall changes in this study, typically with small decreases apart from in the case of the Thurso at Halkirk, which exhibits increases in mean annual flow. The magnitude of changes in annual flows are similar in this study compared with those described above, with change in annual flows typically in the range -24 to +7%.

Seasonally, the Future Flows Hydrology scenarios show greater reductions in flow across most of the UK in spring compared with the SHETRAN results. Future Flows Hydrology results exhibit decreases in spring flow of up to 40%, whereas SHETRAN shows decreases in flow of up to 16%. The Future Flows results for summer indicate decreases throughout the country, although the range of changes is large and spans both positive and negative changes to flows (from +20% to -80%). The largest decreases according to the Future Flows results are found in the north and west of Great Britain. SHETRAN also shows consistent decreases in summer flows of up to 42% across the country, with the largest decreases found in the west but not necessarily in the north. Changes in autumn river flow according to the Future Flows project are dominated by reductions in flow in most cases, but there is a mixed pattern spatially, with percentage changes ranging from +60% to -80%. In contrast, SHETRAN shows a much clearer picture of very large reduction in flows in the south of England (of up to 43%) that become less severe in the north, with a small increase in flows in the northernmost catchment. The Future Flows study shows changes in winter flow to be mixed with drier, similar or wetter signals within a -20% to +40% range. SHETRAN results for the selected

catchments mainly show limited changes or modest increases in winter flows, with the exception of a notable reduction in flow in the future period for the Witham at Claypole Mill.

The results of the present study and Future Flows thus appear to be similar in overall terms, although it is hard to be definitive here, as the use of 11 RCM ensemble members in the Future Flows project produced a very mixed pattern of future changes in terms of magnitude, direction and spatial patterns. This comparison is also complicated by the different approaches to deriving input scenarios between the two projects. The Future Flows project uses the 11 RCM simulations from the HadRM3-PPE-UK ensemble directly to derive model inputs, after bias correction and downscaling from 25km to 1km resolution. In contrast, the UKCP09 projections only use these RCM runs to downscale a much greater number of GCM simulations (280). Various sources of uncertainty are then incorporated in the UKCP09 projections through a Bayesian framework and timescaling procedure, which is used to derive probabilistic projections (Murphy et al., 2009). UKCP09 is therefore a more comprehensive way of simulating future projections and, as probabilities are assigned, directions of change are more apparent. In addition, without using the same meteorological driving data, it is difficult to fully assess how the different hydrological model structures applied in Future Flows compare with SHETRAN as a basis for assessing climate change impacts. It is uncertain whether or not the calibrated parameters used in conceptual models are robust in the cases of notable shifts in future rates of PET and rainfall, in which case using a physically-based hydrological model may provide more robustness. However, further investigation of this issue is required, perhaps through direct comparison of SHETRAN and the models used in the Future Flows project.

# Chapter 7

# Improving the representation of geology

## 7.1 The need for proper representation of geology in hydrological models

A national physically based hydrological model for Great Britain has been set up and demonstrated to perform reasonably in terms of simulating streamflow hydrographs, flow regimes and catchment hydrographs across a range of basin types. However, assessment of the standard configuration described in Chapter 5 shows that catchments in which a significant proportion of streamflow is derived from groundwater are not consistently well modelled. As SHETRAN is a coupled surface-subsurface model that has been applied successfully to model regions with notable groundwater interaction previously (Adams and Parkin (2002), Parkin et al. (2007), Koo and O'Connell (2006)), it seems reasonable to expect that better performance could be attained in catchments where groundwater influences are significant. Starting with the premise that the variably saturated subsurface flow descriptions in SHETRAN are appropriate for modelling these catchments in principal, the causes of comparatively poor performance in some catchments could be more likely to be related to model inputs and parameters.

As discussed in Chapter 3, the model setup has been based on freely available datasets with full coverage of Great Britain. With respect to representation of the subsurface, the following limitations are associated with this approach:

- A single bedrock layer
- Bedrock has a uniform thickness of 20m

- Spatial distribution of hydraulic properties assigned according to the BGS hydrogeology map. This map has only 4 categories of aquifer: high productivity, moderate productivity, low productivity and no groundwater
- Hydraulic properties assigned from previous modelling experience, not from measurements or the literature
- No representation of any kind of geological complexity, including stratigraphy, lenses, faults etc.
- Surface extent of geological units does not reflect subsurface extent
- No representation of superficial deposits.

This chapter investigates the possibility of overcoming some of the limitations of the initial approach with respect to defining the layer structure and properties of the subsurface in the national modelling system. In particular, a high quality 3D geological model covering much of GB recently developed by the British Geological Survey (BGS) (Mathers et al., 2014) is examined and incorporated into SHETRAN, in order to begin the process of implementing a more accurate description of the subsurface and testing the implications for model performance. This novel integration of geological and hydrological models could yield significant benefits in the future, for example in integrated flood and water resources management. The discussion below introduces the geological model and how it has been coupled with the national SHETRAN modelling system, before analysing results of initial simulations of the integrated system.

This chapter firstly discusses the datasets available for incorporation into SHETRAN for GB to improve the representation of geology in the modelling system and the processing undertaken to achieve this. A set of sensitivity tests using various combinations of these datasets are then outlined and the results are discussed.

## 7.2 The BGS 3D geology model

The 3D geological model of Great Britain developed by the BGS was created by first constructing a series of geological cross-sections covering the whole country (Mathers et al., 2014). Each cross section was created in the geological modelling software GSI 3D (Kessler et al., 2009), based on existing map cross sections which in turn were based on seismic data and borehole information as well as the national 2 dimensional geological map (Mathers et al., 2012). 121 cross-sections were constructed, with a total length of over 20000 kilometres, and the depth of the cross-sections varies between 1.5 and 6

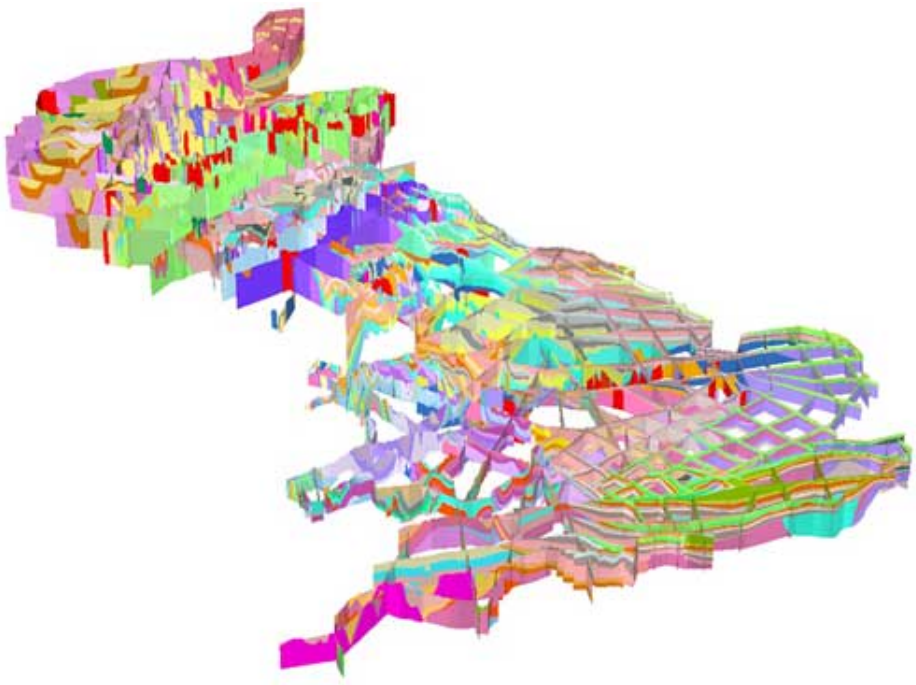


*Figure 7.1: BGS fence diagram on which the 3D geological model is based.*

kilometres. The cross-sections were then brought together to form a fence diagram of the whole of Great Britain (see Figure 7.1). This fence diagram was then interpolated to produce a three-dimensional static model of the geology (Watson et al., 2015).

This geological model provides a significant advance over the data used in the standard SHETRAN model configuration, for which the only information available on a national scale at the time of configuration was the 2D national geological map, which does not give any indication of layer depth or the order of layers of different strata in the subsurface. However, there are some limitations associated with the BGS 3D geological model to note. Most importantly, no hydraulic parameters are associated with this model and so parameters had to be assigned from previous modelling experience. Secondly, at present the model only comprises the major geological units and not the finer detail of areas. Faults are not individual objects in this model; they are simply shown by a break or offset in the geological units, although future versions of the fence diagram will support a more flexible representation of faults. In addition, thicknesses and layers of superficial deposits are not included within the 3D geological model, but they are available as separate map layers from the BGS resulting in potentially inconsistent thicknesses. The digital elevation model used is consistent with that used in SHETRAN for GB but due to different resampling methods used in SHETRAN for GB and the BGS data, the elevations at a given point may vary. Geology for this model also does not cover the whole country, as the 3D model was calculated from the base

Permian rocks upwards, while in some sections false horizontal bases were created to limit the model extent vertically. Furthermore, there has not yet been a full uncertainty study undertaken to assess the model; it should be noted that each cross-section has been interpolated only once by one of several experts and so the lines of these sections are very much interpreted. Cross-section lines could be drawn in various positions so that the final structure represents one of many possible realisations, and the model does not reflect the density of hard information used such as borehole logs and geophysical surveys. It is therefore difficult to tell how many boreholes have been used to construct each cross-section and thus in which locations data may be dense or sparse.

## 7.3 BGS superficial deposits data

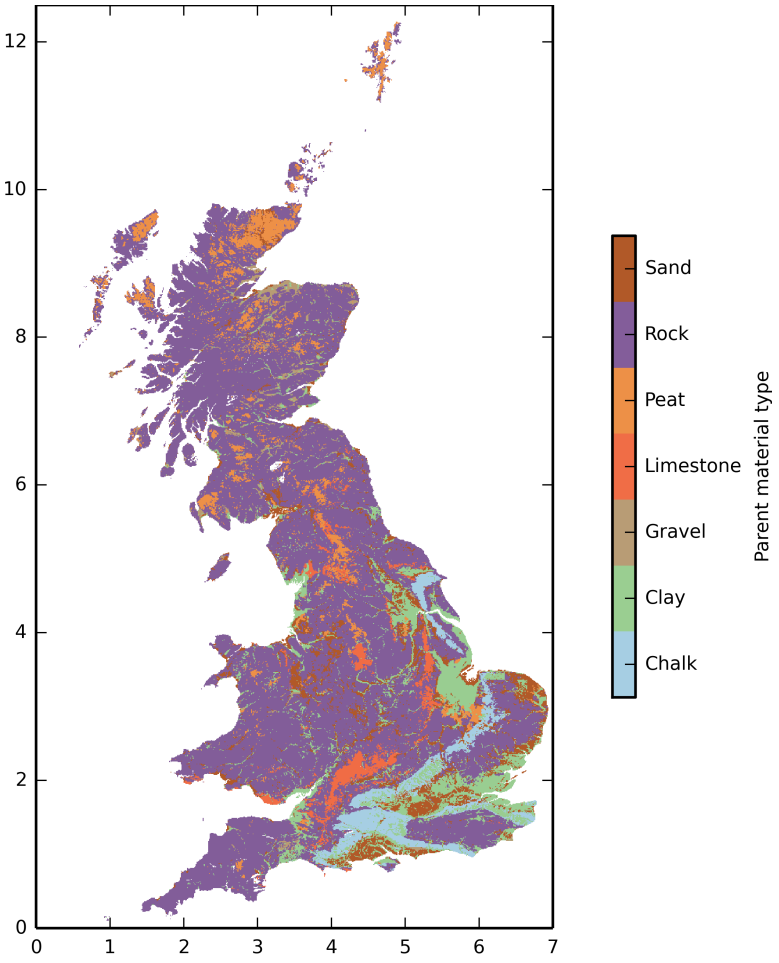



*Figure 7.2: Map of simplified parent material data.*

Superficial deposit information would be a useful addition to the SHETRAN for GB modelling system. The datasets available for use in this project are the BGS Par-

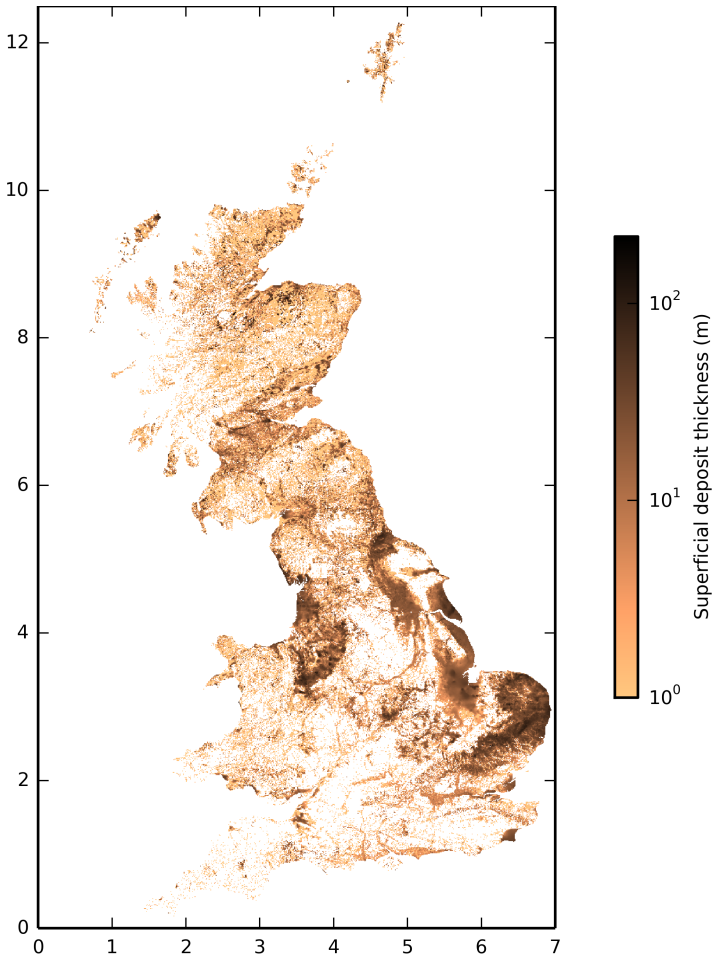



*Figure 7.3: Map of superficial deposits thickness.*

ent Materials Map (Lawley, 2009) and the The National Superficial Deposit Thickness Model (Lawley and Garcia-Bajo, 2009). The distribution of superficial deposit types was taken from the Parent Materials map. A parent material is defined as a geological deposit from which a soil develops; it is the closest geological deposit to the ground surface and therefore this map only provides information about the top layer of superficial deposits and not any deeper layers overlying bedrock. The soil Parent Materials map is derived from the BGS 1:50000 scale geological map of Great Britain, with extra information taken from the BGS rock classification scheme volumes 1 to 4 (Gillespie and Styles (1999), Robertson (1999), Hallsworth and Knox (1999), McMillan and Powell (1999)). The map indicates the most likely parent material type for a given grid square, i.e. the parent material most likely to be encountered at that location if the overlying soil were removed. The map was provided as a shapefile containing 319 different parent material types. Unfortunately none of these parent material types are currently associated with any assigned parameters suitable for a hydrological model,

| Parent material | saturated water content | residual water content | saturated conductivity | alpha | n |
|---|---|---|---|---|---|
| Chalk | 0.3 | 0.2 | 0.1 | 0.1 | 5 |
| Clay | 0.3 | 0.2 | 0.001 | 0.1 | 5 |
| Gravel | 0.3 | 0.2 | 0.01 | 0.1 | 5 |
| Limestone | 0.3 | 0.2 | 0.1 | 0.1 | 5 |
| Peat | 0.766 | 0.01 | 8 | 0.013 | 1.2039 |
| Rock | 0.3 | 0.2 | 0.001 | 0.1 | 5 |
| Sand | 0.403 | 0.025 | 60 | 0.0383 | 1.3774 |

*Table 7.1: Properties assigned to the parent material types.*

and so insufficient information was available to justify retaining such a large number of separate material types. The categories were therefore simplified to 7 different parent material types outlined in Table 7.1, with initial parameters assigned on the basis of previous modelling experience.

The Advanced Superficial Thickness Model (ASTM) of superficial deposit thickness was chosen to provide depths of the superficial deposits. This model indirectly uses borehole data and map information and a digital elevation model, interpolated using the Natural Neighbour method (Sibson, 1981) to create a smooth national dataset of deposit thicknesses. While the interface between soil and parent material can be either be sharp or a continuum, for the purposes of this project the soil layer from the European soil database map was taken to have a clear boundary at the base specified. The thickness of the superficial deposits was then determined from the superficial deposit map minus the thickness of the soil. The data was provided as a 50m ASCII grid and was resampled to 1 kilometre grid size to be consistent with SHETRAN for GB.

## 7.4 EA Transmissivity data

The hydraulic properties of the chalk are known to vary significantly in space, which has a very large influence on groundwater flow patterns and interactions with surface water drainage systems (MacDonald and Allen, 2001). Currently no national map or gridded datasets are available that describe this spatial variation, such that a substitute was sought in the form of the transmissivity distributions utilised in the Environment Agency's regional MODFLOW groundwater models of chalk areas (Shepley et al., 2012). The transmissivity distributions in these models are the result of bringing together available data from pumping tests reported in the Aquifer Properties Manual (MacDonald and Allen, 2001) and other sources, as well as expert judgement of leading hydrogeologists. While some degree of calibration of transmissivity has been

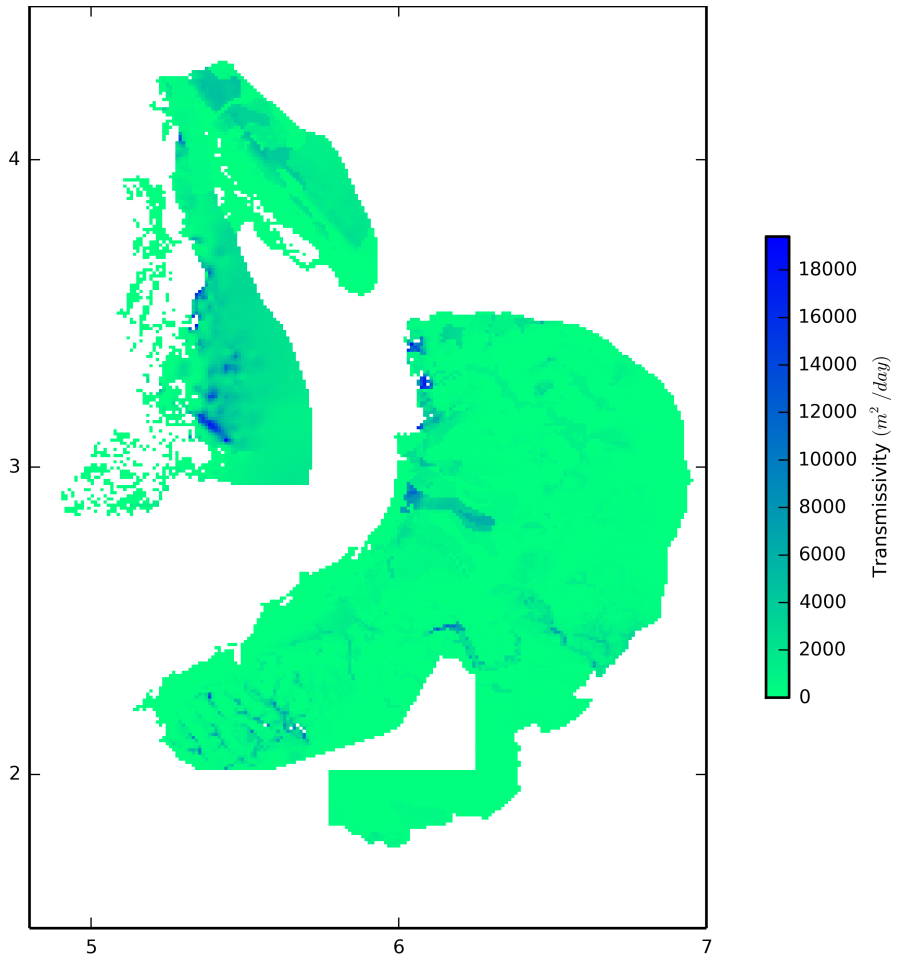



*Figure 7.4: Map of transmissivity values from the regional EA groundwater models.*

undertaken to provide robust results in the modflow models, the original data has been closely respected (Shepley et al., 2012).

The maps of transmissivity made available by the Environment Agency required conversion to hydraulic conductivity and some additional processing for input to SHETRAN. The first step in this was to resample the transmissivity maps to a one kilometre grid square in line with the resolution used in this study. Then, as each of the regional groundwater models contains several MODFLOW layers to represent chalk strata that each have an associated transmissivity, the total transmissivity for each grid square was calculated by adding together the maps for each chalk layer. The total thickness of the chalk for each grid square was then calculated on the basis of the 3D geological model layers and the transmissivity was divided by this thickness to give an approximate hydraulic conductivity for the chalk as a whole. As no water table elevation data was available to determine the typical saturated thickness of the chalk, it was necessary to calculate hydraulic conductivity in this simplified way. This approach is therefore

clearly approximate and does not account for the widely observed decrease of hydraulic conductivity with depth (Shepley et al., 2012), but as a first step in representing lateral variation in chalk hydraulic properties it is considered useful. An obvious next step and further work would therefore be investigating different simple conceptual models of the chalk, such as decreasing conductivity with depth or having a band of high conductivity near the top of the chalk.

Although the Environment Agency has groundwater models for all of the principal aquifers in the country, the focus of this chapter is to improve particularly the performance of the chalk catchments, therefore this was the only data obtained and tested.

## 7.5 Incorporation of information and modifications to SHETRAN- new software

The text file output from GSI3D is organised by geological layers, beginning with the top rock layer (as specified in a separate Generalised Vertical Section (GVS) file) and providing coordinates of each 1km grid square and the elevation (in metres above sea level) of the base of that layer. Information for all underlying layers is then presented sequentially. The surface elevations from the DEM can also be extracted in this way. A number of processing steps were required to integrate the BGS 3D geological model output in this format into SHETRAN, which requires the user to specify the subsurface (including soils) as columns, one per grid square. These columns are recorded in a SHETRAN library file, which contains the equivalent of the GVS file but with hydraulic parameters specified in addition to the depth and type of each layer in each of the columns. The locations of each column are recorded in an ASCII map provided as input to SHETRAN.

The conversion process has been automated by a Python script that requires the GSI3D output file, a GVS file with hydraulic parameters and an existing soil map to create input for SHETRAN. An alternative approach could be to output all of the required data from GSI directly, but the limitation of this method would be poor representation of soils, which are difficult to incorporate properly in the geological model due to their relative thinness. Instead, soil data from the European Soils Database outlined in Chapter 3 are incorporated in the subsurface description for each column by first removing the top part of the uppermost layer of geology according to the soil thickness. The required soil type is then inserted and the total thickness of rock reduced relative to the GSI3D output by a thickness corresponding to the depth of soil. When superficial

deposits are included, the thickness and type of deposit are read in from the national superficial deposits maps first and the corresponding thickness is removed from the top of the bedrock, allowing the superficial deposits to be inserted. The soil is then added by removing the required depth of superficial deposits and replacing this with the required soil types.

Modifications to SHETRAN were made to allow it to simulate deep aquifers (up to 360m) and utilise up to 20 different layers in the subsurface column. The structure of the library file was reformatted to resemble the GVS file more, so that layer types and properties are now detailed at the top and the column structures are described below. This has the advantage of producing tidier library files, in which the hydraulic properties of a given layer can be changed once instead of every time it appears in a column. These modifications were incorporated in a new, 64-bit version of SHETRAN.

It should be noted that the BGS geological model only contains information on the structure and stratigraphy of subsurface formations, not the hydraulic properties associated with different geological layers. The specification of parameters needed in hydrological or hydrogeological simulations is therefore an additional task. The approach taken to specifying these parameters is described below for each of the sensitivity tests undertaken, from which it became apparent that further work on estimating subsurface properties is required to profit most from the more accurate structural geology added into the SHETRAN modelling system.

## 7.6 Different model structures

A series of sensitivity tests with respect to subsurface representation were conducted on the basis of available data, including the 3D geological model described above. The tests examined increases in complexity from retaining a 20m thick aquifer from the standard run described in Chapter 5, through to full 3D representation of geology including superficial deposits.

Figure 7.5 shows the different sensitivity tests examined in this chapter. Test 1 is the standard run of SHETRAN for GB described in Chapters 3 and 5, in which geology below the soil layer is simulated as a 20m aquifer assigned one of four aquifer types that reflect hydrogeological productivity in general terms. This test forms the reference against which other changes to geology in SHETRAN are compared. Test 2 retains the 20m aquifer thickness, but instead of using four general aquifer types, this test incorporates transmissivity parameters for the chalk only, taken from the Aquifer Properties

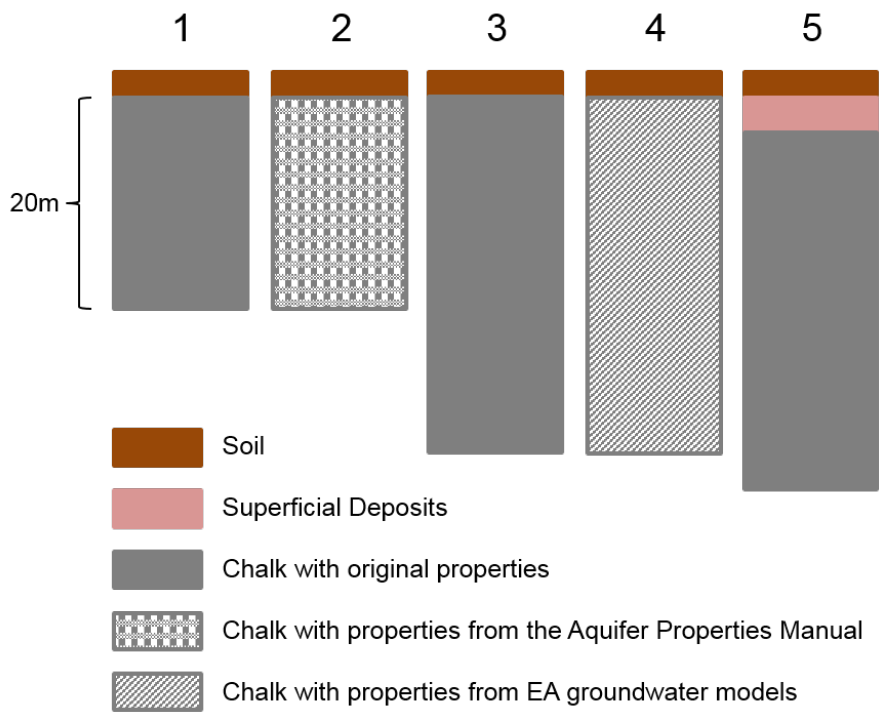



*Figure 7.5: Schematic of different sensitivity tests explored in this chapter.*

Manual (MacDonald and Allen, 2001). Test 3 adopts a different approach by implementing the BGS 3D geology model, while Test 4 extends this using the 3D geological model but additionally applying transmissivity values from EA regional groundwater models in chalk areas. Test 5 is based on Test 3, but additionally incorporates information about superficial deposits overlying the solid geology.

The analysis has been conducted only for catchments that experience a change in their setup relative to the standard run and so the standard run statistics vary from test to test.

### 7.6.1 Simple BGS hydrogeology map

| Aquifer type | saturated water content | residual water content | saturated conductivity | alpha | n |
|---|---|---|---|---|---|
| Highly Productive Aquifer | 0.3 | 0.2 | 0.1 | 0.01 | 5 |
| Moderately Productive Aquifer | 0.3 | 0.2 | 0.01 | 0.01 | 5 |
| Low Productivity Aquifer | 0.3 | 0.2 | 0.001 | 0.01 | 5 |
| No Groundwater | 0.3 | 0.2 | 0.0001 | 0.01 | 5 |

*Table 7.2: Properties assigned to the simple geology types.*

The model was initially set up with a uniform 20m thickness of bedrock classified as highly productive aquifer, moderately productive aquifer, low productivity aquifer and

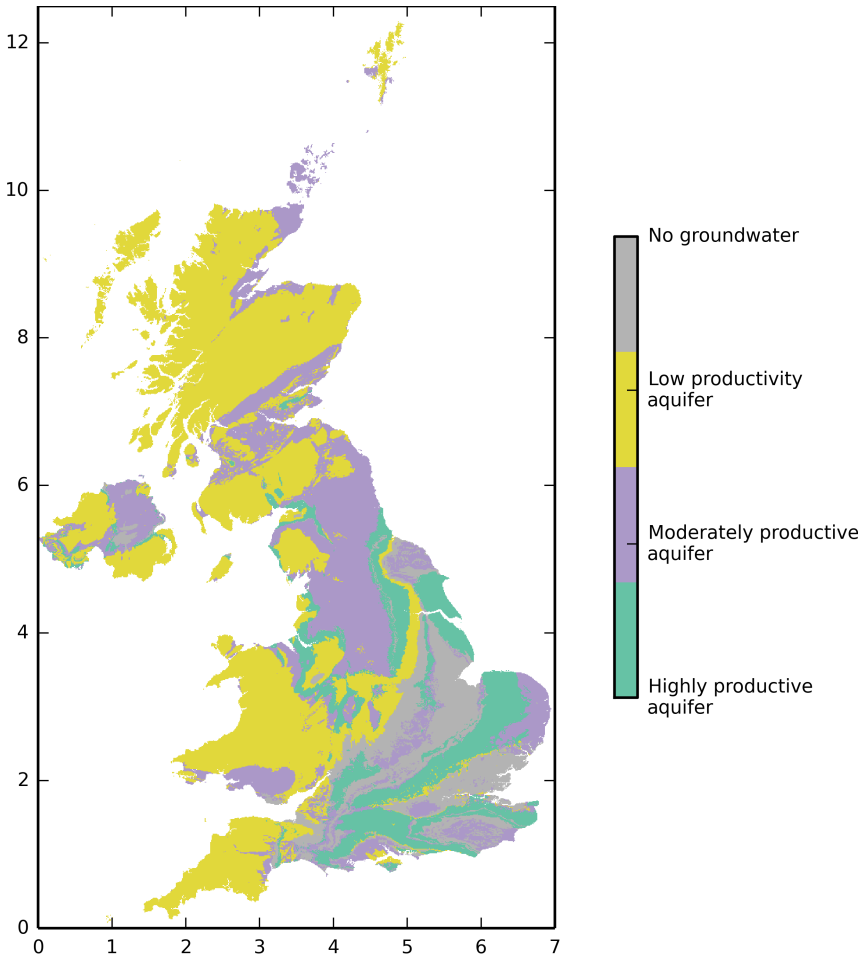



*Figure 7.6: BGS hydrogeology map.*

no groundwater according to the BGS hydrogeology map (see Figure 7.6). Properties were assigned to each of these categories on the basis of properties used for each type of aquifer in previous work and therefore considered relatively standard for SHETRAN (Birkinshaw, 2011). These properties are summarised in Table 7.2. The results from applying this configuration have already been discussed in Chapter 5 and are not repeated in detail here. However, it is worth noting again that reasonable results were produced across much of the country with the notable exception of catchments underlain by major chalk aquifers. This forms the benchmark against which subsequent sensitivity tests are evaluated.

### 7.6.2 Aquifer Properties Manual parameters

One of the clear simplifications in the standard run (corresponding to the initial test in this chapter) is that parameters are applied uniformly across the large chalk regions, which is not really consistent with the highly heterogeneous nature of the chalk and

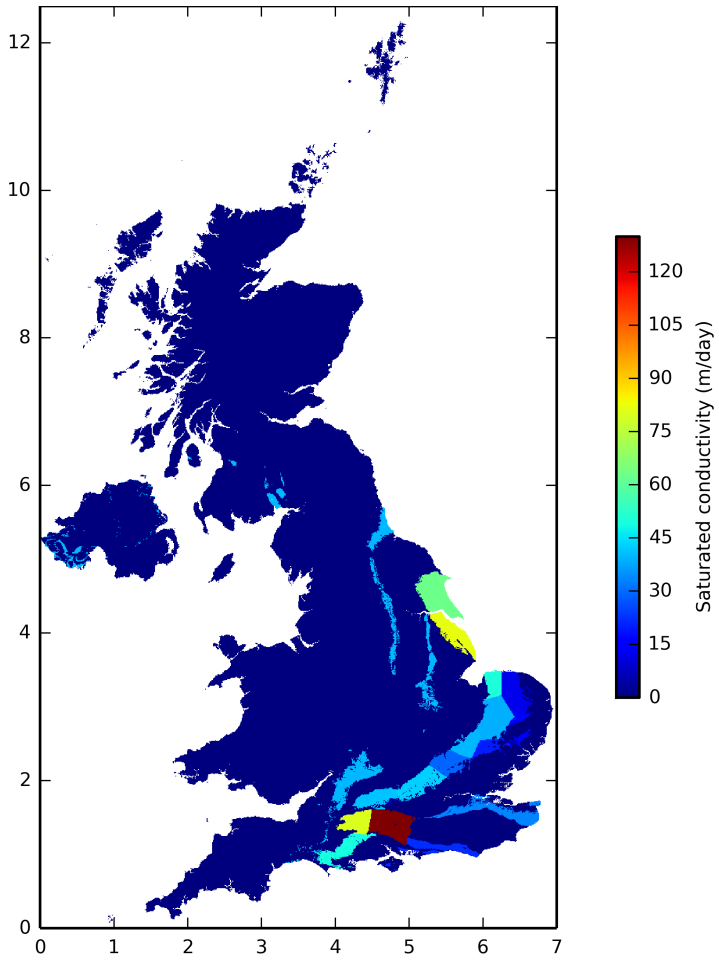



*Figure 7.7: Map showing saturated conductivity (m/day) values with chalk divisions from the Aquifer Properties Manual.*

the known patterns of spatial variation in transmissivity that span several orders of magnitude (Shepley et al., 2012). The most comprehensive source of information on the hydraulic properties of the chalk is the Aquifer Properties Manual (MacDonald and Allen, 2001). This manual represents the culmination of significant efforts to compile diverse data from pumping tests and laboratory experimentation, which is analysed and used to provide regional characterisations of the chalk (and other) aquifers across the UK. The sub-regional differences in chalk transmissivity reported in the Aquifer Properties Manual were therefore utilised to create hydraulic property maps for input to SHETRAN that show more spatial differentiation than used in the initial approach (see Figure 7.7 compared to Figure 7.6 and Table 7.3 compared to 7.2). This represents a likely improvement in the quality of information provided to SHETRAN compared with the standard run, but the properties of the chalk are known to vary over significantly smaller scales. As such, this run is intended to evaluate whether performance

| Geology code | Aquifer type | Saturated conductivity (m/day) |
|---|---|---|
| 1 | Highly productive aquifer through pores | 0.1 |
| 2 | Highly productive aquifer through cracks | 40 |
| 3 | Moderately productive aquifer through pores | 0.01 |
| 4 | Moderately productive aquifer through cracks | 0.01 |
| 5 | Low productivity aquifer through pores | 0.001 |
| 6 | Low productivity aquifer through cracks | 0.001 |
| 7 | No groundwater | 0.0001 |
| 8 | Highly productive aquifer through cracks | 62.5 |
| 9 | Highly productive aquifer through cracks | 82 |
| 10 | Highly productive aquifer through cracks | 50 |
| 11 | Highly productive aquifer through cracks | 12.5 |
| 12 | Highly productive aquifer through cracks | 39 |
| 13 | Highly productive aquifer through cracks | 15.75 |
| 14 | Highly productive aquifer through cracks | 40 |
| 15 | Highly productive aquifer through cracks | 20 |
| 16 | Highly productive aquifer through cracks | 29 |
| 17 | Highly productive aquifer through cracks | 43 |
| 18 | Highly productive aquifer through cracks | 41.5 |
| 19 | Highly productive aquifer through cracks | 11.5 |
| 20 | Highly productive aquifer through cracks | 33.5 |
| 21 | Highly productive aquifer through cracks | 80 |
| 22 | Highly productive aquifer through cracks | 130 |
| 23 | Highly productive aquifer through cracks | 22 |
| 24 | Highly productive aquifer through cracks | 49.25 |

*Table 7.3: Saturated conductivity properties from the Aquifer Properties Manual.*

begins to improve through the use of readily available data from the Aquifer Properties Manual, rather than whether very accurate simulations can be obtained with this specific distribution of chalk transmissivity.

71 catchments were affected by inclusion of hydraulic properties from the Aquifer Properties Manual. Relative to the standard run, the mean change in NSE across these catchments is an increase of 2.45, with 52% of catchments improving with respect to NSE scores. This overall increase in NSE conceals some interesting variation between catchments, however. Notably, 27 of the catchments that had properties changed as a result of including data from the APM already had an NSE of greater than 0.5 in the standard run, but several of these catchments actually worsened using the new properties (see Figure 7.8). Conversely, a number of the catchments that originally performed worst experience improvements with use of data from the Aquifer Properties Manual. This can be seen in the change in the quartiles and median of the NSE distribution shown in Table 7.4, with the 25th and 50th percentiles increasing while the 75th percentile decreases. This reflects the fact that catchments with initially good

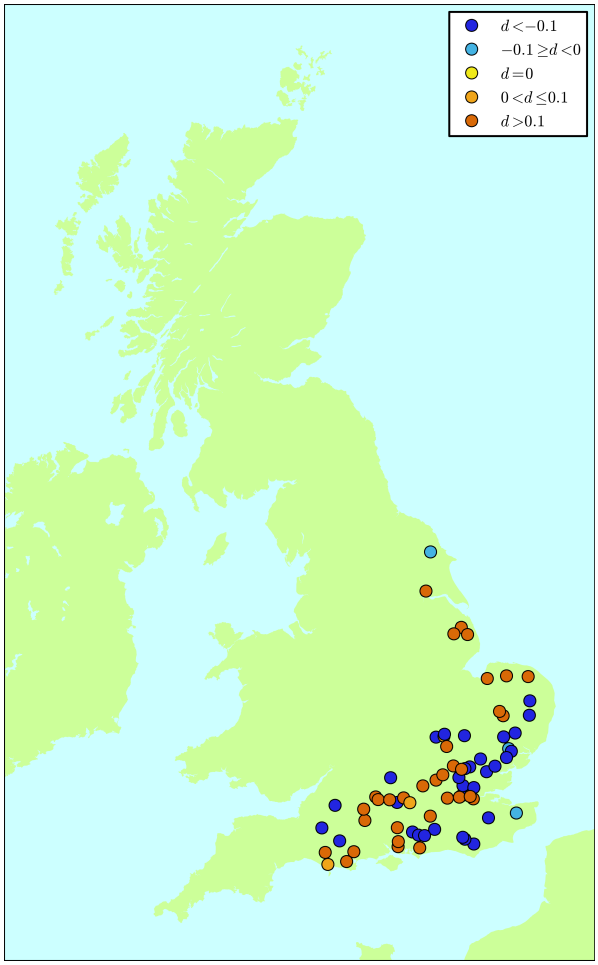



*Figure 7.8: Change in NSE from standard to APM parameters.*

performance were predominantly flashier with only a small area of highly productive aquifer or extensive cover by impervious superficial deposits such as clay, which are not represented in the model structure in this test. It is therefore consistent with expectations that increasing hydraulic conductivity to the levels apparent in major, unconfined chalk aquifers decreases the performance of the model in certain catchments.

However, rather than attempting to improve already well-simulated catchments, the main aim of changing the hydraulic conductivity parameters of the chalk in this way was to see if the characteristic response of chalk catchments - of high baseflow and relatively slowly-varying flow variation - could be reproduced by the model through a simple improvement in initially poorly simulated catchments. This was achieved for the worst performing catchments, as can be seen in the example of the Allen at Walford Mill. From comparing Figure 7.9 with Figure 7.10, it can be seen that substantial gains in the accuracy of simulation were made. Nevertheless, whilst NSE improves for catchments such as this one through dramatically reducing peak flows relative to the

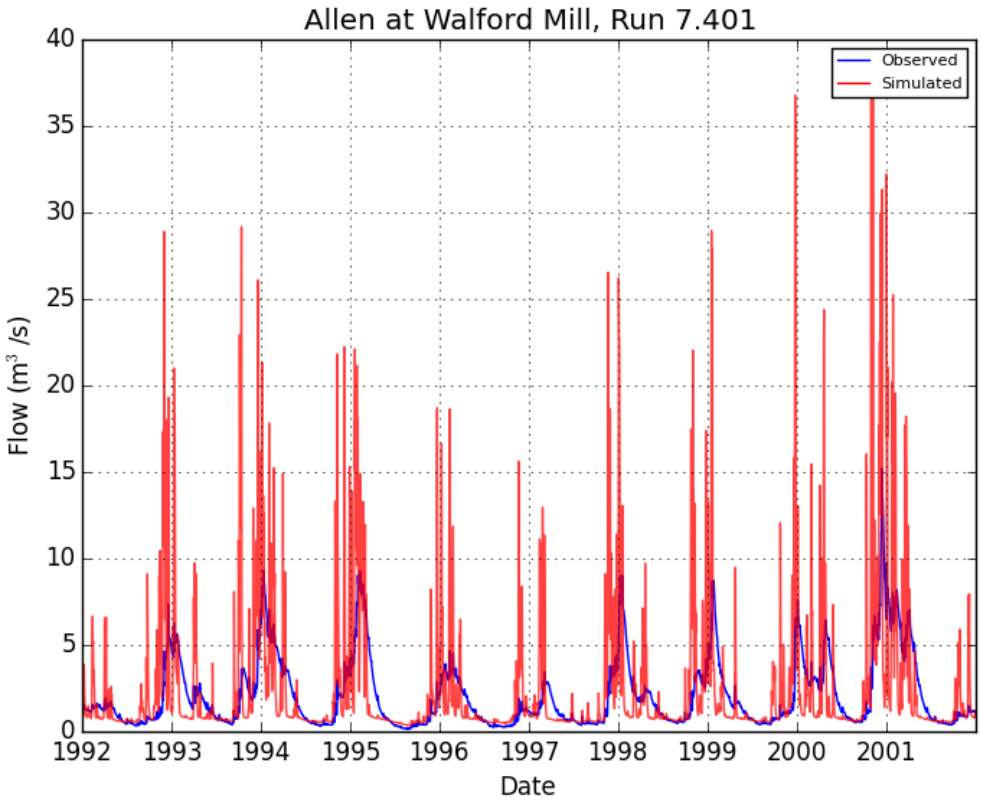



*Figure 7.9: Allen at Walford Mill hydrograph for the standard simulation.*

standard run, the water balance of the catchments does not necessarily improve, as the quantity of flow is still not correct in all cases. This explains the fact that the water balance bias and mmfd statistics in Table 7.4 do not show significant improvement. In some cases the performance statistics are of limited explanatory power in determining overall improvements in performance or are somewhat contradictory. For example, the Mimram at Panshanger Park shows the largest increase in water balance bias, but the shape of the hydrograph has improved greatly, which is reflected in an improved NSE score from -22.34 to -7.91.

The main conclusion from this test is that hydraulic properties derived from the Aquifer Properties Manual can be useful in improving the modelling system, particularly if used in conjunction with an accurate 3D geological model of the subsurface incorporating superficial deposits. The absence of such a detailed subsurface model in this test is likely to explain the variation in degree of improvement between catchments, which highlights how it is important to correctly simulate regional groundwater flow patterns in catchments with major aquifers, even at a relatively coarse scale.

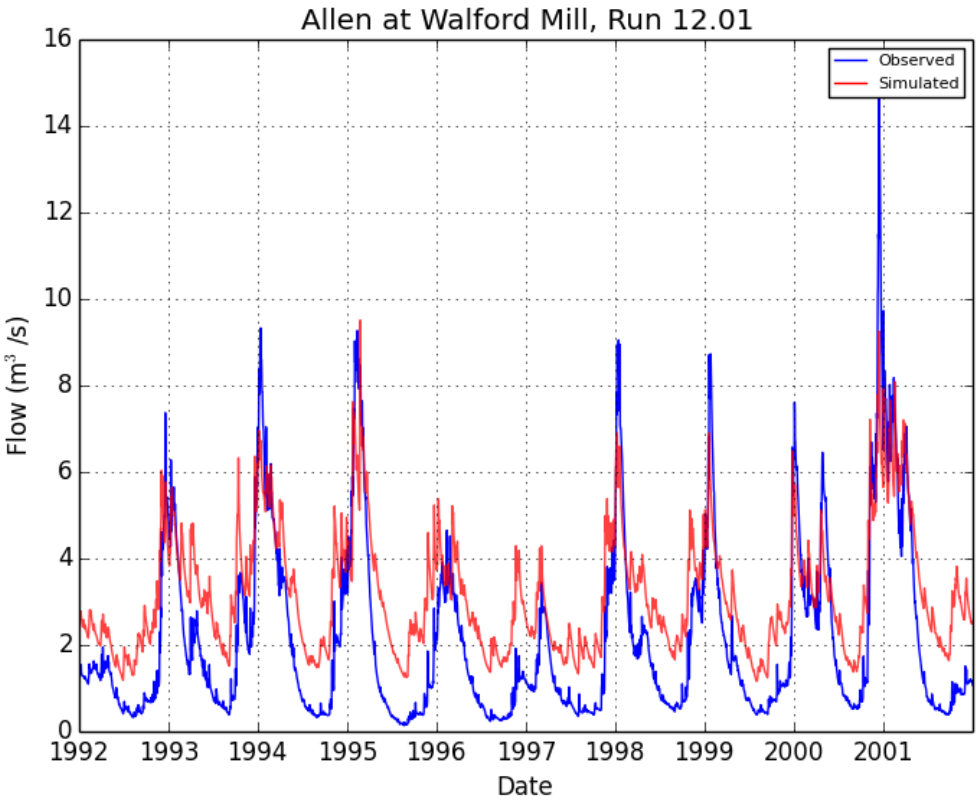



*Figure 7.10: Allen at Walford Mill for the APM simulation.*

### 7.6.3 3D geology

The results of the first tests presented above appear to indicate that more detailed and accurate representations of subsurface structure could be desirable. As such, the BGS 3D model of solid geology was incorporated into the national SHETRAN modelling system, as described in section 7.5 above. In the first instance this was conducted without adding in superficial deposits, which forms a separate test reported shortly. In many catchments across the UK where underlying bedrock has low hydraulic conductivity and storage coefficients, incorporating the 3D geological model is not expected to make a significant difference, as groundwater flow and storage is intrinsically limited in these geological formations. Of more interest are catchments overlying major aquifers, such as chalk and sandstone for example, where subsurface transmissivity can vary significantly as a result of heterogeneity in aquifer thickness as well as hydraulic conductivity. The parameters ascribed to each new lithology were based upon the hydrogeology code associated with each layer, which were consistent with the original hydrolgeology types used in the standard run. Therefore the parameter values applied to each lithology in the 3D model are drawn from the values associated with those codes in the standard

| Metric | Statistic | Standard | APM |
|---|---|---|---|
| NSE | Mean NSE Change | - | 2.45 |
| | % Improved | - | 52.1 |
| | Median | -0.05 | 0.23 |
| | Change in Median | - | 0.28 |
| | 25th Percentile | -3.15 | -0.72 |
| | 75th Percentile | 0.63 | 0.46 |
| | % Band 1 | 11.3 | 0.0 |
| | % Band 2 | 18.3 | 9.3 |
| | % Band 3 | 70.4 | 90.7 |
| Water Balance | % Improved | - | 21.1 |
| | Median | 16.1 | 22.3 |
| | Change in Median | - | 6.2 |
| | 25th Percentile | 0.7 | -12.4 |
| | 75th Percentile | 53.5 | 84.3 |
| | % Band 1 | 28.2 | 14.7 |
| | % Band 2 | 21.1 | 21.3 |
| | % Band 3 | 50.7 | 64.0 |
| mmfd | % Improved | - | 32.4 |
| | Median | 87.6 | 101.1 |
| | Change in Median | - | 13.5 |
| | 25th Percentile | 55.8 | 53.7 |
| | 75th Percentile | 150.6 | 191.4 |
| | % Band 1 | 5.6 | 0.0 |
| | % Band 2 | 8.5 | 14.7 |
| | % Band 3 | 85.9 | 85.3 |

*Table 7.4: Results from including APM values.*

run (see Table 7.5).

Incorporating the BGS 3D geological model results in changes to the subsurface representation of 158 catchments, rather than the 306 modelled in the standard run. This is because the 3D structural geology was only derived to the base of the Permian period, which excludes all of the catchments in Scotland. For the 158 catchments in which revised geology is applied, it can be seen from Table 7.6 that the mean change in NSE relative to the standard run is positive, showing that on average catchments are improving. The percent of catchments with higher NSE following the incorporation of 3D geology is 60%. However, the median NSE remains the same at 0.67. The 25th percentile of the NSE distribution increases notably from 0.08 to 0.3, while the 75th percentile increases only marginally from 0.77 to 0.79, which indicates that the majority of improvement is seen in catchments that performed poorly in the standard run. This point is reinforced by the change in percent of catchments in NSE bands. With the 3D geological model incorporated, the percentage of catchments in bands 2 and 3 decrease

| Name | saturated water content | residual water content | saturated conductivity | alpha | n |
|---|---|---|---|---|---|
| CRAG-PESA | 0.3 | 0.2 | 0.01 | 0.1 | 5 |
| SOLT-CLSISA | 0.3 | 0.2 | 0.001 | 0.1 | 5 |
| BRBA-SSCL | 0.3 | 0.2 | 0.01 | 0.1 | 5 |
| EOMIO-CLSSG | 0.3 | 0.2 | 0.01 | 0.1 | 5 |
| THAM-CLSSG | 0.3 | 0.2 | 0.0001 | 0.1 | 5 |
| LMBE-CLSSG | 0.3 | 0.2 | 0.001 | 0.1 | 5 |
| TAB-SSCL | 0.3 | 0.2 | 0.01 | 0.1 | 5 |
| WHCK-CHLK | 0.3 | 0.2 | 0.1 | 0.1 | 5 |
| GYCK-CHLK | 0.3 | 0.2 | 0.1 | 0.1 | 5 |
| GUGS-MDSL | 0.3 | 0.2 | 0.0001 | 0.1 | 5 |
| LGS-STMD | 0.3 | 0.2 | 0.1 | 0.1 | 5 |
| W-MDSS | 0.3 | 0.2 | 0.0001 | 0.1 | 5 |
| PB-LSMD | 0.3 | 0.2 | 0.01 | 0.1 | 5 |
| LOCR-SSML | 0.3 | 0.2 | 0.01 | 0.1 | 5 |
| WWAK-MDSS | 0.3 | 0.2 | 0.0001 | 0.1 | 5 |
| KLOX-MDSS | 0.3 | 0.2 | 0.0001 | 0.1 | 5 |
| IOGO-SLAR | 0.3 | 0.2 | 0.01 | 0.1 | 5 |
| RAG-SDSM | 0.3 | 0.2 | 0.01 | 0.1 | 5 |
| LI-MSLS | 0.3 | 0.2 | 0.0001 | 0.1 | 5 |
| MMG-MDSS | 0.3 | 0.2 | 0.001 | 0.1 | 5 |
| SSG-SDSM | 0.3 | 0.2 | 0.1 | 0.1 | 5 |
| ZG-DLDO | 0.3 | 0.2 | 0.1 | 0.1 | 5 |
| CCO-MDSS-2 | 0.3 | 0.2 | 0.0001 | 0.1 | 5 |
| APY-SCON | 0.3 | 0.2 | 0.01 | 0.1 | 5 |
| PUND-MDSS | 0.3 | 0.2 | 0.0001 | 0.1 | 5 |
| PUND-SCON | 0.3 | 0.2 | 0.1 | 0.1 | 5 |
| No Groundwater | 0.3 | 0.2 | 0.0001 | 0.1 | 5 |
| Low Productivity Aquifer | 0.3 | 0.2 | 0.001 | 0.1 | 5 |
| Moderately Productive Aquifer | 0.3 | 0.2 | 0.01 | 0.1 | 5 |
| Highly Productive Aquifer | 0.3 | 0.2 | 0.1 | 0.1 | 5 |

*Table 7.5: Properties assigned to the 3D geology model.*

slightly, by 1.5% and 2.2% respectively, while the percentage in band 1 increases by 3.7%.

Figure 7.11 shows the difference in NSE between the standard run and the run including 3D geology. From this figure it is apparent that the biggest increases in NSE are located in chalk regions. Interestingly, however, the largest decreases in NSE between the two runs are also located in the chalk catchments. Figure 7.12 confirms that, while there are increases in NSE in a number of locations, the chalk catchments still perform relatively poorly overall. As there is a good degree of confidence in the thickness and structure of the solid geology applied, one of the most likely reasons for poor NSE scores is inaccurate hydraulic property parameters leading to unrealisitic transmissivities. Furthermore, as this run does not include superficial deposits, their critical role in modulating runoff

| Metric | Statistic | Standard | 3D geology |
|---|---|---|---|
| NSE | Mean NSE Change | - | 0.44 |
| | % Improved | - | 58.9 |
| | Median | 0.67 | 0.67 |
| | Change in Median | - | 0.00 |
| | 25th Percentile | 0.08 | 0.30 |
| | 75th Percentile | 0.77 | 0.79 |
| | % Band 1 | 19.6 | 23.3 |
| | % Band 2 | 38.6 | 37.1 |
| | % Band 3 | 41.8 | 39.6 |
| Water Balance | % Improved | - | 41.1 |
| | Median | 7.2 | 7.3 |
| | Change in Median | - | 0.1 |
| | 25th Percentile | -1.2 | -3.7 |
| | 75th Percentile | 22.7 | 25.5 |
| | % Band 1 | 45.6 | 40.9 |
| | % Band 2 | 22.8 | 23.9 |
| | % Band 3 | 31.6 | 35.2 |
| mmfd | % Improved | - | 50.0 |
| | Median | 48.7 | 47.4 |
| | Change in Median | - | -1.3 |
| | 25th Percentile | 31.6 | 29.7 |
| | 75th Percentile | 82.0 | 94.1 |
| | % Band 1 | 9.5 | 8.8 |
| | % Band 2 | 29.7 | 32.7 |
| | % Band 3 | 60.8 | 58.5 |

*Table 7.6: Results from including 3D geology.*

processes is not simulated. Additional investigation of both of these issues is required.

When looking at the results of including the 3D geological model in terms of water balance bias, only 41% of catchments improved (Table 7.6). The median water balance bias remains the same, but the spread of biases widens slightly according to the interquartile range, from -1.2 to -3.7 at the lower end and from 22.7 to 25.5 at the upper end. The percentage of catchments in band 1 decreases by 5%, with resultant increases in the number of catchments in bands reflecting higher water balance biases. There is little change in mmfd overall.

While the results of the first trial of incorporating the 3D geological model are variable, the approach does appear to show significant potential. One example of this is the Mimram at Panshanger Park, which shows some of the largest improvements resulting from the addition of 3D geology (Figure 7.13). This catchment in the south-east of England is primarily underlain by unconfined chalk. Higher NSE is apparent in the run using the 3D geological model (-10.7) compared with the standard run (-22.3), as peak

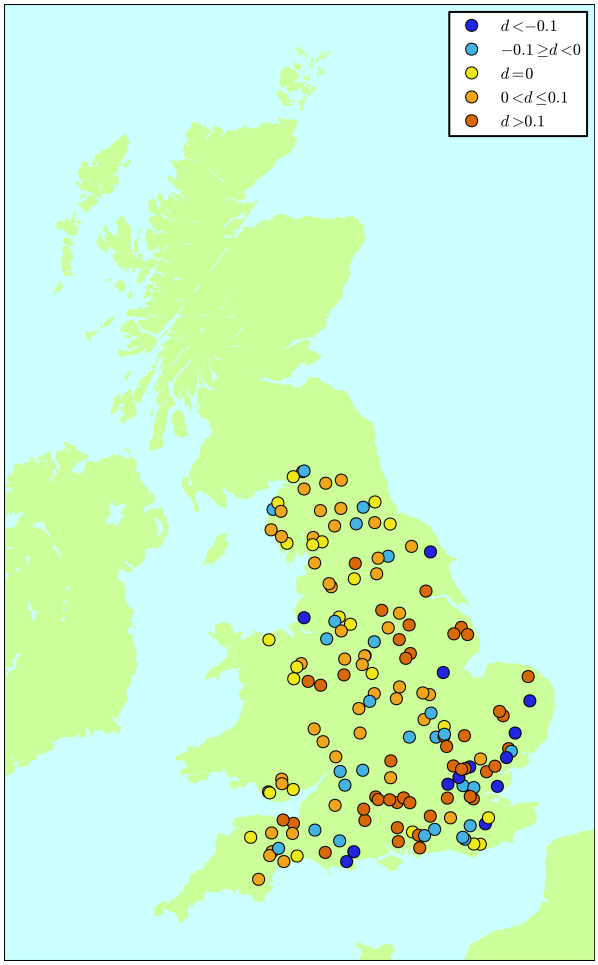



*Figure 7.11: Map of NSE change from the standard to GSI simulations (where d is change in NSE).*

flows decrease as a result of increased storage in the thicker chalk and a higher proportion of runoff deriving from the slower saturated subsurface flow pathway. However, Figure 7.13 indicates that the shape of the simulated hydrograph is still not completely consistent with observations, as it remains too peaky. This indicates that, whilst increasing storage by increasing the thickness of the aquifer helps, this approach by itself is not sufficient for accurate simulation of chalk catchments, as appropriate estimates of hydraulic parameters are still required. For chalk catchments, spatial patterns of variation in hydraulic properties need to be accounted for, which includes both areal and vertical heterogeneity, particularly in terms of the differences between interfluve and valley areas, as well as variations (typically decreases) in hydraulic conductivity with depth (MacDonald and Allen, 2001).

This run forms an initial attempt at incorporating the most realisitic information available with regards the subsurface in the form of the BGS 3D geological model. This

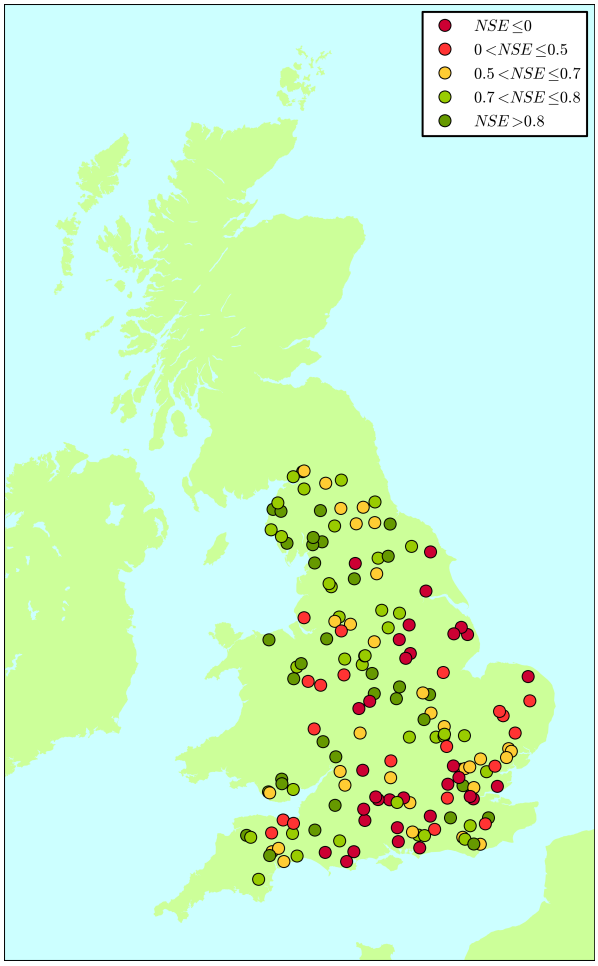



*Figure 7.12: NSE of the GSI simulations.*

integration has not been carried out before, and it is clear that significant further work is required in order to achieve good representations across all catchment types. Fortunately a methodology and work flow is now in place for further experimentation, which will allow refinements to the geological model and hydraulic properties to be incorporated. Two further tests have been conducted on the basis of this methodology, which are discussed in turn below, with further work anticipated for the future.

### 7.6.4 Inclusion of MODFLOW T

Of the catchments that have been modelled in this study, only 24 overlap with the data from EA regional chalk groundwater models available for this study, such that the impact of including parameters applied in the MODFLOW models could only be examined for a small number of catchments. The method for deriving hydraulic conductivity inputs for SHETRAN from these transmissivity distributions is explained above

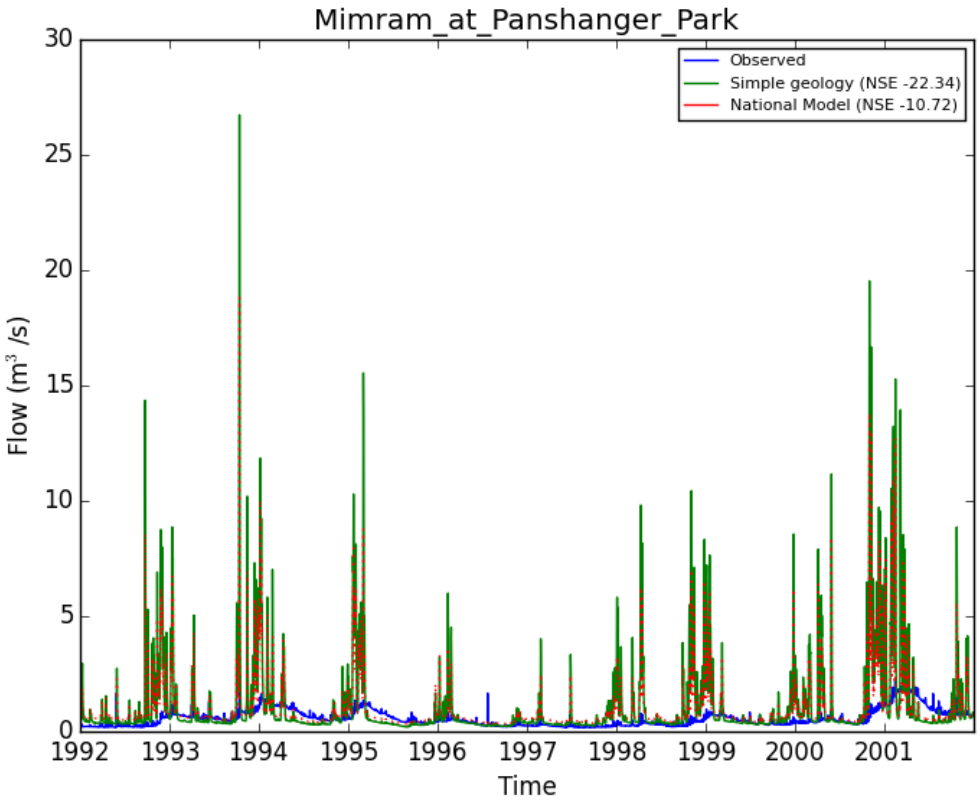



*Figure 7.13: Hydrograph of the Mimran at Panshanger Park showing the observations (blue), standard simulation (green) and simulation with 3D geology (red). The NSE improves from -22.3 to -10.7.*

in section 7.5. It is important to note that the data from EA models has been used in conjunction with the 3D geological model representation of the subsurface but without the incorporation of superficial deposits to test only the impact of including measured transmissivity values. As no superficial deposits are included the largest improvements are expected to be seen in catchments where there is no significant influence from superficial deposits, however, all of the catchments affected in this test are influenced by superficial deposits and so there is no way of testing this hypothesis.

With a much smaller sample size (24) for analysing the results of this test relative to the standard run, the changes to the model performance statistics in relative terms are larger than those calculated for previous tests. On average, NSE increases by 1.37 across the catchments analysed, with 58% of catchments showing an improvment in NSE when incorporating chalk transmissivity values and distributions from the EA groundwater models Table 7.7. The median NSE increased from 0.54 in the standard run to 0.60, and both the 25th percentile and 75th percentile NSE values increased

| Metric | Statistic | Standard | 3D + EA T |
|---|---|---|---|
| NSE | Mean NSE Change | - | 1.37 |
| | % Improved | - | 58.3 |
| | Median | 0.54 | 0.60 |
| | Change in Median | - | 0.05 |
| | 25th Percentile | -0.59 | -0.07 |
| | 75th Percentile | 0.72 | 0.78 |
| | % Band 1 | 16.7 | 12.5 |
| | % Band 2 | 16.7 | 37.5 |
| | % Band 3 | 66.7 | 50.0 |
| Water Balance | % Improved | - | 41.7 |
| | Median | 20.1 | 25.5 |
| | Change in Median | - | 5.4 |
| | 25th Percentile | 2.9 | 4.6 |
| | 75th Percentile | 37.3 | 77.4 |
| | % Band 1 | 29.2 | 12.5 |
| | % Band 2 | 16.7 | 20.8 |
| | % Band 3 | 54.2 | 66.7 |
| mmfd | % Improved | - | 45.8 |
| | Median | 69.7 | 95.1 |
| | Change in Median | - | 25.4 |
| | 25th Percentile | 59.0 | 50.8 |
| | 75th Percentile | 112.9 | 178.7 |
| | % Band 1 | 8.3 | 0.0 |
| | % Band 2 | 4.2 | 16.7 |
| | % Band 3 | 87.5 | 83.3 |

*Table 7.7: Results from including T values from the EA regional groundwater models.*

from -0.59 to -0.07 and 0.72 to 0.78 respectively, reflecting a general improvement in NSE. However, it should be noted that the percentage of catchments with band 1 NSE classifications decreases, as does the percentage of catchments with band 3 NSE scores, which indicates that the some of the initially best-performing catchments worsen and some of the worst improve.

Figure 7.14 shows the change in NSE relative to the standard run that comes from using a more detailed description of 3D geology and EA groundwater model transmissivity values. There is no simple geographical pattern to the results immediately apparent from this figure, with neighbouring catchments subject to NSE changes in opposing directions in some cases, as can be seen in the East Yorkshire chalk area for example. There are no clear explanations as to why this is the case. There are complex patterns of superficial deposits over most catchments (as described in the National River Flow Archive), there are certainly issues relating to averaging transmissivity uniformly over depth and also with running models of groundwater catchments to only their surface

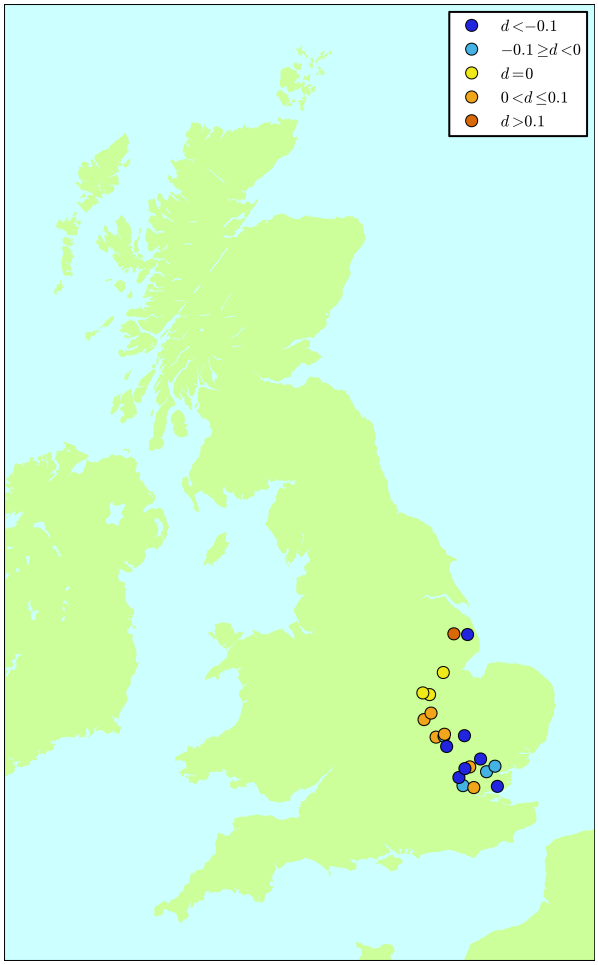



*Figure 7.14: NSE change due to T values from the EA regional groundwater models.*

water extents.

Figures 7.15 and 7.16 show an example of the improvement that can be obtained from incorporating into SHETRAN the transmissivity values from the EA groundwater models with the BGS geological model. This indicates that, while the shape of the hydrograph is still not entirely consistent with observations, it is much improved and more accurately reflects a baseflow-dominated hydrological regime.

Table 7.7 also shows that the water balance bias does not improve with the inclusion of transmissivity from the EA groundwater models. 42% of catchments show improvements in water balance bias, but there is a general shift in the median and quartiles to higher biases (albeit with a limited sample size). However, in this test there is a large change in the mmfd. The median mmfd increases from 70 to 95, which represents a deterioration in the correspondence of simulated and observed flow regimes. The 25th percentile of the mmdf distributions decreases whist the 75th increases, indicating that spread in mmfd increases but that there is not a general shift to poorer mmfd values.

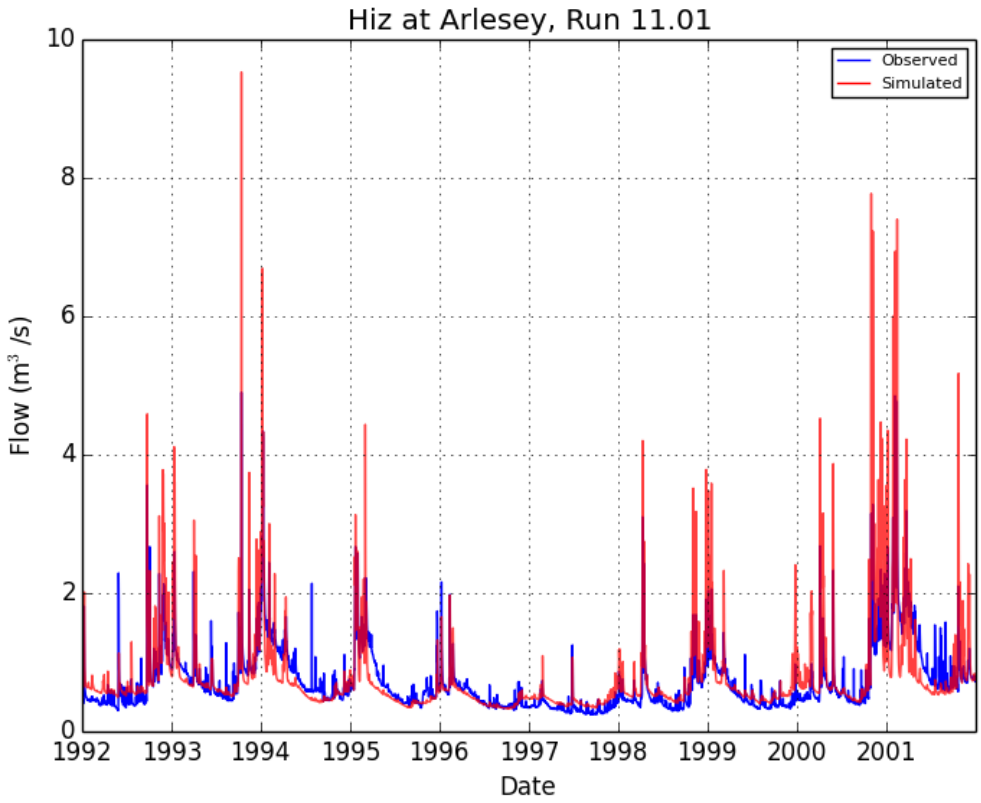



*Figure 7.15: Hydrograph for the Hiz at Arlesey with 3D geology. NSE = -1.31.*

Using the EA model T values improves many aspects of the simulated hydrographs but often in ways that the performance statistics do not capture. Including this data is clearly not a panacea for correcting groundwater dominated catchment performance but is certainly a useful data source for conductivity values. These values must be used in conjunction with other information (superficial deposits, flow pathways, hydrostratigraphy etc.) to provide a real improvement.

### 7.6.5 3D geology and superficial deposits

As described above, the Parent Material map was used as a proxy for superficial deposit type. One of the main limitations of this dataset is that it does not provide any stratigraphic information about superficial deposits, i.e. in cases where there could be some layering. Superficial deposits based on the Parent Material map were combined with the representations of soil and solid geology based on the BGS 3D model, as outlined above. As superficial deposits cover much of the country, comparisons between the standard run and this run (incorporating both 3D solid geology and superficial

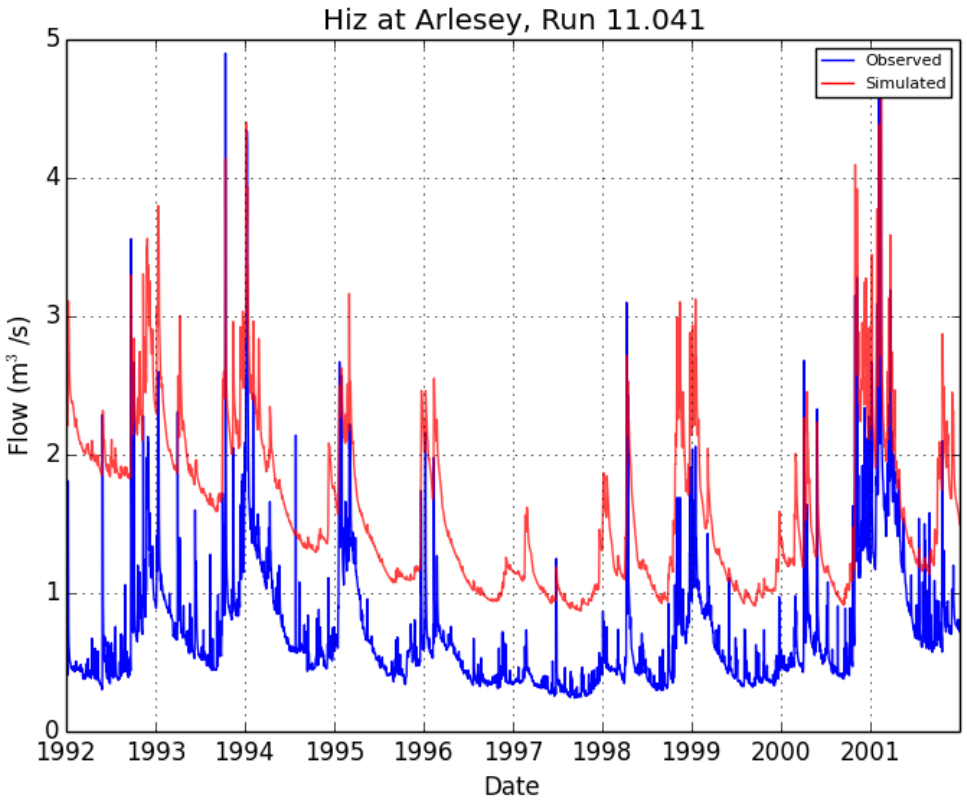



*Figure 7.16: Hydrograph for the Hiz at Arlesey with 3D geology and transmissivity values from the EA groundwater models. NSE = 0.52.*

deposits) can be made for 258 catchments. The hydraulic properties of the superficial deposits can be found in Table 7.1.

Combining the 3D geology representation with superficial deposits results in a mean increase in NSE of 0.75 across all catchments. 51% of catchments show improved NSE with the inclusion of both superficial deposits and 3D geology. Similar to the results from implementing the 3D geological model alone, the shift in performance seems to be at the lower end of the NSE distribution, with the 25th percentile increasing from 0.46 to 0.52 whilst the 75th percentile only changes from 0.79 to 0.80 from the standard to superficial deposit runs respectively. The percentage of catchments in bands 2 and 3 of NSE decreases from 42.2% to 41.4% for band 2 and from 35.7% to 31.9% whilst band 1 increases from 22.1% to 26.6%.

Figure 7.18 shows that the largest increases in NSE between the standard run and the run including both 3D geology and superficial deposits are found in regions of chalk catchments, as well as more generally across England and Wales. Scotland shows the least number of improved catchments, but still some. This is probably due to a

| Metric | Statistic | Standard | 3D and superficial |
|---|---|---|---|
| NSE | Mean NSE Change | - | 0.75 |
| | % Improved | - | 51.6 |
| | Median | 0.69 | 0.70 |
| | Change in Median | - | 0.01 |
| | 25th Percentile | 0.46 | 0.52 |
| | 75th Percentile | 0.79 | 0.80 |
| | % Band 1 | 22.1 | 26.6 |
| | % Band 2 | 42.2 | 41.4 |
| | % Band 3 | 35.7 | 31.9 |
| Water Balance | % Improved | - | 51.6 |
| | Median | 4.8 | 3.5 |
| | Change in Median | - | -1.3 |
| | 25th Percentile | -2.7 | -3.7 |
| | 75th Percentile | 17.7 | 15.0 |
| | % Band 1 | 53.1 | 50.2 |
| | % Band 2 | 22.9 | 26.2 |
| | % Band 3 | 24.0 | 23.6 |
| mmfd | % Improved | - | 51.6 |
| | Median | 41.0 | 39.1 |
| | Change in Median | - | -1.9 |
| | 25th Percentile | 27.7 | 27.7 |
| | 75th Percentile | 70.3 | 70.9 |
| | % Band 1 | 11.6 | 12.2 |
| | % Band 2 | 36.8 | 38.8 |
| | % Band 3 | 51.6 | 49.0 |

*Table 7.8: Results from including 3D geology and superficial deposits.*

variety of factors: the 3D geology, the shallower catchments so that infiltration is more prominent; the more complex geology and superficial deposit types.

Notably, Table 7.8 suggests that the changes in water balance arising from incorporating improvements to both solid geology and superficial deposits may provide more benefits than simply using the 3D geology by itself. The water balance bias improves in 52% of catchments relative to the standard run, with the median bias decreasing slightly (in the 3D geology only run this increases slightly from 7.2 to 7.3). However, when including 3D geology and superficial deposits, the percentage of catchments in band 1 and 3 both decrease slightly relative to the standard run while the number of catchments in band 2 increases, which suggests that the downward shifts in water balance bias are not completely convincing, with some of the poorest performing catchments improving slightly but some of the better ones worsening. The mmfd statistics are very similar for both runs, which suggests that changes between the tests are general and not seasonally or percentile specific.

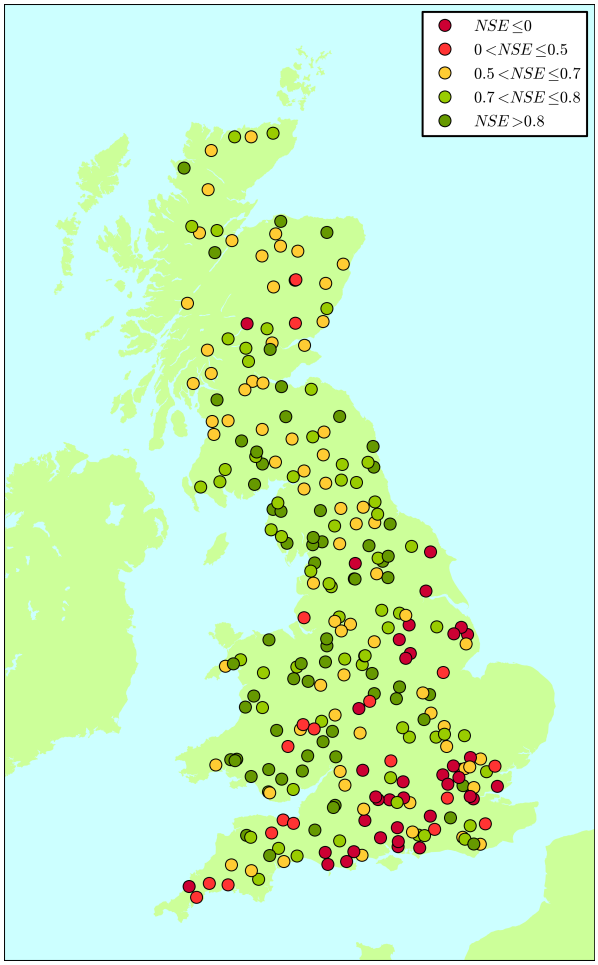



*Figure 7.17: NSE of superficial deposits test.*

The catchments that show the largest improvements in this simulation are due to the inclusion of the 3D geology rather than the inclusion of superficial deposits. This suggests that there are some limitations in the degree to which the Parent Material map used here adequately captures the distribution of superficial deposits across the country, although this issue is not necessarily simple to separate from that of difficulties in estimating the hydraulic properties of these deposits. One example of the limitations of the data used here for deriving superficial deposits is the case of permeable deposits such as gravel overlying impermeable bedrock, which was discussed in Chapter 5. Gravel deposits acting as local or minor aquifers may play a highly important role in catchment response, for example by increasing baseflow compared with the case where the deposits were absent. The inclusion of superficial deposits in the model structure should resolve problems like this, but the national data available at this stage in the form of the Parent Material map does not seem to represent situations such as this in sufficient detail. For example, the National River Flow Archive describes the catchment geology of the

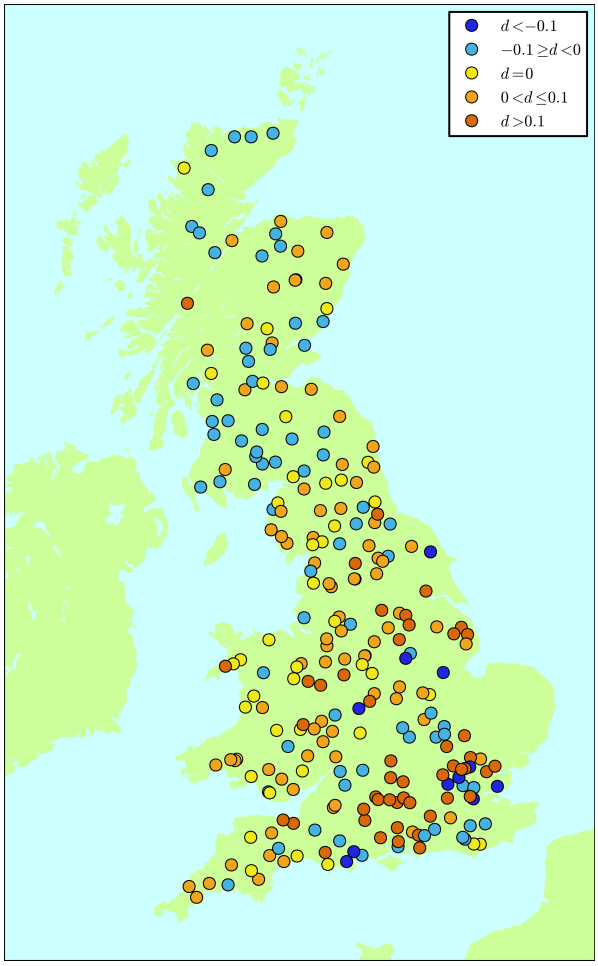



*Figure 7.18: Change in NSE to inclusions of superficial deposits.*

Lugg at Byton as being 'impermeable formations covered by extensive alluvial gravel deposits in the valleys'. Yet the Parent Material map shows this area as various types of mudstone (simplified to 'rock' for SHETRAN. See Table 7.1), giving no indication of the presence of significant gravel deposits. The Parent Material map is therefore unlikely to be the ideal choice for this modelling system, and further work is required to identify or derive a more appropriate solution.

## 7.7 Conclusions

This chapter has explored the potential for improving a national modelling system based on SHETRAN through incorporating more detailed information on subsurface structure and properties than previously available. The tests outlined here indicate that notable improvements in performance of the modelling system in terms of hydrograph shape may be obtained by incorporating more realistic stratigraphy and hydraulic prop-

erties. This is apparent from the test run using hydraulic properties for the chalk from the Aquifer Properties Manual (MacDonald and Allen, 2001) (1), in which some of the catchments that performed worst in the standard run improved substantially. Similarly, applying hydraulic conductivity based on transmissivity maps from EA regional groundwater models leads to some considerable improvements in model performance, such as the Hiz at Arlesey. It is also the case that some gains in simulation accuracy were made from utilising more accurate layering and thicknesses of lithological units from the BGS 3D geological model. In this run 60% of catchments have higher NSE following incorporation of 3D geology.

However, the importance of accurate data with regards to both geological structure and hydraulic properties is evident from the way that performance changes across catchments are mixed in each of these runs. For example, The inclusion of EA transmissivity data results in an improvement in NSE in 58% of the catchments tested and worsen in 42%. This means that for some catchments performance appears to worsen with inclusion of theoretically more accurate information, which suggests that some compensating errors may exist in the standard run or that some of the additional information is not completely accurate. As the BGS geological model is considered to be a high quality data source - although some further verification is still required - the implication here is that further work is required to ensure that all of the available data on hydraulic properties in particular are incorporated into the model, as this is likely to be critical in improving some of the poorly performing catchments. For chalk regions, refining hydraulic properties would be likely to mean a better treatment of the vertical variation in hydraulic conductivity, similar to the MODFLOW-VKD implementation (Environment Agency, 2003). It would potentially also be beneficial to identify which catchments would be better simulated on the basis of groundwater catchment boundaries rather than surface watersheds.

In addition, the analysis above suggests that the Parent Materials map is unlikely to form an adequate basis for specifying superficial deposits in the modelling system. This is demonstrated by the fact that National River Flow Archive descriptors to not match up with the Parent Materials map. Instead, a 3D model of superficial deposits could potentially be of significant value when combined with the BGS solid geology model, but again the effectiveness of this approach would be limited by the quality of hydraulic property information that could be applied. However, given the extensive areas of superficial deposits overlying many regions of the UK, rectifying their representation in the SHETRAN modelling system could bring substantial improvements in simulation of overall catchment flow regimes.

The work carried out to couple the BGS 3D geological model with a national SHETRAN modelling system opens up the possibility for further improving parameter values and the representation of superficial deposits. Simultaneously simulating large numbers of catchments and assessing the effects of different refinements is now relatively straightforward as a result of the system developed in this research, although additional work on metrics to assess model performance is desirable to overcome some of the limitations of different skill scores so clear from the analysis in this chapter. Further refinements for regions with significant aquifers should also be carried out collaboratively by geologists, hydrogeologists and hydrological modellers to ensure that the highest quality data and experience are brought to bear on the modelling system.

# Chapter 8

# Conclusion

## 8.1 Summary of results

The preceding chapters have presented the development and analysis of a robust, multi-purpose hydrological model for Great Britain. This physically-based modelling system was created using SHETRAN and national datasets, allowing any catchment - gauged or ungauged - to be set up in a matter of seconds, a process that could previously take weeks or even months. This efficient method of model setup is facilitated by a newly developed and easy-to-use graphical interface. For national scale assessments and analysis, the modelling system utilises automatic setup scripts and has additionally been configured to make use of available parallel processing options. This workflow facilitates for rapid simulation of 306 catchments for multi-year periods using hourly time steps and relatively high spatial resolution, which could easily be increased if required.

In addition, a new national, gridded hourly rainfall dataset has been produced for input to the modelling system, in order to explore the effects on model outputs of using more accurate sub-daily meteorological forcings. The dataset was created for the period 1990 to 2006 by disaggregating the UKCP09 5km gridded daily rainfall dataset using data from over 1300 hourly rain gauges in conjunction with nearest neighbour interpolation. Comparing the hourly records with daily rainfall data showed that both datasets are consistent in timing and amount of rainfall for the majority of gauges, although several common quality control issues were identified for some rain gauge records. Depending on the nature of the issues, these records were removed prior to gridding.

Extensive analysis of the performance of the national modelling system and its sensitivity to several of the most important input datasets and parameters has been conducted.

The initial run undertaken gave acceptable model performance (NSE>0.5) in the majority of catchments, with some exceptions located generally in parts of Scotland and the south-east of England. Structural improvements were then made to the SHETRAN system to include snow melt processes particularly relevant in upland regions, as well as better representation of lakes and handling of sinks in the input DEM. These modifications resulted in clear improvements leading to definition of the standard configuration of the national SHETRAN modelling system for Great Britain, against which subsequent sensitivity tests and scenario runs are analysed. The standard configuration produces acceptable results across much of the country where it performs satisfactorily (NSE>0.5) for 72% of catchments and well (NSE>0.7) for 48%. The major exception to this level of performance is seen in some catchments in the south of England, particularly those underlain by significant aquifers and so where hydrological regimes are highly influenced by subsurface processes and their interaction with surface hydrology. Some of the limitations of currently available national datasets with respect to capturing important features of local hydrological cycles are also apparent from the sensitivity analysis, important examples of which include insufficient data on superficial deposits and artificial influences, such as abstractions or flow regulation.

Comparisons with other UK multi-catchment studies and other national models confirm that SHETRAN for GB simulates a large proportion of catchments well and indeed better than calibrated conceptual models for many of the best modelled catchments. However, limitations in subsurface representation lead to comparatively poor simulations in catchments with significant groundwater systems. Importantly, the results presented in this study are from an uncalibrated, physically-based model, whereas the comparison studies use calibrated models. As such the high level of performance of SHETRAN for GB is very encouraging.

Sensitivity to the Strickler coefficient representing surface roughness was examined, as were sensitivities to rainfall and potential evapotranspiration (PET) quantities and distributions. In tests where relatively small perturbations were applied to the standard datasets, such as increasing PET by 5% or using an exponential distribution for sub-daily rainfall, mixed changes in model performance were typically obtained, with a fairly even split between catchments showing overall improvements and declines in performance. These tests appear to indicate a degree of robustness in terms of system setup and inputs, with variation in catchment performance responses to changes suggesting little systematic bias and good overall simulation if not necessarily always locally optimal. This is reflected when larger changes are applied, such as a 20% increase in rainfall or a Strickler coefficient of 5, which leads to large overall decreases in

model performance. In conjunction with the higher variability of responses arising from smaller perturbations, this is taken as an indication that generally appropriate parameters and input datasets are being used in the system. Furthermore, where the changes made had a physical basis - particularly using realistic hourly meteorological as inputs - model performance tended to increase slightly at the national scale. The largest percentage of improved catchment simulations resulted from incorporating hourly rainfall based on the new dataset created during this project, as well as PET disaggregated to hourly intervals. Improvements were also seen when the AE/PE ratio was increased by 0.1, PET increased by 5% or rainfall decreased by 10%, which reflects the possibility that calculated PET values may be too low or rainfall slightly too high, although the latter is not consistent with the expectation of undercatch for a number of raingauges.

Throughout the sensitivity analysis it was generally found that groundwater-dominated catchments in the south-east of England were not adequately simulated. The reasons for this were deemed to be largely related to inadequate descriptions of superficial and solid geology in the standard model datasets, as well as insufficient characterisations of hydraulic properties and their patterns of spatial variation. For example, the initial hydrogeological input to the modelling system was a 2D hydrogeological map that only distinguishes between high, medium and low productivity aquifers, which does not provide enough data to determine aquifer transmissivity in a meaningful way. Possible improvements to the representation of geology in the modelling system were therefore investigated. In particular, the BGS national 3D geological model, which accurately characterises stratigraphy and strata thickness to the base of the Permian, was incorporated into the modelling system to provide a more realistic description of subsurface structural (solid) geology. The BGS parent materials map showing the type of deposit from which soils are likely to be derived was additionally incorporated as a proxy for superficial deposits, although this approach was found to be unsuitable for appropriate representation of the hydrological role of these deposits. High quality information on superficial deposits is therefore vital for further improving model realism and performance. In addition, more accurate hydraulic conductivity distributions were incorporated from both the Aquifer Properties Manual (MacDonald and Allen, 2001) and from EA regional groundwater models, which confirmed that more realistic parameter values can improve hydrograph shapes if not other deficiencies in simulation of some catchments. A substantial body of further work is still required to improve the representation of permeable catchments in the modelling system, with particular challenges remaining in terms of data for the distribution and properties of superficial deposits, as well as identifying appropriate hydraulic property parameters and their

patterns of variation in some key aquifers such as the chalk.

A climate change impact study was also conducted for 20 catchments across the country to demonstrate how the modelling system could be used for a full national climate change impact assessment study. In order to facilitate this, code was developed to apply outputs from the UKCP09 spatial weather generator as inputs to the SHETRAN modelling system. A medium emissions scenario for the 2050s was chosen as an exemplar case. The results from this preliminary investigation suggest that higher projected PET rates, in combination with modest increases in rainfall, could lead to a general decrease in mean annual river flows of between -24% to +7% across the country under this scenario. The largest reductions in mean annual flow are found in the south of England, while modest increases in flow are found in Scotland, which reflects spatial variation in the balance between - and expected changes in - rainfall and PET across Great Britain. Seasonally, SHETRAN shows changes in flow of +1% to -16% in spring, -12% to -42% in summer, +5% to -43% in autumn and +17% to -22% in winter. Similar to the case of changes in mean annual flows, there are notable geographical patterns to these results, with higher flows in the 2050s relative to the baseline typically found in the west and north of Great Britain, while lower future flows are more common in the south and east. The results also show that flows typically decrease across all flow percentiles except for the uppermost (highest) flow percentiles. However, the 10-year return period flood peak is shown to increase dramatically from -5% to +26% across the country. Comparison of these results with current projections from the Future Flows Project (Prudhomme et al., 2012) shows overall consistency between projections using two different types of hydrological model.

## 8.2 Results in the context of important hydrological questions

The national, physically-based modelling framework developed in the research has been applied largely for gauged catchments to date. However, there are a number of ways in which the work relates to the ongoing issues associated with prediction in ungauged basins (Sivapalan, 2003; Hrachowitz et al., 2013). Rather than using some approaches typically applied to simulating ungauged basins, such as regionalisation on the basis of catchment characteristics and calibrated parameters for available gauged catchments (e.g. Hundecha and Bárdossy (2004)), the focus here has been on testing the sensitivity of universal parameterisations and datasets. Taking this approach starts to explore

questions of whether calibration and statistical relationships between catchment properties and parameters are required for simulating a diverse array of catchments across Great Britain if a physically-based model is used. The good overall performance obtained for the majority of catchments simulated indicates a fairly high degree of model robustness taking a universal approach. This may be interpreted as indicating that the parameters and datasets are appropriate and reasonably close to a global optimum, which is encouraging for applications in ungauged basins. However, there are of course a number of catchments that are still relatively poorly simulated, as well as issues of slightly lower performance in the validation period and the fact that this approach does optimise performance for any given catchment.

It has been possible to identify and constrain quite closely those situations under which the modelling system is likely to provide reasonable or poor flow predictions, which helps to move the system a bit closer to operational usefulness. This knowledge may inform decisions as to whether the modelling system is appropriate for particular applications in ungauged basins. For example, the analysis suggests that, at the present stage of model development, responsive catchments dominated by surface flow regimes are more likely to be well simulated than catchments with significant hydrogeological dimensions, particularly in chalk regions. Cluster analysis helps to confirm and quantify this characterisation of variation in system performance, providing some degree of confidence in predicting the performance of an ungauged catchment based on a relatively small number of catchment descriptors. While the system does not yet provide a suitable means of flow prediction in all types of catchments in Great Britain, this work helps to identify where it could form a realistic possibility. It also highlights where physical improvements to the modelling system and its input data should be made if applications in ungauged basins are to become more tractable.

In addition, it is also interesting to consider how the results of this study fit with the debates regarding the strengths and limitations of physically-based spatially-distributed hydrological models. As discussed in Chapter 2, several arguments about the theoretical underpinnings of physically-based models have been put forward in the literature, which tend to be around issues of scale, nonlinearity, data requirements, catchment uniqueness and parameter estimation particularly (Beven, 1989, 2001). Full exploration of these arguments has not been possible within the scope of this project, but importantly, a physically-based model appears to perform well across a large number of catchments without any local tuning of parameters. This is an encouraging result for the structure and equations underpinning SHETRAN, but of course poorly simulated catchments need to be addressed. The investigations undertaken so far point towards

data availability still forming a significant challenge in such catchments, particularly with respect to characterisation of the subsurface. Initial results from including more realistic geology and hydraulic properties suggest that better physical information could resolve some of these problems, but further work along this line is required before this can be fully appraised. Estimation of subsurface parameters of both soil and geological formations is likely to remain a significant challenge, but further compilation and synthesis of hydraulic testing data is an important avenue of investigation. At any rate, some of the practical barriers to employing physically-based models for various applications have certainly been diminished through this work, which has additionally enabled new possibilities, such as using SHETRAN for comparative studies on large samples.

The potential applications arising from the development of this system are indeed significant. Extension of the climate change analysis presented here is very plausible and forms an interesting new approach to national scale assessment. Coupling with the UKCP09 weather generator provides some advantages over the use of change factors applied in a number of other studies. For example, it produces a set of simulations so that a range of futures can be assessed. Even taking advantage of the resources now available for parallel processing, climate change impact assessment using this method still forms a relatively computationally expensive undertaking, but the methodology is robust as it is physically based. Theory would suggest that results from a model such as SHETRAN should be the most reliable available for climate change projections, but more work is required to compare the differences in predictions with other types of models. There is a similar requirement for land use change studies, but in both cases it is hoped that such analysis can be conducted relatively efficiently with the tools now developed.

This work has therefore made SHETRAN a more approachable tool for hydrological applications and has already had a positive impact beyond this project alone. The system has been used to set up catchment models for several unrelated studies already, such as a joint project by the BGS and EA on groundwater flooding in the East Yorkshire chalk. This forms part of an ongoing project to improve representation of groundwater in the modelling system and apply it for managing groundwater flood risk. The model has also been used by Nottingham University to look at impacts from future land cover changes, while the Met Office is currently exploring how to couple SHETRAN with the land surface scheme JULES, in order to improve the representation of hydrology in the latter. Three PhD projects are also currently using the modelling system, while the rapid setup facility has been used to help a number of Masters students with models

for their dissertations. Furthermore, this system has great potential to become a national resource for various applications in flood and water management, for example by examining scenarios arising from different pressures on water quantity and quality at local, regional or national scales. The modelling system described here is on a par with other national modelling capabilities (Henriksen et al., 2003; Habets et al., 2008) and has a different set of uses than both the G2G (Bell et al., 2009) and CLASSIC-GB (Crooks et al., 2014) models applied in the UK which focus solely on modelling river flows.

## 8.3 Future work

There are a number of directions for future research and applications that could utilise the modelling system developed and other work presented in this thesis, some of which are already underway. These future lines of work are grouped by improvements to software, data, geology and general improvements.

Software:

- The modelling system has so far only been tested at a 1km spatial resolution. This is a typical resolution for a national scale study (Henriksen et al., 2003; Bell et al., 2009), but if it is to become a truly multi-purpose system then the model should be run and evaluated at different resolutions. So far the model setup system provides capabilities to produce catchment models across Great Britain and 100m and 500m resolutions, but these models have not yet been tested. Doing this would certainly make an interesting sensitivity test, extending prior work by Zhang (2012) further. Zhang's study highlighted the importance of temporal resolution for SHETRAN simulations, but the effects of different spatial resolution on SHETRAN outputs remain unexplored for the range of catchment types that could be investigated using the modelling system.
- The graphical user-interface for setting up catchments currently requires the user to upload a shapefile of the catchment boundary in order to create a model. A script has been written and tested for generating catchment boundaries automatically from a DEM given an outlet point, which will be incorporated into the user-interface to further improve its ease of use.
- The modelling system is currently only available for desktop computers, although other servers and networks can of course be used for running simulations. It is

planned that the system will be put on an internal intranet and then later onto the internet with appropriate licences and support.

Data:

- CEH have recently released a new 1km resolution daily gridded rainfall product, CEH-GEAR (Tanguy et al., 2014). This is a high quality, freely available rainfall dataset, and so the potential for using it as meteorological forcing for the modelling system should be investigated. Initial comparisons with the 5km UKCP09 dataset applied to date have shown that it is very similar to CEH-GEAR, but the licence conditions of the latter allow for it to be more readily distributed and applied.
- The hourly rainfall dataset is being updated to be based on the CEH-GEAR gridded dataset in conjunction with records from hourly raingauges, which will allow for the dataset to be freely distributed (hosted by CEH). Further analysis of the hourly gridded rainfall dataset is currently being undertaken with respect to climatology, preservation of extremes and comparisons of interpolation methods.
- The rainfall dataset employed in the model is considered to be the best available at the time of model setup, but the lack of readily available potential evapotranspiration (PET) data for the UK meant that PET had to be calculated with some approximations due to data limitations. This is considered to be the best option at this time, as indicated by the overall accuracy of simulated water balances. However, it would certainly be of value to see if the reasonably small water balance bias could be rectified by use of an alternative dataset, such as the MORECS PET products created by the Met Office (Hough and Jones, 1997). Similarly, the possibility of coupling SHETRAN with JULES as described above would also be a useful and valuable comparison to further test the adequacy of the current PET dataset.

Geology:

- It is intended that further collaboration with BGS/EA will be carried out to integrate superficial deposits and appropriate hydrogeological parameters with the BGS 3D geological model. Once this is developed and incorporated into the SHETRAN system, groundwater level data for representative observation boreholes will be obtained from the BGS/EA and used to evaluate internal catchment model performance. These data would also allow for further assessment of the physical realism of the standard configuration of the current system.

- The performance of catchment models has not yet been evaluated with respect to groundwater level data from observation boreholes. This is a result of the absence of a nationally available dataset for use in this project, although future collaboration with BGS and EA may allow for a systematic comparison of observed and simulated groundwater levels. Multi-dataset evaluation has been shown to be very important for constraining model parameters (e.g. Anderton et al., 2002), which could add significantly to confidence in model behaviour or alternatively highlight areas of deficient data or locally inappropriate parameter sets.
- A robust method for delineating groundwater catchments where they differ from surface water catchments and applying appropriate subsurface boundary conditions would also be useful for improving the modelling system. It may well be possible to draw lessons from the EA's regional groundwater models covering the major aquifers in the UK (Shepley et al., 2012), where fluctuating groundwater catchment boundaries in some regions are accounted for using buffer zones to allow the model to calculate flow divides, thereby reducing the influence of specified boundary conditions.

General:

- No formal uncertainty analysis has been undertaken during the work to date, as the emphasis has been placed on sensitivity testing. As noted by Hrachowitz et al. (2013), uncertainty analysis is now widely recognised as an important component of the workflow in hydrological modelling, although there are ongoing debates regarding the best ways in which this should be done (e.g. Beven and Binley, 1992; Clark et al., 2011). It is anticipated that uncertainty analysis would be an interesting and important additional avenue of investigation for the modelling system, which could be explored in a different project with additional time and computing resources.
- The SHETRAN modelling system presented here has not been extensively evaluated with respect to other commonly applied models or indeed or other national models of Great Britain, such as the G2G (Bell et al., 2009) and CLASSIC-GB (Crooks et al., 2014) models. Some comparisons are made with the results from the Future Flows Project (Prudhomme et al., 2012), but more comprehensive comparisons of performance for a historical period would be a useful next step. As some of the existing national modelling systems are designed and calibrated with accuracy in mind, such as Habets et al. (2008) used for flood forecasting, it would be interesting to evaluate how large the differences in performance are between

a physically-based model using universal parameterisation and more conceptual approaches benefiting from more intensive calibration efforts.

- Work is ongoing to update the UKCP09 spatial weather generator. Upon completion of the new version of the weather generator, a full and rigorous climate climate impact study could be undertaken by using it in conjunction with the SHETRAN modelling system. Ideally this would include a large sample of catchments from across Great Britain, as well as a range of scenarios and time slices. This would provide a more comprehensive dataset for analysing possible hydrological futures and their variation across the country, as well as the effects of uncertainty from various sources.
- There are also current plans to couple SHETRAN for GB to a hydrodynamic model (Hi-PIMS, Liang and Smith (2015)) for modelling both fluvial and multi-source flooding. This could pave the way for creating a national groundwater flooding map derived from a physically-based model. Coupling with Hi-PIMS could also prompt exploration of whether SHETRAN could be written to run using GPUs, which would speed up run times and therefore enable the simulation of larger areas and/or finer resolutions, as well as more scenarios. Separate plans are being considered for coupling SHETRAN for GB to JULES, which could particularly improve PET estimates.
- The promising results from applying SHETRAN at the national scale in Great Britain also raise the question of whether the model could form a useful system in other countries globally, particularly given ongoing development of global datasets and technologies such as remote sensing. This would of course be a very long-term process requiring more computing power, while many issues could arise in terms of obtaining appropriate driving and validation data. However, if levels of performance of a physically-based modelling system can be related to catchment types in a similar way to the results presented here, it could be that SHETRAN could help to fill some of the gaps in particular data-poor contexts.

# References


Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connell, P. E., and Rasmussen, J. (1986). An introduction to the European Hydrological System – Système Hydrologique Europeén, (SHE), 1: History and philosophy of a physically-based, distributed modelling system. *Journal of Hydrology*, 87(12):45–59.

Adams, R. and Parkin, G. (2002). Development of a coupled surface-groundwater-pipe network model for the sustainable management of karstic groundwater. *Environmental Geology*, 42(5):513–517.

Ali, G., Tetzlaff, D., Soulsby, C., McDonnell, J. J., and Capell, R. (2012). A comparison of similarity indices for catchment classification using a cross-regional dataset. *Advances in Water Resources*, 40(0):11–22.

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M. (1998). Crop evapotranspiration - guidelines for computing crop water requirements - FAO irrigation and drainage paper 56. Report, FAO - Food and Agriculture Organization of the United Nations.

Anderton, S., Latron, J., and Gallart, F. (2002). Sensitivity analysis and multi-response, multi-criteria evaluation of a physically based distributed model. *Hydrological Processes*, 16(2):333–353.

Arnell, N. W. (1999). Climate change and global water resources. *Global Environmental Change*, 9:S31–S49.

Arnold, J. G., Srinivasan, R., Muttiah, R. S., and Williams, J. R. (1998). Large area hydrologic modeling and assessment part i: Model development 1. *Journal of the American Water Resources Association*, 34(1):73–89.

Bates, B., Kundzewicz, Z. W., Wu, S., Palutikof, J., et al. (2008). Climate change and water: Technical paper of the Intergovernmental Panel on Climate Change.

Bathurst, J., Ewen, J., Parkin, G., O'Connell, P., and Cooper, J. (2004). Validation of catchment models for predicting land-use and climate change impacts. 3. Blind validation for internal and outlet responses. *Journal of Hydrology*, 287(1):74–94.

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

Bell, V. A., Kay, A. L., Jones, R. G., and Moore, R. J. (2007b). Use of a grid-based hydrological model and regional climate model outputs to assess changing flood risk. *International Journal of Climatology*, 27(12):1657–1671.

Bell, V. A., Kay, A. L., Jones, R. G., Moore, R. J., and Reynard, N. S. (2009). Use of soil data in a grid-based hydrological model to estimate spatial variation in changing flood risk across the UK. *Journal of Hydrology*, 377(3–4):335–370.

Bell, V. A. and Moore, R. J. (1998). A grid-based distributed flood forecasting model for use with weather radar data: Part 1. Formulation. *Hydrology and Earth System Sciences*, 2(2-3):265–281.

Bergström, S., Singh, V., et al. (1995). The HBV model. In Singh, V., editor, *Computer Models of Watershed Hydrology.*, pages 443–476. Water Resources Publications, Fort Collins, Colorado U.S.A.

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

Beven, K. (1989). Changing ideas in hydrology – the case of physically-based models. *Journal of Hydrology*, 105(12):157–172.

Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources*, 16(1):41–51.

Beven, K. (1999a). How far can we go in distributed hydrological modelling? *Hydrology and Earth System Sciences*, 5(1):1–12.

Beven, K. (1999b). Uniqueness of place and process representations in hydrological modelling. *Hydrology and Earth System Sciences*, 4(2):203–213.

Beven, K. (2001). Dalton medal lecture: How far can we go in distributed hydrological modelling? *Hydrology and Earth System Sciences*, 5(1):1–12.

Beven, K. (2006). Searching for the holy grail of scientific hydrology. *Hydrology and Earth System Sciences Discussions*, 10(5):609–618.

Beven, K. (2012). *Rainfall-Runoff Modelling: The Primer*. Wiley, West Sussex, UK.

Beven, K. and Binley, A. (1992). The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, 6(3):279–298.

Beven, K., Calver, A., and Morris, E. (1987). The Institute of Hydrology distributed model. Technical report, Institute of Hydrology.

Beven, K. and Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, 249(14):11–29.

Beven, K. and Kirkby, M. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Journal*, 24(1):43–69.

Beven, K., Lamb, R., Quinn, P., Romanowicz, R., Freer, J., Singh, V., et al. (1995). Topmodel. In Singh, V., editor, *Computer models of watershed hydrology*, pages 627–668. Water Resources Publications, Fort Collins, Colorado U.S.A.

Beven, K. and O'Connell, P. (1982). On the role of physically-based distributed modelling in hydrology. Technical report, Institute of Hydrology.

Birkinshaw, S. J. (2010). Technical note: Automatic river network generation for a physically-based river catchment model. *Hydrology and Earth System Sciences*, 14(9):1767–1771.

Birkinshaw, S. J. (2011). SHETRAN version 4: data requirements, data processing and parameter values. Technical report, Newcastle University.

Birkinshaw, S. J., Bathurst, J. C., Iroumé, A., and Palacios, H. (2011). The effect of forest cover on peak flow and sediment discharge – an integrated field and modelling study in central–southern Chile. *Hydrological Processes*, 25(8):1284–1297.

Birkinshaw, S. J. and Webb, B. (2010). Flow pathways in the Slapton Wood catchment using temperature as a tracer. *Journal of Hydrology*, 383(3):269–279.

Blenkinsop, S., Lewis, E., Chan, S., and Fowler, H. (2016). Quality control of an hourly precipitation dataset and climatology of extremes for the UK. *Journal of Climatology*. In press.

Booker, D. J. and Woods, R. A. (2014). Comparing and combining physically-based and empirically-based approaches for estimating the hydrology of ungauged catchments. *Journal of Hydrology*, 508(0):227–239.

Boorman, D., Hollis, J., and Lilly, A. (1995). Hydrology of soil types: a hydrologically-based classification of the soils of the United Kingdom. *Institute of Hydrology Report*, 126.

Boughton, W. and Droop, O. (2003). Continuous simulation for design flood estimation– a review. *Environmental Modelling and Software*, 18(4):309–318.

Bovolo, C. and Bathurst, J. (2012). Modelling catchment-scale shallow landslide occurrence and sediment yield as a function of rainfall return period. *Hydrological Processes*, 26(4):579–596.

Box, G. and Jenkins, G. (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day Series in Time Series Analysis. Wiley, New Jersey.

British Geological Survey (2012). 1:625 000 scale digital hydrogeological data. `http://www.bgs.ac.uk/downloads/start.cfm?id=2258`. Accessed: 2012-02-24.

British Geological Survey (2014). 1:625 000 scale digital hydrogeological data. `http://www.bgs.ac.uk/products/hydrogeology/maps.html`. Accessed: 2014-09-19.

Brunner, P. and Simmons, C. T. (2012). Hydrogeosphere: a fully integrated, physically based hydrological model. *Groundwater*, 50(2):170–176.

Burnash, R., Ferral, R., McGuire, R., and McGuire, R. (1973). *A Generalized Streamflow Simulation System: Conceptual Modeling for Digital Computers*. U.S. Department of Commerce, National Weather Service, and State of California, Department of Water Resources.

Cameron, D., Beven, K. J., Tawn, J., Blazkova, S., and Naden, P. (1999). Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty). *Journal of Hydrology*, 219(3):169–187.

Centre for Ecology and Hydrology (2012). Land cover map 2007. `https://gateway.ceh.ac.uk/home`. Accessed: 2012-02-21.

Centre for Ecology and Hydrology (2014). Licence for access to data supplied by the National River Flow Archive from the Centre for Ecology and Hydrology. `http://www.ceh.ac.uk/data/nrfa/data/nrfa_catchment_licence.pdf`. Accessed: 2014-09-19.

Chan, S., Kendon, E., Fowler, H., Blenkinsop, S., and Roberts, N. (2014). Projected increases in summer and winter UK sub-daily precipitation extremes from high-resolution regional climate models. *Environmental Research Letters*, 9(8):084019.

Christensen, N. S., Wood, A. W., Voisin, N., Lettenmaier, D. P., and Palmer, R. N. (2004). The effects of climate change on the hydrology and water resources of the Colorado River basin. *Climatic Change*, 62(1-3):337–363.

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

Clark, M. P., Kavetski, D., and Fenicia, F. (2011). Pursuing the method of multiple working hypotheses for hydrological modeling. *Water Resources Research*, 47(9):W09301.

Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., Wagener, T., and Hay, L. E. (2008). Framework for understanding structural errors (FUSE): A modular framework to diagnose differences between hydrological models. *Water Resources Research*, 44(12):W00B02.

Coxon, G., Freer, J., Westerberg, I., Wagener, T., Woods, R., and Smith, P. (2015). A novel framework for discharge uncertainty quantification applied to 500 UK gauging stations. *Water Resources Research*, 51(7):5531–5546.

Crawford, N. H. and Linsley, R. K. (1966). Digital simulation in hydrology– Stanford Watershed Model 4. Technical report, Stanford University.

Crooks, S. M., Kay, A. L., Davies, H. N., and Bell, V. A. (2014). From catchment to national scale rainfall-runoff modelling: Demonstration of a hydrological modelling framework. *Hydrology*, 1(1):63–88.

Dawson, C. and Wilby, R. (2001). Hydrological modelling using artificial neural networks. *Progress in Physical Geography*, 25(1):80–108.

Dawson, C. W., Abrahart, R. J., Shamseldin, A. Y., and Wilby, R. L. (2006). Flood estimation at ungauged sites using artificial neural networks. *Journal of Hydrology*, 319(1):391–409.

Debele, B., Srinivasan, R., and Yves Parlange, J. (2007). Accuracy evaluation of weather data generation and disaggregation methods at finer timescales. *Advances in Water Resources*, 30(5):1286–1300.

Deckers, D. L., Booij, M. J., Rientjes, T. H., and Krol, M. S. (2010). Catchment variability and parameter estimation in multi-objective regionalisation of a rainfall–runoff model. *Water Resources Management*, 24(14):3961–3985.

Department for Communities and Local Government (2009). Planning policy statement 25: Development and flood risk practice guide. Technical report, Department for Communities and Local Government.

Department for Communities and Local Government (2014). Winter 2013/14 severe weather recovery progress report. Technical report, Department for Communities and Local Government.

Department for Environment, Food and Rural Affairs (2014). UK climate projections licence agreement. `http://www.metoffice.gov.uk/climatechange/`

science/monitoring/ukcp09/UKCIP08_license_agreement_130709.pdf. Accessed: 2014-09-19.

Dornes, P. F., Tolson, B. A., Davison, B., Pietroniro, A., Pomeroy, J. W., and Marsh, P. (2008). Regionalisation of land surface hydrological model parameters in subarctic and arctic environments. *Physics and Chemistry of the Earth, Parts A/B/C*, 33(17):1081–1089.

Downer, C. W. and Ogden, F. L. (2004). GSSHA: Model to simulate diverse stream flow producing processes. *Journal of Hydrologic Engineering*, 9(3):161–174.

Duan, Q., Sorooshian, S., and Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research*, 28(4):1015–1031.

Engeland, K., Gottschalk, L., and Tallaksen, L. (2001). Estimation of regional parameters in a macro scale hydrological model. *Nordic Hydrology*, 32(3):161–180.

Engman, E. T. (1986). Roughness coefficients for routing surface runoff. *Journal of Irrigation and Drainage Engineering*, 112(1):39–53.

Environment Agency (2003). Enhancements to MODFLOW. User guide for MODFLOW VKD – a modified version of MODFLOW-96 to include variations in hydraulic proerties with depth (user guide v24). Report, Environment Agency of England and Wales National Groundwater and Contaminated Land Centre.

Environment Agency (2014). 2m composite LIDAR digital terrain model. http://data.gov.uk/dataset/2m-composite-lidar-digital-terrain-model. Accessed: 2014-09-23.

European Commission, Joint Research Centre (2012). European soil database raster library 1km x 1km. http://eusoils.jrc.ec.europa.eu/ESDB_Archive/ESDB_data_1k_raster_intro/ESDB_1k_raster_data_intro.html. Accessed: 2012-02-24.

European Commission, Joint Research Centre (2014). European soil database raster library 1km x 1km: Notification regarding these data. http://eusoils.jrc.ec.europa.eu/library/Data/_Datarequest/ESDB_RasterLibrary.cfm. Accessed: 2014-09-19.

Ewen, J. (1990). Basis for the sub-surface contaminant migration component of the catchment water flow, contaminant transport, and contaminant migration modelling system SHETRAN-UK. In *NIREX Research Report NSS/R229*. NIREX Harwell.

Ewen, J., O'Connell, E., Bathurst, J. C., Birkinshaw, S. J., Kilsby, C. G., Parkin, G., and O'Donnell, G. (2012). Physically-based modelling, uncertainty, and pragmatism – comment on: 'Système Hydrologique Europeén (SHE): review and perspectives after 30 years development in distributed physically-based hydrological modelling' by Jens Christian Refsgaard, Børge Storm and Thomas Clausen. *Hydrology Research*, 43(6):945–947.

Ewen, J., Parkin, G., and O'Connell, P. E. (2000). SHETRAN: distributed river basin flow and transport modeling system. *Journal of Hydrologic Engineering*, 5(3):250–258.

Fang, X., Pomeroy, J., Ellis, C., MacDonald, M., DeBeer, C., and Brown, T. (2013). Multi-variable evaluation of hydrological model predictions for a headwater basin in the Canadian Rocky Mountains. *Hydrology and Earth System Sciences*, 17(4):1635–1659.

Fang, X., Pomeroy, J., Westbrook, C., Guo, X., Minke, A., and Brown, T. (2010). Prediction of snowmelt derived streamflow in a wetland dominated prairie basin. *Hydrology and Earth System Sciences*, 14(6):991–1006.

Fenicia, F., Kavetski, D., and Savenije, H. H. (2011). Elements of a flexible approach for conceptual hydrological modelling: 1. Motivation and theoretical development. *Water Resources Research*, 47(11):W11510.

Freeze, R. A. and Harlan, R. L. (1969). Blueprint for a physically-based, digitally-simulated hydrologic response model. *Journal of Hydrology*, 9(3):237–258.

Frey, B. J. and Dueck, D. (2007). Clustering by passing messages between data points. *Science*, 315(5814):972–976.

Gillespie, M. and Styles, M. (1999). BGS rock classification scheme, Volume 1. Classification of igneous rocks. Technical report, British Geological Survey.

Goderniaux, P., Brouyère, S., Fowler, H. J., Blenkinsop, S., Therrien, R., Orban, P., and Dassargues, A. (2009). Large scale surface–subsurface hydrological model to assess climate change impacts on groundwater reserves. *Journal of Hydrology*, 373(1):122–138.

Götzinger, J. and Bárdossy, A. (2007). Comparison of four regionalisation methods for a distributed hydrological model. *Journal of Hydrology*, 333(24):374–384.

Grayson, R. B., Moore, I. D., and McMahon, T. A. (1992). Physically based hydrologic modeling: 1. A terrain-based model for investigative purposes. *Water Resources Research*, 28(10):2639–2658.

Gupta, H., 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(12):80–91.

Gupta, H., Perrin, C., Bloschl, G., Montanari, A., Kumar, R., Clark, M., and Andréassian, V. (2014). Large-sample hydrology: a need to balance depth with breadth. *Hydrology and Earth System Sciences*, 18(2):463–477.

Ha, E. and Yoo, C. (2007). Use of mixed bivariate distributions for deriving inter-station correlation coefficients of rain rate. *Hydrological Processes*, 21(22):3078–3086.

Habets, F., Boone, A., Champeaux, J.-L., Etchevers, P., Franchisteguy, L., Leblois, E., Ledoux, E., Le Moigne, P., Martin, E., Morel, S., et al. (2008). The SAFRAN-ISBA-MODCOU hydrometeorological model applied over France. *Journal of Geophysical Research*, 113(D6):D06113.

Habib, E., Krajewski, W. F., and Ciach, G. J. (2001). Estimation of rainfall interstation correlation. *Journal of Hydrometeorology*, 2(6):621–629.

Hall, J. W., Dawson, R., Sayers, P., Rosu, C., Chatterton, J., and Deakin, R. (2003). A methodology for national-scale flood risk assessment. *Proceedings of the ICE-Water and Maritime Engineering*, 156(3):235–247.

Hallsworth, C. and Knox, R. (1999). BGS rock classification scheme. Volume 3, Classification of sediments and sedimentary rocks. Technical report, British Geological Survey.

He, Y., Bárdossy, A., and Zehe, E. (2011). A catchment classification scheme using local variance reduction method. *Journal of Hydrology*, 411(12):140–154.

Henriksen, H. J., Troldborg, L., Nyegaard, P., Sonnenborg, T. O., Refsgaard, J. C., and Madsen, B. (2003). Methodology for construction, calibration and validation of a national hydrological model for Denmark. *Journal of Hydrology*, 280(14):52–71.

Hock, R. (2003). Temperature index melt modelling in mountain areas. *Journal of Hydrology*, 282(1):104–115.

Hock, R. (2005). Glacier melt: a review of processes and their modelling. *Progress in Physical Geography*, 29(3):362–391.

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*, 1(2):227–239.

Howard, L. (1818). *The Climate of London: deduced from Meteorological observations,*

*made at different places in the neighbourhood of the metropolis*, volume 1. W. Phillips, London UK.

Hrachowitz, M., Savenije, H., Blöschl, G., McDonnell, J., Sivapalan, M., Pomeroy, J., Arheimer, B., Blume, T., Clark, M., Ehret, U., et al. (2013). A decade of Predictions in Ungauged Basins (PUB)– a review. *Hydrological Sciences Journal*, 58(6):1198–1255.

Hundecha, Y. and Bárdossy, A. (2004). Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model. *Journal of Hydrology*, 292(14):281–295.

Hundecha, Y., Ouarda, T. B. M. J., and Bárdossy, A. (2008). Regional estimation of parameters of a rainfall-runoff model at ungauged watersheds using the spatial structures of the parameters within a canonical physiographic-climatic space. *Water Resources Research*, 44(1):W01427.

Institute of Hydrology (1999). *Flood Estimation Handbook*. Institute of Hydrology, Wallingford, UK.

Jimenez Cisneros, B. E., Oki, T., Arnell, N. W., Benito, G., Cogley, J. G., Doll, P., Jiang, T., and Mwakalila, S. S. (2014). *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge.

Joint Nature Conservation Committee (2014). NVC survey data and distribution maps. `http://jncc.defra.gov.uk/page-4267`. Accessed: 2014-09-19.

Jones, P., Kilsby, C., Harpham, C., Glenis, V., and Burton, A. (2009). UK climate projections science report: Projections of future daily climate for the UK from the weather generator. Technical report, University of Newcastle, UK.

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

Kay, A. L. and Jones, D. A. (2010). Transient changes in flood frequency and timing in Britain under potential projections of climate change. *International Journal of Climatology*, 32:489–502.

Kendon, E. J., Roberts, N. M., Fowler, H. J., Roberts, M. J., Chan, S. C., and Senior, C. A. (2014). Heavier summer downpours with climate change revealed by weather forecast resolution model. *Nature Climate Change*, 4:570–576.

Kendon, E. J., Roberts, N. M., Senior, C. A., and Roberts, M. J. (2012). Realism of rainfall in a very high-resolution regional climate model. *Journal of Climate*, 25(17):5791–5806.

Kessler, H., Mathers, S., and Sobisch, H.-G. (2009). The capture and dissemination of integrated 3D geospatial knowledge at the British Geological Survey using GSI3D software and methodology. *Computers and Geosciences*, 35(6):1311–1321.

Kilsby, C. G., Jones, P. D., Burton, A., Ford, A. C., Fowler, H. J., Harpham, C., James, P., Smith, A., and Wilby, R. L. (2007). A daily weather generator for use in climate change studies. *Environmental Modelling and Software*, 22(12):1705–1719.

Kjeldsen, T. R. (2007). *Flood Estimation Handbook, Supplementary Report No. 1: The revitalised FSR/FEH rainfall-runoff method*. CEH, Wallingford.

Klemes, V. (1986). Operational testing of hydrological simulation models. *Hydrological Sciences Journal*, 31(1):13–24.

Kokkonen, T. S., Jakeman, A. J., Young, P. C., and Koivusalo, H. J. (2003). Predicting daily flows in ungauged catchments: model regionalization from catchment descriptors at the Coweeta Hydrologic Laboratory, North Carolina. *Hydrological Processes*, 17(11):2219–2238.

Koo, B. and O'Connell, P. (2006). An integrated modelling and multicriteria analysis approach to managing nitrate diffuse pollution: 2. A case study for a chalk catchment in England. *Science of the Total Environment*, 358(1):1–20.

Koutsoyiannis, D. (2003). Rainfall disaggregation methods: Theory and applications. In *Workshop on Statistical and Mathematical Methods for Hydrological Analysis, Rome*, volume 5270.

Lamb, R. (1999). Calibration of a conceptual rainfall-runoff model for flood frequency estimation by continuous simulation. *Water Resources Research*, 35(10):3103–3114.

Lawley, R. (2009). The soil-parent material database: a user guide. Technical report, British Geological Survey.

Lawley, R. and Garcia-Bajo, M. (2009). The national superficial deposit thickness model.(version 5). Technical report, British Geological Survey.

Ledbetter, R., Prudhomme, C., and Arnell, N. (2012). A method for incorporating climate variability in climate change impact assessments: Sensitivity of river flows in the Eden catchment to precipitation scenarios. *Climatic Change*, 113(3-4):803–823.

Li, M., Shao, Q., Zhang, L., and Chiew, F. H. S. (2010). A new regionalization approach and its application to predict flow duration curve in ungauged basins. *Journal of Hydrology*, 389(12):137–145.

Liang, Q. and Smith, L. S. (2015). A high-performance integrated hydrodynamic modelling system for urban flood simulations. *Journal of Hydroinformatics*, 4:518–533.

Liedekerke, M. V., Jones, A., and Panagos, P. (2006). ESDBv2 raster library - a set of rasters derived from the European Soil Database distribution v2.0. Technical report, European Commission and the European Soil Bureau Network.

Limbrick, K., Whitehead, P., Butterfield, D., and Reynard, N. (2000). Assessing the potential impacts of various climate change scenarios on the hydrological regime of the River Kennet at Theale, Berkshire, south-central England, UK: an application and evaluation of the new semi-distributed model, INCA. *Science of The Total Environment*, 251–252(0):539 – 555.

Lindstrom, G., Pers, C., Rosberg, J., Stromqvist, J., and Arheimer, B. (2010). Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales. *Hydrology Research*, 41(3–4):295–319.

Litzkow, M. J., Livny, M., and Mutka, M. W. (1988). Condor-a hunter of idle workstations. In *Distributed Computing Systems, 1988., 8th International Conference on*, pages 104–111. Institute of Electrical and Electronics Engineers.

Lopez, A., Fung, F., New, M., Watts, G., Weston, A., and Wilby, R. L. (2009). From climate model ensembles to climate change impacts and adaptation: A case study of water resource management in the southwest of England. *Water Resources Research*, 45(8):W08419.

Loucks, D. P., Van Beek, E., Stedinger, J. R., Dijkman, J. P., and Villars, M. T. (2005). *Water resources systems planning and management: an introduction to methods, models and applications*. UNESCO, Paris.

MacDonald, A. M. and Allen, D. J. (2001). Aquifer properties of the Chalk of England. *Quarterly Journal of Engineering Geology and Hydrogeology*, 34(4):371–384.

Mathers, S., Terrington, R., Waters, C., and Leslie, A. (2012). Metadata report for national bedrock fence diagram GB3D_v2012. Technical report, British Geological Survey.

Mathers, S., Terrington, R., Waters, C., and Leslie, A. (2014). GB3D–a framework for the bedrock geology of Great Britain. *Geoscience Data Journal*, 1(1):30–42.

McMillan, A. and Powell, J. (1999). BGS rock classification scheme. Volume 4, Classification of artificial (man-made) ground and natural superficial deposits: applications to geological maps and datasets in the UK. Technical report, British Geological Survey.

Mellor, D., Sheffield, J., O'Connell, P., and Metcalfe, A. (2000). A stochastic space-time rainfall forecasting system for real time flow forecasting ii: Application of SHETRAN and ARNO rainfall runoff models to the Brue catchment. *Hydrology and Earth System Sciences*, 4(4):617–626.

Merz, R. and Blöschl, G. (2004). Regionalisation of catchment model parameters. *Journal of Hydrology*, 287(1):95–123.

Met Office Integrated Data Archive System (2012). Land and marine surface stations data (1853-current), NCAS British Atmospheric Data Centre. `http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__dataent_ukmo-midas`. Accessed: 2012-03-01.

Michelson, D. B. (2004). Systematic correction of precipitation gauge observations using analyzed meteorological variables. *Journal of Hydrology*, 290(3):161–177.

Montanari, A. and Koutsoyiannis, D. (2012). A blueprint for process-based modeling of uncertain hydrological systems. *Water Resources Research*, 48(9):W09555.

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

Morris, D. G., Flavin, R. W., and Moore, R. V. (1990). *A digital terrain model for hydrology*. Proc 4th Int. Symposium on Spatial Data Handling, Zurich.

Morton, D., Rowland, C., Wood, C., Meek, L., Marston, C., Smith, G., Wadsworth, R., and Simpson, I. C. (2011). Countryside survey: Final report for LCM2007 the new UK land cover map. Technical report, Centre for Ecology and Hydrology.

Murphy, J. M., Sexton, D., Jenkins, G., Booth, B., Brown, C., Clark, R., Collins, M., Harris, G., Kendon, E., Betts, R., et al. (2009). UK climate projections science report: climate change projections. Technical report, Meteorological Office Hadley Centre.

Nachtergaele, F. and Batjes, N. (2012). *Harmonized world soil database*. FAO, Rome.

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):282–290.

National River Flow Archive (2012). Catchment data. `http://www.ceh.ac.uk/data/nrfa/data/search.html`. Accessed: 2012-03-07.

National River Flow Archive (2014a). Gauged daily flows. http://www.ceh.ac.uk/data/nrfa/data/gauged_flow.html. Accessed: 2014-09-19.

National River Flow Archive (2014b). Hydrometric areas for Great Britain and Northern Ireland. https://catalogue.ceh.ac.uk/documents/1957166d-7523-44f4-b279-aa5314163237. Accessed: 2014-09-19.

National Soil Resources Institute (2001). Natmap soilscapes: Easy soils map. http://www.landis.org.uk/data/nmsoilscapes.cfm. Accessed: 2014-09-19.

Natural England (2014). Natural England GIS digital boundary datasets. http://www.gis.naturalengland.org.uk/pubs/gis/GIS_register.asp. Accessed: 2014-09-19.

NERC (1975). *The Flood Studies Report (5 Volumes)*. Natural Environment Research Council, London.

O'Connell, P. (1991). A historical perspective. In *Recent Advances in the Modeling of Hydrologic Systems*, pages 3–30. Springer.

Ordnance Survey (2012). Mapping data and geographic information from Ordnance Survey. https://www.ordnancesurvey.co.uk/opendatadownload/products.html. Accessed: 2012-03-01.

Ordnance Survey (2013a). Land-form PANORAMA user guide and technical specification v5.2. Technical report, Ordnance Survey.

Ordnance Survey (2013b). Meridian 2 user guide and technical specification v6.0. Technical report, Ordnance Survey.

Ordnance Survey (2013c). OS terrain 50 user guide and technical specification v1.0. Technical report, Ordnance Survey.

Ordnance Survey (2014). OS open data licence. http://www.ordnancesurvey.co.uk/docs/licences/os-opendata-licence.pdf. Accessed: 2014-09-19.

Orellana, B., Pechlivanidis, I., McIntyre, N., Wheater, H., and Wagener, T. (2008). A toolbox for the identification of parsimonious semi-distributed rainfall-runoff models: Application to the upper Lee catchment. In *International Congress on Environmental Modelling and Software*, pages 1–8, Barcelona, Spain.

Ouarda, T. B. M. J., Cunderlik, J. M., St-Hilaire, A., Barbet, M., Bruneau, P., and Bobe, B. (2006). Data-based comparison of seasonality-based regional flood frequency methods. *Journal of Hydrology*, 330(1-2):329–339.

Oudin, L., Kay, A., Andrassian, V., and Perrin, C. (2010). Are seemingly physically similar catchments truly hydrologically similar? *Water Resources Research*, 46(11):W11558.

Pachauri, R. K., Allen, M., Barros, V., Broome, J., Cramer, W., Christ, R., Church, J., Clarke, L., Dahe, Q., Dasgupta, P., et al. (2014). Climate change 2014: Synthesis report. Contribution of working groups i, ii and iii to the fifth assessment report of the Intergovernmental Panel on Climate Change. Technical report, IPCC.

Paniconi, C. and Putti, M. (2015). Physically based modeling in catchment hydrology at 50: Survey and outlook. *Water Resources Research*, 51(9):7090–7129.

Parkin, G., Birkinshaw, S., Younger, P., Rao, Z., and Kirk, S. (2007). A numerical modelling and neural network approach to estimate the impact of groundwater abstractions on river flows. *Journal of Hydrology*, 339(1):15–28.

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):193–214.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Perry, M. and Hollis, D. (2005). The generation of monthly gridded datasets for a range of climatic variables over the UK. *International Journal of Climatology*, 25(8):1041–1054.

Perry, M., Hollis, D., and Elms, M. (2009). The generation of daily gridded datasets of temperature and rainfall for the UK. Technical report, Met Office National Climate Information Centre.

Price, D., Hudson, K., Boyce, G., Schellekens, J., Moore, R. J., Clark, P., Harrison, T., Connolly, E., and Pilling, C. (2012). Operational use of a grid-based model for flood forecasting. *Proceedings of the ICE-Water Management*, 165(2):65–77.

PricewaterhouseCoopers (2014). Revised estimated cost of continuing UK floods. `http://pwc.blogs.com/press_room/2014/02/pwc-revise-estimated-cost-of-continuing-uk-floods-to-630m.html`. Accessed: 2015-12-01.

Prudhomme, C., Dadson, S., Morris, D., Williamson, J., Goodsell, G., Crooks, S., Boelee, L., Davies, H., Buys, G., Lafon, T., et al. (2012). Future flows climate: an ensemble of 1-km climate change projections for hydrological application in Great Britain. *Earth System Science Data Discussions*, 5(1):475–490.

Prudhomme, C., Haxton, T., Crooks, S., Jackson, C., Barkwith, A., Williamson, J., Kelvin, J., Mackay, J., Wang, L., and Young, A. (2013). Future flows hydrology: an ensemble of daily river flow and monthly groundwater levels for use for climate change impact assessment across Great Britain. *Earth System Science Data*, 5(1):101–107.

Pui, A., Sharma, A., Mehrotra, R., Sivakumar, B., and Jeremiah, E. (2012). A comparison of alternatives for daily to sub-daily rainfall disaggregation. *Journal of Hydrology*, 470:138–157.

Rahiz, M. and New, M. (2013a). 21st century drought scenarios for the UK. *Water Resources Management*, 27(4):1039–1061.

Rahiz, M. and New, M. (2013b). Does a rainfall-based drought index simulate hydrological droughts? *International Journal of Climatology*, 34(9):2853–2871.

Refsgaard, J. (1996). Terminology, modelling protocol and classification of hydrological model codes. In *Distributed hydrological modelling*, pages 17–39. Springer, Netherlands.

Refsgaard, J. and Storm, B. (1995). MIKE SHE. In Singh, V., editor, *Computer models of watershed hydrology*, pages 1–8, Fort Collins, Colorado U.S.A. Water Resources Publications.

Refsgaard, J., Storm, B., and Clausen, T. (2010). Système Hydrologique Européen (SHE): review and perspectives after 30 years development in distributed physically-based hydrological modelling. *Hydrology Research*, 41(5):355–377.

Refsgaard, J., Storm, B., and Clausen, T. (2012). Physically-based modelling, good modelling practice including uncertainty – reply to comment by Ewen et al . (2012). *Hydrology Research*, 43(6):948–950.

Rhodes, B. C. (2008). Pyephem. `http://rhodesmill.org/pyephem/`. Accessed: 2014-09-23.

Robertson, S. (1999). BGS rock classification scheme. Volume 2, Classification of metamorphic rocks. Technical report, British Geological Survey.

Rockwood, D., Davis, E., and Anderson, J. (1972). *User Manual for COSSARR Model*. U.S. Army Engineering Division, North Pacific.

Rutter, N., Essery, R., Pomeroy, J., Altimir, N., Andreadis, K., Baker, I., Barr, A., Bartlett, P., Boone, A., Deng, H., et al. (2009). Evaluation of forest snow processes models (SnowMIP2). *Journal of Geophysical Research: Atmospheres*, 114(D6):D06111.

Schulla, J. and Jasper, K. (2015). Model description WaSiM-ETH. Technical report, Institute for Atmospheric and Climate Science, Swiss Federal Institute of Technology, Zürich.

Seibert, J. (1999). Regionalisation of parameters for a conceptual rainfall-runoff model. *Agricultural and Forest Meteorology*, 98:279–293.

Seibert, J. (2003). Reliability of model predictions outside calibration conditions. *Nordic Hydrology*, 34(5):477–492.

Seibert, J. and McDonnell, J. J. (2010). Land-cover impacts on streamflow: a change-detection modelling approach that incorporates parameter uncertainty. *Hydrological Sciences Journal*, 55(3):316–332.

Serinaldi, F. (2008). Analysis of inter-gauge dependence by kendalls $\tau$k, upper tail dependence coefficient, and 2-copulas with application to rainfall fields. *Stochastic Environmental Research and Risk Assessment*, 22(6):671–688.

Shaw, E. M., Beven, K. J., Chappell, N. A., and Lamb, R. (2010). *Hydrology in practice*. Spon Press, Oxford, UK.

Shepley, M. G., Whiteman, M., Hulme, P., and Grout, M. (2012). Introduction: groundwater resources modelling: a case study from the UK. *Geological Society, London, Special Publications*, 364(1):1–6.

Sherman, L. (1932). Streamflow from rainfall by unit-graph method. *Engineering News Record*, 108:501–505.

Sibson, R. (1981). A brief description of natural neighbour interpolation. *Interpreting Multivariate Data*, 21:21–36.

Simpson, I. and Jones, P. (2014). Analysis of UK precipitation extremes derived from Met Office gridded data. *International Journal of Climatology*, 34(7):2438–2449.

Singh, R., Archfield, S. A., and Wagener, T. (2014). Identifying dominant controls on hydrologic parameter transfer from gauged to ungauged catchments: a comparative hydrology approach. *Journal of Hydrology*, 517:985–996.

Singh, V. (1995). *Computer models of watershed hydrology*. Water Resources Publications, Fort Collins, Colorado U.S.A.

Sittner, W. T., Schauss, C. E., and Monro, J. C. (1969). Continuous hydrograph synthesis with an API-type hydrologic model. *Water Resources Research*, 5(5):1007–1022.

Sivapalan, M. (2003). Prediction in ungauged basins: a grand challenge for theoretical hydrology. *Hydrological Processes*, 17(15):3163–3170.

Strömqvist, J., Arheimer, B., Dahné, J., Donnelly, C., and Lindström, G. (2012). Water and nutrient predictions in ungauged basins: set-up and evaluation of a model at the national scale. *Hydrological Sciences Journal*, 57(2):229–247.

Sugawara, M., Watanabe, I., Ozaki, E., and Katsuyame, Y. (1983). *Reference Manual for the TANK Model*. National Research Center for Disaster Prevention, Tokyo.

Tanguy, M., Dixon, H., Prosdocimi, I., Morris, D. G., and Keller, V. D. J. (2014). Gridded estimates of daily and monthly areal rainfall for the United Kingdom (1890-2012) [CEH-GEAR]. Technical report, NERC Environmental Information Data Centre.

The Met. Office (2012). UKCP09: Download data sets. `http://www.metoffice.gov.uk/climatechange/science/monitoring/ukcp09/download/index.html`. Accessed: 2012-02-08.

The National Archives (2014). Non-commercial government licence for public sector information. `http://www.nationalarchives.gov.uk/doc/non-commercial-government-licence/non-commercial-government-licence.htm`. Accessed: 2014-09-19.

Therrien, R., McLaren, R., Sudicky, E., and Panday, S. (2010). Hydrogeosphere: A three-dimensional numerical model describing fully-integrated subsurface and surface flow and solute transport. Technical report, Groundwater Simulations Group, University of Waterloo.

Todini, E. (1996). The ARNO rainfall–runoff model. *Journal of Hydrology*, 175(1):339–382.

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):892–898.

Vogel, R. (2005). *Regional calibration of watershed models*. CRC Press, Boca Raton.

Vörösmarty, C. J., Green, P., Salisbury, J., and Lammers, R. B. (2000). Global water resources: vulnerability from climate change and population growth. *Science*, 289(5477):284–288.

Walsh, C. and Kilsby, C. (2007). Implications of climate change on flow regime affecting Atlantic salmon. *Hydrology and Earth System Sciences*, 11(3):1127–1143.

Water Framework Directive (2000). EU Water Framework Directive. Technical report, Directive 2000/60/EC.

Watson, C., Richardson, J., Wood, B., Jackson, C., and Hughes, A. (2015). Improving geological and process model integration through TIN to 3D grid conversion. *Computers & Geosciences*, 82:45–54.

Wheater, H. (2002). Progress in and prospects for fluvial flood modelling. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 360(1796):1409–1431.

Wheater, H. S., Jakeman, A. J., and Beven, K. J. (1993). Progress and directions in rainfall-runoff modelling. In Jakeman, A. J., Beck, M. B., and McAleer, M. J., editors, *Modelling Change in Environmental Systems*, pages 101–132. Wiley.

Wilby, R. L. and Dessai, S. (2010). Robust adaptation to climate change. *Weather*, 65(7):180–185.

Wilby, R. L. and Harris, I. (2006). A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK. *Water Resources Research*, 42(2):W02419.

Yoo, C. and Ha, E. (2007). Effect of zero measurements on the spatial correlation structure of rainfall. *Stochastic Environmental Research and Risk Assessment*, 21(3):287–297.

Young, P. C. and Beven, K. J. (1994). Data-based mechanistic modelling and the rainfall-flow non-linearity. *Environmetrics*, 5(3):335–363.

Zhang, G. (2007). *Modelling Hydrological Response at the Catchment Scale*. Eburon Uitgeverij BV.

Zhang, R. (2012). Impacts of spatial and temporal scales on a distributed hydrological model. (Unpublished). Institute of Mediterranean Agrarian and Environmental Sciences, University of Évora.

Zhang, R., Santos, C. A., Moreira, M., Freire, P. K., and Corte-Real, J. (2013). Automatic calibration of the SHETRAN hydrological modelling system using MSCE. *Water Resources Management*, 27(11):4053–4068.

Zuzel, J. F. and Cox, L. M. (1975). Relative importance of meteorological variables in snowmelt. *Water Resources Research*, 11(1):174–176.

# Appendices

# Appendix A

# Catchment characteristics and results for standard simulation of SHETRAN for GB

*Table A.1: Catchment characteristics and NSE for standard run.*

| Catchment | Station | Area | Avg. Ann. Rainfall (mm) | Avg. Ann PET (mm) | DPSBAR | BFI | mean flow (mm/day) | NSE |
|---|---|---|---|---|---|---|---|---|
| Aire at Kildwick Bridge | 27035 | 284 | 1211 | 514 | 47 | 0.37 | 1.96 | 0.89 |
| Aldbourne at Ramsbury | 39101 | 55 | 851 | 598 | 32 | 0.97 | 0.37 | -12.36 |
| Allan Water at Kinbuck | 18001 | 163 | 1409 | 504 | 57 | 0.44 | 2.79 | 0.76 |
| Allen at Walford Mill | 43018 | 174 | 899 | 612 | 24 | 0.91 | 0.96 | -1.96 |
| Allt Deveron at Cabrach | 9005 | 67 | 1166 | 490 | 58 | 0.49 | 2.06 | 0.57 |
| Almond at Almondbank | 15013 | 178 | 1560 | 493 | 125 | 0.44 | 2.59 | 0.6 |
| Almond at Craigiehall | 19001 | 373 | 947 | 529 | 22 | 0.39 | 1.43 | 0.79 |
| Alt at Kirkby | 69032 | 98 | 876 | 615 | 7 | 0.53 | 1.24 | 0.23 |
| Ancholme at Toft Newton | 29009 | 27 | 622 | 572 | 11 | 0.53 | 0.44 | 0.68 |
| Anker at Polesworth | 28026 | 370 | 671 | 580 | 12 | 0.51 | 0.72 | 0.76 |
| Annan at Brydekirk | 78003 | 924 | 1386 | 484 | 59 | 0.43 | 2.84 | 0.86 |
| Ardle at Kindrogan | 15014 | 107 | 1400 | 476 | 98 | 0.4 | 2.63 | 0.53 |
| Arrow at Broom | 54007 | 322 | 712 | 599 | 18 | 0.53 | 0.75 | 0.65 |
| Arrow at Titley Mill | 55013 | 127 | 1053 | 569 | 64 | 0.56 | 1.59 | 0.67 |
| Avon at Amesbury | 43005 | 325 | 794 | 602 | 19 | 0.91 | 0.93 | -0.45 |
| Avon at Delnashaugh | 8004 | 542 | 1212 | 477 | 90 | 0.55 | 2.36 | 0.6 |
| Avon at Polmonthill | 17005 | 196 | 988 | 529 | 24 | 0.41 | 1.83 | 0.64 |
| Avon at Stareton | 54019 | 351 | 673 | 582 | 13 | 0.49 | 0.63 | 0.9 |
| Axe at Whitford | 45004 | 290 | 1070 | 602 | 42 | 0.48 | 1.59 | 0.79 |
| Ayr at Mainholm | 83006 | 575 | 1258 | 512 | 34 | 0.29 | 2.38 | 0.71 |
| Babingley at Castle Rising | 33054 | 48 | 681 | 584 | 16 | 0.95 | 0.91 | -1 |
| Bain at Goulceby Bridge | 30011 | 67 | 737 | 562 | 20 | 0.74 | 0.44 | -3.4 |
| Bedburn Beck at Bedburn | 24004 | 77 | 968 | 516 | 65 | 0.46 | 1.38 | 0.69 |
| Bedford Ouse at Bedford | 33002 | 1474 | 661 | 596 | 13 | 0.53 | 0.62 | 0.86 |

*Table A.1: Catchment characteristics and NSE for standard run.*

| Catchment | Station | Area | Avg. Ann. Rainfall (mm) | Avg. Ann PET (mm) | DPSBAR | BFI | mean flow (mm/day) | NSE |
|---|---|---|---|---|---|---|---|---|
| Bedford Ouse at Roxton | 33039 | 1671 | 652 | 596 | 13 | 0.57 | 0.59 | 0.82 |
| Bervie at Inverbervie | 13001 | 126 | 880 | 494 | 43 | 0.56 | 1.48 | 0.69 |
| Beult at Stile Bridge | 40005 | 279 | 710 | 608 | 11 | 0.23 | 0.64 | 0.87 |
| Beverley Brook at Wimbledon Common | 39005 | 38 | 637 | 658 | 13 | 0.65 | 1.22 | -0.42 |
| Black Cart Water at Milliken Park | 84017 | 107 | 1787 | 511 | 52 | 0.37 | 3.81 | 0.92 |
| Blackwater at Loch Dee | 80006 | 14 | 2581 | 498 | 95 | 0.46 | 7.31 | 0.69 |
| Blackwater at Swallowfield | 39007 | 358 | 723 | 613 | 14 | 0.67 | 0.75 | 0.38 |
| Blyth at Hartford Bridge | 22006 | 273 | 716 | 530 | 15 | 0.36 | 0.67 | 0.83 |
| Bollin at Wilmslow | 69012 | 74 | 931 | 572 | 37 | 0.61 | 1.47 | 0.46 |
| Box at Polstead | 36003 | 57 | 582 | 595 | 14 | 0.63 | 0.33 | 0.03 |
| Boyd at Bitton | 53017 | 49 | 843 | 613 | 32 | 0.44 | 0.97 | 0.83 |
| Braan at Hermitage | 15023 | 211 | 1436 | 492 | 88 | 0.43 | 2.86 | 0.79 |
| Brain at Guithavon Valley | 37009 | 64 | 601 | 596 | 13 | 0.67 | 0.52 | 0.3 |
| Bran at Dosmucheran | 4006 | 136 | 2045 | 472 | 117 | 0.28 | 4.45 | 0.81 |
| Brett at Cockfield | 36009 | 24 | 618 | 592 | 10 | 0.32 | 0.46 | 0.61 |
| Brock at U/S A6 | 72007 | 33 | 1403 | 535 | 62 | 0.34 | 2.24 | 0.69 |
| Brompton Beck at Snainton Ings | 27073 | 13 | 784 | 539 | 37 | 0.91 | 1.78 | -0.38 |
| Browney at Burn Hall | 24005 | 182 | 752 | 536 | 39 | 0.5 | 0.79 | 0.75 |
| Brue at Lovington | 52010 | 140 | 895 | 608 | 27 | 0.48 | 1.18 | 0.84 |
| Bull at Lealands | 41029 | 40 | 838 | 598 | 21 | 0.37 | 0.96 | 0.83 |
| Bure at Ingworth | 34003 | 171 | 698 | 589 | 11 | 0.83 | 0.58 | -5.86 |
| Calder at Whalley Weir | 71004 | 319 | 1216 | 529 | 51 | 0.42 | 2.31 | 0.76 |
| Canons Brook at Elizabeth Way | 38007 | 20 | 656 | 607 | 13 | 0.4 | 0.76 | 0.46 |
| Carron at Headswood | 17001 | 125 | 1582 | 515 | 57 | 0.34 | 2.42 | 0.7 |
| Carron at New Kelso | 93001 | 140 | 2541 | 461 | 160 | 0.26 | 6.63 | 0.62 |
| Cary at Somerton | 52011 | 84 | 732 | 621 | 15 | 0.38 | 0.84 | 0.88 |
| Ceiriog at Brynkinalt Weir | 67005 | 111 | 1290 | 547 | 89 | 0.54 | 2.42 | 0.79 |
| Chater at Fosters Bridge | 31010 | 69 | 691 | 580 | 26 | 0.53 | 0.64 | 0.81 |
| Chelmer at Churchend | 37011 | 76 | 616 | 589 | 13 | 0.43 | 0.40 | 0.59 |
| Chelmer at Springfield | 37008 | 194 | 602 | 590 | 13 | 0.57 | 0.47 | 0.58 |
| Cheriton Stream at Sewards Bridge | 42008 | 84 | 933 | 599 | 25 | 0.96 | 0.67 | -16.22 |
| Cherwell at Enslow Mill | 39021 | 558 | 700 | 590 | 20 | 0.66 | 0.59 | 0.15 |
| Chess at Rickmansworth | 39088 | 97 | 784 | 603 | 27 | 0.95 | 0.51 | -18.69 |
| Clyde at Blairston | 84005 | 1725 | 1208 | 502 | 43 | 0.45 | 2.16 | 0.73 |
| Cocker at Southwaite Bridge | 75004 | 119 | 1945 | 519 | 135 | 0.43 | 3.93 | 0.85 |
| Coln at Bibury | 39020 | 115 | 882 | 590 | 35 | 0.93 | 1.02 | -0.46 |
| Colne at Lexden | 37005 | 236 | 593 | 598 | 14 | 0.52 | 0.39 | 0.81 |
| Conder at Galgate | 72014 | 28 | 1175 | 531 | 48 | 0.35 | 2.10 | 0.8 |
| Coquet at Morwick | 22001 | 582 | 897 | 504 | 50 | 0.44 | 1.28 | 0.79 |
| Cothi at Felin Mynachdy | 60002 | 294 | 1675 | 567 | 64 | 0.42 | 3.40 | 0.85 |
| Cray at Crayford | 40016 | 125 | 696 | 629 | 22 | 0.73 | 0.34 | -19.39 |
| Cree at Newton Stewart | 81002 | 371 | 1811 | 507 | 65 | 0.27 | 3.66 | 0.77 |
| Creedy at Cowley | 45012 | 265 | 937 | 601 | 32 | 0.46 | 1.19 | 0.75 |
| Crimple at Burn Bridge | 27051 | 8 | 801 | 555 | 27 | 0.31 | 1.14 | 0.61 |
| Culm at Wood Mill | 45003 | 225 | 1010 | 596 | 34 | 0.54 | 1.43 | 0.71 |
| Currypool Stream at Currypool Farm | 52016 | 18 | 983 | 611 | 62 | 0.71 | 1.00 | -0.33 |

*Table A.1: Catchment characteristics and NSE for standard run.*

| Catchment | Station | Area | Avg. Ann. Rainfall (mm) | Avg. Ann PET (mm) | DPSBAR | BFI | mean flow (mm/day) | NSE |
|---|---|---|---|---|---|---|---|---|
| Cynon at Abercynon | 57004 | 107 | 1913 | 562 | 81 | 0.4 | 3.48 | 0.83 |
| Dane at Rudheath | 68003 | 405 | 861 | 560 | 26 | 0.52 | 1.05 | 0.85 |
| Darent at Hawley | 40012 | 184 | 741 | 603 | 31 | 0.73 | 0.30 | -11.71 |
| Dart at Austins Bridge | 46003 | 254 | 1857 | 581 | 56 | 0.52 | 3.84 | 0.7 |
| Dean Water at Cookston | 15008 | 178 | 837 | 516 | 31 | 0.59 | 1.29 | 0.39 |
| Dearne at Barnsley Weir | 27023 | 119 | 786 | 560 | 37 | 0.48 | 1.01 | 0.64 |
| Dee at Manley Hall | 67015 | 1019 | 1448 | 541 | 78 | 0.53 | 2.64 | 0.9 |
| Dee at Mar Lodge | 12007 | 295 | 1493 | 453 | 127 | 0.44 | 3.62 | 0.62 |
| Dee at New Inn | 67018 | 55 | 2035 | 532 | 88 | 0.27 | 4.86 | 0.77 |
| Dee at Polhollick | 12003 | 699 | 1414 | 462 | 119 | 0.49 | 2.90 | 0.6 |
| Derwent at Camerton | 75002 | 660 | 1818 | 519 | 103 | 0.49 | 3.50 | 0.89 |
| Derwent at Chatsworth | 28043 | 345 | 1219 | 538 | 78 | 0.55 | 1.58 | 0.63 |
| Deveron at Muiresk | 9002 | 970 | 967 | 509 | 48 | 0.57 | 1.52 | 0.79 |
| Don at Doncaster | 27021 | 1272 | 827 | 573 | 37 | 0.56 | 1.08 | 0.72 |
| Don at Parkhill | 11001 | 1276 | 944 | 495 | 53 | 0.68 | 1.43 | 0.56 |
| Doniford Stream at Swill Bridge | 51001 | 75 | 953 | 608 | 50 | 0.68 | 1.20 | 0.28 |
| Dove at Izaak Walton | 28046 | 86 | 1139 | 533 | 57 | 0.79 | 1.92 | 0.76 |
| Dove at Kirkby Mills | 27042 | 63 | 974 | 531 | 86 | 0.58 | 1.52 | 0.72 |
| Dove at Marston on Dove | 28018 | 884 | 966 | 538 | 38 | 0.61 | 1.36 | 0.82 |
| Dove at Rocester Weir | 28008 | 397 | 1057 | 534 | 48 | 0.62 | 1.62 | 0.79 |
| Dover Beck at Lowdham | 28060 | 72 | 681 | 552 | 23 | 0.75 | 0.19 | -24.2 |
| Dowles Brook at Oak Cottage | 54034 | 41 | 718 | 589 | 35 | 0.4 | 0.77 | 0.64 |
| Duddon at Duddon Hall | 74001 | 90 | 2263 | 524 | 137 | 0.29 | 4.82 | 0.81 |
| Dulnain at Balnaan Bridge | 8009 | 270 | 1111 | 484 | 65 | 0.44 | 1.94 | 0.61 |
| Dun at Hungerford | 39028 | 101 | 836 | 596 | 20 | 0.95 | 0.60 | -14.16 |
| Duneaton at Maidencots | 84022 | 117 | 1389 | 488 | 45 | 0.42 | 2.37 | 0.65 |
| Dwyfawr at Garndolbenmaen | 65007 | 53 | 2092 | 556 | 94 | 0.38 | 4.29 | 0.81 |
| Dyfi at Dyfi Bridge | 64001 | 467 | 1978 | 543 | 109 | 0.39 | 4.32 | 0.87 |
| Earn at Forteviot Bridge | 16004 | 792 | 1467 | 500 | 93 | 0.52 | 3.17 | 0.89 |
| East Avon at Upavon | 43014 | 89 | 821 | 598 | 19 | 0.89 | 0.79 | -3.35 |
| East Dart at Bellever | 46005 | 23 | 2290 | 558 | 60 | 0.44 | 4.68 | 0.68 |
| Eastburn Beck at Crosshills | 27084 | 45 | 1117 | 510 | 73 | 0.35 | 1.66 | 0.8 |
| Eastwood Brook at Eastwood | 37033 | 10 | 556 | 654 | 15 | 0.32 | 0.42 | -0.1 |
| Eden at Kirkby Stephen | 76014 | 70 | 1470 | 499 | 89 | 0.26 | 3.23 | 0.78 |
| Eden at Temple Sowerby | 76005 | 616 | 1181 | 507 | 53 | 0.37 | 2.08 | 0.81 |
| Ehen at Braystones | 74005 | 134 | 1771 | 520 | 89 | 0.42 | 3.37 | 0.71 |
| Ellen at Bullgill | 75017 | 105 | 1179 | 533 | 45 | 0.49 | 1.91 | 0.75 |
| Elwy at Pont-y-Gwyddel | 66006 | 192 | 1245 | 556 | 57 | 0.45 | 1.93 | 0.85 |
| Enborne at Brimpton | 39025 | 152 | 824 | 604 | 19 | 0.53 | 0.74 | 0.53 |
| Enrick at Mill of Tore | 6008 | 110 | 1360 | 490 | 71 | 0.3 | 2.60 | 0.83 |
| Erch at Pencaenewydd | 65005 | 21 | 1409 | 553 | 78 | 0.54 | 2.47 | 0.37 |
| Ericht at Craighall | 15025 | 443 | 1280 | 483 | 90 | 0.48 | 2.53 | 0.75 |
| Eridge Stream at Hendal Bridge | 40020 | 55 | 888 | 601 | 28 | 0.45 | 1.12 | 0.74 |
| Esk at Canonbie | 77002 | 499 | 1485 | 473 | 65 | 0.38 | 3.05 | 0.81 |
| Esk at Cropple How | 74007 | 70 | 2313 | 514 | 132 | 0.3 | 5.54 | 0.75 |
| Ewe at Poolewe | 94001 | 441 | 2318 | 466 | 129 | 0.64 | 5.83 | 0.9 |

*Table A.1: Catchment characteristics and NSE for standard run.*

| Catchment | Station | Area | Avg. Ann. Rainfall (mm) | Avg. Ann PET (mm) | DPSBAR | BFI | mean flow (mm/day) | NSE |
|---|---|---|---|---|---|---|---|---|
| Ewelme Brook at Ewelme | 39065 | 12 | 715 | 602 | 33 | 0.98 | 0.32 | -26.7 |
| Ewenny at Keepers Lodge | 58009 | 66 | 1367 | 585 | 42 | 0.57 | 2.41 | 0.5 |
| Exe at Pixton | 45009 | 149 | 1425 | 585 | 66 | 0.52 | 2.57 | 0.4 |
| Exe at Thorverton | 45001 | 607 | 1312 | 591 | 48 | 0.5 | 2.27 | 0.68 |
| Fal at Tregony | 48003 | 91 | 1277 | 599 | 32 | 0.68 | 1.91 | 0.16 |
| Falloch at Glen Falloch | 85003 | 82 | 2832 | 472 | 154 | 0.16 | 6.29 | 0.68 |
| Feugh at Heugh Head | 12008 | 231 | 1193 | 491 | 85 | 0.45 | 2.19 | 0.63 |
| Fowey at Restormel | 48011 | 169 | 1491 | 585 | 45 | 0.62 | 2.43 | 0.08 |
| Frome at Ebley Mill | 54027 | 199 | 858 | 597 | 48 | 0.87 | 1.09 | 0.62 |
| Frome at Yarkhill | 55018 | 148 | 748 | 596 | 28 | 0.52 | 0.67 | 0.75 |
| Gairn at Invergairn | 12006 | 153 | 1220 | 476 | 116 | 0.54 | 2.20 | 0.17 |
| Gannel at Gwills | 49004 | 43 | 1132 | 604 | 26 | 0.67 | 1.42 | -0.38 |
| Gifford Water at Lennoxlove | 20007 | 64 | 823 | 508 | 46 | 0.58 | 1.03 | 0.62 |
| Gipping at Stowmarket | 35008 | 127 | 597 | 594 | 10 | 0.39 | 0.43 | 0.7 |
| Glaslyn at Beddgelert | 65001 | 68 | 3032 | 543 | 162 | 0.32 | 7.30 | 0.74 |
| Goyt at Marple Bridge | 69017 | 185 | 1182 | 554 | 69 | 0.53 | 1.73 | 0.61 |
| Great Eau at Claythorpe Mill | 29002 | 78 | 719 | 558 | 22 | 0.88 | 0.70 | -1.67 |
| Great Stour at Horton | 40011 | 345 | 770 | 610 | 22 | 0.69 | 0.79 | 0.43 |
| Greet at Southwell | 28072 | 59 | 656 | 557 | 19 | 0.71 | 0.44 | -0.93 |
| Greta at Rutherford Bridge | 25006 | 87 | 1135 | 509 | 35 | 0.21 | 2.26 | 0.7 |
| Gwash South Arm at Manton | 31025 | 26 | 729 | 573 | 28 | 0.27 | 0.60 | 0.52 |
| Gwili at Glangwili | 60006 | 134 | 1641 | 576 | 52 | 0.47 | 3.31 | 0.81 |
| Halladale at Halladale | 96001 | 209 | 1136 | 470 | 33 | 0.27 | 2.09 | 0.64 |
| Harpers Brook at Old Mill Bridge | 32003 | 74 | 657 | 590 | 16 | 0.48 | 0.49 | 0.7 |
| Harwood Beck at Harwood | 25012 | 26 | 1474 | 486 | 91 | 0.24 | 3.30 | 0.59 |
| Hayle at St Erth | 49002 | 48 | 1104 | 589 | 27 | 0.83 | 1.78 | -0.96 |
| Hebden Beck at Hebden | 27032 | 25 | 1544 | 506 | 59 | 0.43 | 0.62 | -34.4 |
| Hiz at Arlesey | 33033 | 114 | 608 | 604 | 17 | 0.85 | 0.51 | -1.31 |
| Hodder at Hodder Place | 71008 | 266 | 1620 | 530 | 61 | 0.31 | 2.81 | 0.75 |
| Idle at Mattersey | 28015 | 538 | 664 | 566 | 14 | 0.78 | 0.38 | -6.13 |
| Ingrebourne at Gaynes Park | 37018 | 47 | 609 | 617 | 18 | 0.49 | 0.59 | 0.68 |
| Inver at Little Assynt | 95001 | 145 | 2165 | 450 | 114 | 0.65 | 5.00 | 0.84 |
| Irvine at Newmilns | 83010 | 73 | 1360 | 507 | 41 | 0.28 | 2.85 | 0.64 |
| Irvine at Shewalton | 83005 | 369 | 1275 | 518 | 26 | 0.27 | 2.33 | 0.69 |
| Irwell at Adelphi Weir | 69002 | 561 | 1281 | 557 | 45 | 0.49 | 2.70 | 0.73 |
| Ise Brook at Harrowden Old Mill | 32004 | 197 | 665 | 588 | 16 | 0.54 | 0.58 | 0.76 |
| Ithon at Disserth | 55016 | 363 | 1194 | 548 | 49 | 0.38 | 1.92 | 0.88 |
| Ivel at Blunham | 33022 | 545 | 593 | 600 | 14 | 0.73 | 0.47 | 0.47 |
| Jed Water at Jedburgh | 21024 | 143 | 958 | 496 | 51 | 0.41 | 1.43 | 0.71 |
| Kennet at Theale | 39016 | 1055 | 803 | 600 | 22 | 0.88 | 0.80 | -2.91 |
| Kent at Sedgwick | 73005 | 210 | 1749 | 515 | 85 | 0.41 | 3.79 | 0.83 |
| Kielder Burn at Kielder | 23011 | 57 | 1347 | 483 | 89 | 0.33 | 3.00 | 0.57 |
| Kinnel Water at Redhall | 78004 | 77 | 1449 | 481 | 64 | 0.27 | 3.16 | 0.71 |
| Kirtle Water at Mossknowe | 77004 | 76 | 1213 | 510 | 33 | 0.29 | 2.04 | 0.73 |
| Kym at Meagre Farm | 33012 | 134 | 613 | 594 | 14 | 0.26 | 0.39 | 0.66 |
| Lambourn at Shaw | 39019 | 240 | 778 | 601 | 24 | 0.97 | 0.62 | -11.18 |

*Table A.1: Catchment characteristics and NSE for standard run.*

| Catchment | Station | Area | Avg. Ann. Rainfall (mm) | Avg. Ann PET (mm) | DPSBAR | BFI | mean flow (mm/day) | NSE |
|---|---|---|---|---|---|---|---|---|
| Lavant at Graylingwell | 41023 | 89 | 978 | 604 | 45 | 0.82 | 0.29 | -2.08 |
| Laver at Ripon | 27059 | 82 | 928 | 554 | 44 | 0.43 | 1.13 | 0.7 |
| Leadon at Wedderburn Bridge | 54017 | 291 | 715 | 605 | 20 | 0.49 | 0.61 | 0.73 |
| Lee at Feildes Weir | 38001 | 1056 | 646 | 600 | 16 | 0.59 | 0.36 | -0.52 |
| Leet Water at Coldstream | 21023 | 114 | 679 | 525 | 18 | 0.33 | 0.74 | 0.78 |
| Leri at Dolybont | 64006 | 49 | 1629 | 545 | 74 | 0.49 | 2.36 | 0.84 |
| Leven at Leven Bridge | 25005 | 202 | 758 | 545 | 33 | 0.42 | 0.81 | 0.86 |
| Leven at Newby Bridge FMS | 73010 | 257 | 2287 | 518 | 114 | 0.49 | 4.71 | 0.92 |
| Liddel Water at Rowanburnfoot | 77003 | 320 | 1368 | 485 | 63 | 0.31 | 2.83 | 0.69 |
| Little Eachaig at Dalinlongart | 86001 | 33 | 2468 | 501 | 168 | 0.21 | 4.69 | 0.67 |
| Little Ouse at Abbey Heath | 33034 | 723 | 635 | 599 | 8 | 0.8 | 0.44 | -0.68 |
| Llynfi at Three Cocks | 55025 | 137 | 1134 | 555 | 55 | 0.57 | 1.42 | 0.36 |
| Lod at Halfway Bridge | 41022 | 54 | 868 | 604 | 33 | 0.35 | 0.95 | 0.74 |
| Lossie at Sheriffmills | 7003 | 221 | 886 | 540 | 43 | 0.53 | 1.07 | 0.8 |
| Lossie at Torwinny | 7006 | 18 | 1038 | 531 | 37 | 0.45 | 1.82 | 0.56 |
| Loxwood Stream at Drungewick | 41025 | 94 | 817 | 600 | 21 | 0.22 | 1.04 | 0.54 |
| Luce at Airyhemming | 81003 | 175 | 1522 | 523 | 40 | 0.22 | 3.03 | 0.74 |
| Lud at Louth | 29003 | 56 | 706 | 563 | 23 | 0.9 | 0.70 | -0.78 |
| Lugg at Butts Bridge | 55021 | 377 | 942 | 571 | 53 | 0.66 | 1.33 | 0.46 |
| Lugg at Byton | 55014 | 206 | 1061 | 562 | 71 | 0.65 | 1.63 | 0.21 |
| Luggie Water at Condorrat | 84016 | 33 | 1168 | 528 | 23 | 0.41 | 2.29 | 0.53 |
| Lunan Water at Kirkton Mill | 13005 | 130 | 723 | 519 | 25 | 0.52 | 1.12 | 0.71 |
| Lune at Caton | 72004 | 983 | 1549 | 512 | 72 | 0.32 | 3.16 | 0.78 |
| Lune at Killington New Bridge | 72005 | 223 | 1654 | 510 | 72 | 0.33 | 3.86 | 0.8 |
| Luss Water at Luss | 85004 | 35 | 2282 | 499 | 179 | 0.28 | 6.51 | 0.64 |
| Lymn at Partney Mill | 30004 | 63 | 730 | 562 | 18 | 0.65 | 0.68 | 0.5 |
| Lyne Water at Lyne Station | 21018 | 175 | 1030 | 500 | 51 | 0.57 | 1.54 | 0.81 |
| Lyon at Comrie Bridge | 15011 | 400 | 1976 | 475 | 161 | 0.44 | 2.70 | -0.06 |
| Manifold at Ilam | 28031 | 151 | 1108 | 530 | 52 | 0.53 | 2.04 | 0.78 |
| Meig at Glenmeannie | 4005 | 127 | 1913 | 473 | 179 | 0.25 | 4.71 | 0.53 |
| Mellte at Pontneddfechan | 58006 | 65 | 2056 | 541 | 76 | 0.35 | 4.32 | 0.87 |
| Meon at Mislingford | 42006 | 75 | 917 | 604 | 30 | 0.93 | 1.12 | -1.08 |
| Mersey at Ashton Weir | 69007 | 667 | 1152 | 565 | 56 | 0.54 | 1.65 | 0.56 |
| Midford Brook at Midford | 53005 | 147 | 970 | 622 | 31 | 0.62 | 1.32 | 0.86 |
| Mimram at Panshanger Park | 38003 | 135 | 659 | 605 | 20 | 0.93 | 0.34 | -22.34 |
| Mires Beck at North Cave | 26008 | 42 | 707 | 557 | 26 | 0.87 | 0.48 | -2.71 |
| Mole at Kinnersley Manor | 39069 | 150 | 806 | 603 | 12 | 0.39 | 1.29 | 0.78 |
| Mole at Woodleigh | 50006 | 329 | 1369 | 587 | 46 | 0.47 | 2.31 | 0.74 |
| Monnow at Grosmont | 55029 | 355 | 1021 | 581 | 64 | 0.51 | 1.44 | 0.81 |
| Motray Water at St Michaels | 14005 | 58 | 697 | 553 | 39 | 0.57 | 0.84 | 0.64 |
| Nadder at Wilton | 43006 | 217 | 921 | 607 | 30 | 0.81 | 1.14 | 0.28 |
| Naver at Apigill | 96002 | 475 | 1415 | 464 | 68 | 0.42 | 2.89 | 0.8 |
| Nene Kislingbury at Dodford | 32008 | 108 | 670 | 585 | 16 | 0.57 | 0.50 | 0.8 |
| Ness at Ness-side | 6007 | 1857 | 1866 | 472 | 104 | 0.6 | 4.19 | 0.57 |
| Nevis at Claggan | 90003 | 71 | 2880 | 437 | 228 | 0.26 | 7.99 | 0.47 |
| Nith at Drumlanrig | 79006 | 481 | 1589 | 492 | 64 | 0.34 | 3.08 | 0.84 |

*Table A.1: Catchment characteristics and NSE for standard run.*

| Catchment | Station | Area | Avg. Ann. Rainfall (mm) | Avg. Ann PET (mm) | DPSBAR | BFI | mean flow (mm/day) | NSE |
|---|---|---|---|---|---|---|---|---|
| Nith at Friars Carse | 79002 | 805 | 1542 | 490 | 68 | 0.39 | 2.98 | 0.84 |
| Nith at Hall Bridge | 79003 | 160 | 1622 | 502 | 58 | 0.27 | 3.11 | 0.83 |
| North Esk at Logie Mill | 13007 | 755 | 1161 | 490 | 76 | 0.51 | 2.20 | 0.73 |
| Nunningham Stream at Tilley Bridge | 41001 | 19 | 841 | 602 | 25 | 0.33 | 0.89 | 0.67 |
| Ock at Abingdon | 39081 | 234 | 638 | 613 | 12 | 0.64 | 0.57 | 0.6 |
| Ogmore at Bridgend | 58001 | 156 | 1871 | 572 | 85 | 0.48 | 3.67 | 0.8 |
| Otter at Dotton | 45005 | 196 | 1016 | 595 | 38 | 0.53 | 1.39 | 0.64 |
| Ouse at Gold Bridge | 41005 | 180 | 856 | 595 | 24 | 0.49 | 1.06 | 0.64 |
| Oykel at Easter Turnaig | 3003 | 336 | 1862 | 468 | 84 | 0.22 | 4.21 | 0.53 |
| Perry at Yeaton | 54020 | 181 | 789 | 580 | 15 | 0.66 | 0.76 | -0.12 |
| Petteril at Harraby Green | 76010 | 164 | 962 | 540 | 26 | 0.46 | 1.13 | 0.69 |
| Piddle at Baggs Mill | 44002 | 186 | 997 | 610 | 25 | 0.89 | 1.12 | -2.94 |
| Pincey Brook at Sheering Hall | 38026 | 56 | 618 | 596 | 9 | 0.37 | 0.48 | 0.79 |
| Pointon Lode at Pointon | 30014 | 11 | 608 | 586 | 18 | 0.49 | 0.53 | 0.57 |
| Poulter at Cuckney | 28044 | 37 | 679 | 562 | 29 | 0.92 | 0.72 | -3.76 |
| Quaggy at Manor House Gardens | 39095 | 34 | 642 | 657 | 17 | 0.46 | 0.37 | -6.72 |
| Ravensbourne at Catford Hill | 39056 | 125 | 713 | 634 | 23 | 0.54 | 0.28 | -16.4 |
| Rea at Calthorpe Park | 28039 | 74 | 810 | 632 | 18 | 0.46 | 0.91 | -0.21 |
| Rede at Rede Bridge | 23008 | 348 | 1012 | 496 | 53 | 0.32 | 1.51 | 0.7 |
| Rhee at Burnt Mill | 33021 | 311 | 571 | 597 | 12 | 0.74 | 0.32 | 0.56 |
| Rhymney at Llanedeyrn | 57008 | 187 | 1514 | 580 | 70 | 0.47 | 2.61 | 0.76 |
| Roden at Rodington | 54016 | 266 | 719 | 577 | 9 | 0.62 | 0.62 | 0.12 |
| Roding at Redbridge | 37001 | 304 | 624 | 603 | 14 | 0.39 | 0.51 | 0.89 |
| Rother at Iping Mill | 41011 | 159 | 936 | 605 | 25 | 0.67 | 1.25 | 0.6 |
| Rother at Princes Marsh | 41027 | 40 | 942 | 601 | 29 | 0.6 | 1.10 | 0.58 |
| Rother at Woodhouse Mill | 27025 | 364 | 775 | 571 | 32 | 0.54 | 0.97 | 0.72 |
| Rothley Brook at Rothley | 28056 | 92 | 701 | 579 | 20 | 0.46 | 0.69 | 0.8 |
| Ruchill Water at Cultybraggan | 16003 | 99 | 1957 | 493 | 112 | 0.29 | 4.52 | 0.75 |
| Sapiston at Rectory Bridge | 33013 | 210 | 610 | 602 | 9 | 0.65 | 0.28 | -0.3 |
| Scar Water at Capenoch | 79004 | 147 | 1575 | 489 | 92 | 0.31 | 3.37 | 0.78 |
| Severn at Plynlimon flume | 54022 | 9 | 2656 | 522 | 113 | 0.34 | 5.15 | 0.72 |
| Sherston Avon at Fosseway | 53023 | 76 | 862 | 593 | 17 | 0.65 | 1.11 | 0.71 |
| Silk Stream at Colindeep Lane | 39049 | 32 | 715 | 656 | 21 | 0.33 | 0.67 | -0.12 |
| Skerne at Preston le Skerne | 25020 | 149 | 686 | 545 | 19 | 0.4 | 0.50 | 0.52 |
| Snaizeholme Beck at Low Houses | 27047 | 11 | 1754 | 495 | 107 | 0.18 | 4.38 | 0.71 |
| Soar at Littlethorpe | 28082 | 183 | 661 | 577 | 11 | 0.49 | 0.64 | 0.8 |
| South Esk at Brechin | 13008 | 496 | 1218 | 488 | 91 | 0.54 | 2.18 | 0.74 |
| South Tyne at Featherstone | 23006 | 325 | 1358 | 491 | 74 | 0.32 | 2.82 | 0.68 |
| South Tyne at Haydon Bridge | 23004 | 751 | 1153 | 501 | 61 | 0.34 | 2.11 | 0.7 |
| Sow at Great Bridgford | 28052 | 149 | 780 | 549 | 18 | 0.65 | 0.67 | 0.35 |
| Sprint at Sprint Mill | 73009 | 38 | 2073 | 509 | 120 | 0.32 | 4.49 | 0.81 |
| Stansted Sp at Mountfitchet | 38016 | 49 | 643 | 587 | 16 | 0.98 | 0.10 | -442.49 |
| Stour at Hammoon | 43009 | 524 | 886 | 611 | 19 | 0.31 | 1.26 | 0.75 |
| Stour at Langham | 36006 | 579 | 604 | 595 | 15 | 0.52 | 0.45 | 0.66 |
| Stour at Throop | 43007 | 1066 | 896 | 614 | 21 | 0.65 | 1.11 | 0.73 |
| Strathmore at Allnabad | 96004 | 108 | 2253 | 444 | 115 | 0.2 | 5.87 | 0.59 |

*Table A.1: Catchment characteristics and NSE for standard run.*

| Catchment | Station | Area | Avg. Ann. Rainfall (mm) | Avg. Ann PET (mm) | DPSBAR | BFI | mean flow (mm/day) | NSE |
|---|---|---|---|---|---|---|---|---|
| Stringside at Whitebridge | 33029 | 100 | 652 | 593 | 8 | 0.84 | 0.44 | -4.82 |
| Swale at Crakehill | 27071 | 1359 | 873 | 546 | 37 | 0.47 | 1.32 | 0.88 |
| Sydling Water at Sydling St Nicholas | 44006 | 13 | 1094 | 602 | 54 | 0.88 | 1.26 | -7.27 |
| Taf at Clog-y-Fran | 60003 | 218 | 1475 | 573 | 40 | 0.55 | 2.99 | 0.65 |
| Taff at Pontypridd | 57005 | 464 | 1960 | 560 | 88 | 0.45 | 3.82 | 0.88 |
| Tame at Lea Marston Lakes | 28080 | 801 | 732 | 619 | 13 | 0.69 | 1.49 | 0.06 |
| Tanat at Llanyblodwel | 54038 | 239 | 1379 | 544 | 103 | 0.48 | 2.39 | 0.83 |
| Tas at Shotesham | 34002 | 152 | 634 | 595 | 9 | 0.59 | 0.43 | 0.67 |
| Tawe at Ynystanglws | 59001 | 235 | 1975 | 566 | 78 | 0.36 | 4.51 | 0.81 |
| Tees at Broken Scar | 25001 | 825 | 1152 | 509 | 49 | 0.34 | 1.77 | 0.58 |
| Teise at Stone Bridge | 40009 | 146 | 841 | 602 | 27 | 0.46 | 0.80 | 0.52 |
| Teme at Knightsford Bridge | 54029 | 1502 | 860 | 573 | 46 | 0.55 | 1.01 | 0.8 |
| Teme at Tenbury | 54008 | 1138 | 888 | 568 | 49 | 0.55 | 1.09 | 0.74 |
| Test at Broadlands | 42004 | 1048 | 830 | 607 | 21 | 0.94 | 0.91 | -7.7 |
| Thrushel at Tinhay | 47008 | 111 | 1212 | 591 | 31 | 0.43 | 1.79 | 0.57 |
| Thurso at Halkirk | 97002 | 417 | 1070 | 464 | 21 | 0.45 | 1.88 | 0.76 |
| Tiddy at Tideford | 47009 | 39 | 1287 | 593 | 42 | 0.6 | 2.06 | 0.21 |
| Tillingbourne at Shalford | 39029 | 58 | 823 | 598 | 33 | 0.89 | 0.79 | -13.19 |
| Tima Water at Deephope | 21026 | 30 | 1598 | 458 | 68 | 0.26 | 3.98 | 0.71 |
| Tone at Bishops Hull | 52005 | 205 | 1029 | 608 | 38 | 0.6 | 1.26 | 0.73 |
| Torne at Auckley | 28050 | 137 | 622 | 595 | 12 | 0.7 | 0.56 | -0.21 |
| Tove at Cappenham Bridge | 33018 | 135 | 710 | 585 | 16 | 0.54 | 0.67 | 0.8 |
| Trent at Stoke on Trent | 28040 | 52 | 886 | 550 | 33 | 0.44 | 1.05 | 0.76 |
| Turkey Brook at Albany Park | 38021 | 48 | 682 | 620 | 22 | 0.22 | 0.36 | 0.37 |
| Tyne at Bywell | 23001 | 2183 | 1057 | 503 | 52 | 0.38 | 1.83 | 0.71 |
| Tywi at Nantgaredig | 60010 | 1097 | 1640 | 559 | 64 | 0.46 | 3.12 | 0.85 |
| Uck at Isfield | 41006 | 89 | 848 | 598 | 24 | 0.42 | 1.08 | 0.77 |
| Ure at Kilgram Bridge | 27034 | 511 | 1431 | 514 | 84 | 0.32 | 2.72 | 0.81 |
| Ure at Westwick Lock | 27007 | 922 | 1179 | 533 | 61 | 0.39 | 2.01 | 0.85 |
| Urr at Dalbeattie | 80001 | 198 | 1397 | 514 | 38 | 0.36 | 2.60 | 0.84 |
| Usk at Chain Bridge | 56001 | 919 | 1462 | 556 | 82 | 0.5 | 2.62 | 0.85 |
| Ver at Hansteads | 39014 | 132 | 724 | 606 | 18 | 0.88 | 0.28 | -31.57 |
| Wallington at North Fareham | 42001 | 109 | 876 | 623 | 18 | 0.41 | 0.49 | -1.81 |
| Wansbeck at Mitford | 22007 | 289 | 833 | 513 | 26 | 0.37 | 0.96 | 0.79 |
| Waveney at Needham Mill | 34006 | 373 | 615 | 597 | 8 | 0.46 | 0.42 | 0.72 |
| Wear at Sunderland Bridge | 24001 | 667 | 961 | 518 | 58 | 0.42 | 1.44 | 0.77 |
| Weaver at Ashbrook | 68001 | 613 | 743 | 566 | 12 | 0.54 | 0.79 | 0.77 |
| Weaver at Audlem | 68005 | 201 | 732 | 568 | 12 | 0.54 | 0.67 | 0.77 |
| Wellow Brook at Wellow | 53009 | 76 | 1017 | 624 | 31 | 0.62 | 1.46 | 0.78 |
| Wensum at Fakenham | 34011 | 167 | 697 | 589 | 9 | 0.82 | 0.45 | -6.21 |
| West Avon at Upavon | 43017 | 78 | 783 | 598 | 19 | 0.71 | 0.75 | 0.42 |
| Wey at Broadwey | 44009 | 11 | 986 | 617 | 75 | 0.95 | 2.53 | -0.05 |
| Wey at Tilford | 39011 | 394 | 881 | 604 | 24 | 0.73 | 0.71 | -4.33 |
| Windrush at Newbridge | 39006 | 365 | 793 | 594 | 27 | 0.86 | 0.78 | -0.15 |
| Wistaston Brook at Marshfield Bridge | 68004 | 97 | 741 | 555 | 13 | 0.6 | 0.82 | 0.51 |
| Witham at Claypole Mill | 30001 | 301 | 631 | 578 | 15 | 0.69 | 0.53 | 0.64 |

*Table A.1: Catchment characteristics and NSE for standard run.*

| Catchment | Station | Area | Avg. Ann. Rainfall (mm) | Avg. Ann PET (mm) | DPSBAR | BFI | mean flow (mm/day) | NSE |
|---|---|---|---|---|---|---|---|---|
| Worfe at Burcote | 54024 | 258 | 713 | 580 | 16 | 0.7 | 0.38 | -2.83 |
| Wreake at Syston Mill | 28024 | 413 | 663 | 563 | 17 | 0.41 | 0.59 | 0.78 |
| Wye at Cefn Brwyn | 55008 | 12 | 2516 | 526 | 90 | 0.31 | 5.09 | 0.7 |
| Wye at Hedsor | 39023 | 135 | 770 | 606 | 39 | 0.94 | 0.64 | -15.61 |
| Wye at Redbrook | 55023 | 4030 | 1075 | 570 | 50 | 0.53 | 1.59 | 0.93 |
| Yscir at Pontaryscir | 56013 | 61 | 1428 | 547 | 54 | 0.46 | 2.74 | 0.85 |

# Appendix B

# Supporting evidence for Chapter 6

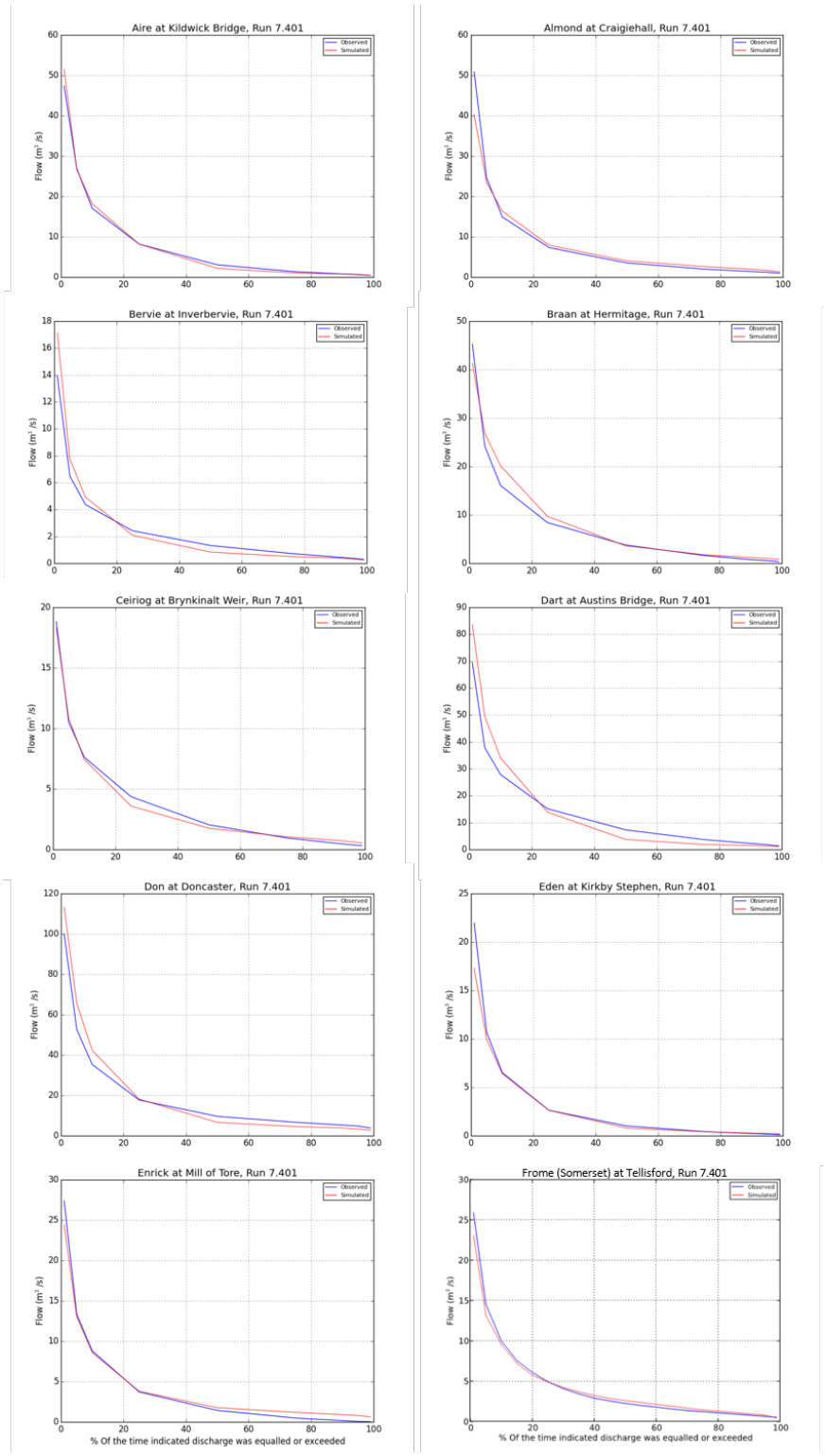



*Figure B.1: Flow duration curves of observed and simulated flows for catchments used in the climate change impact study in Chapter 6.*

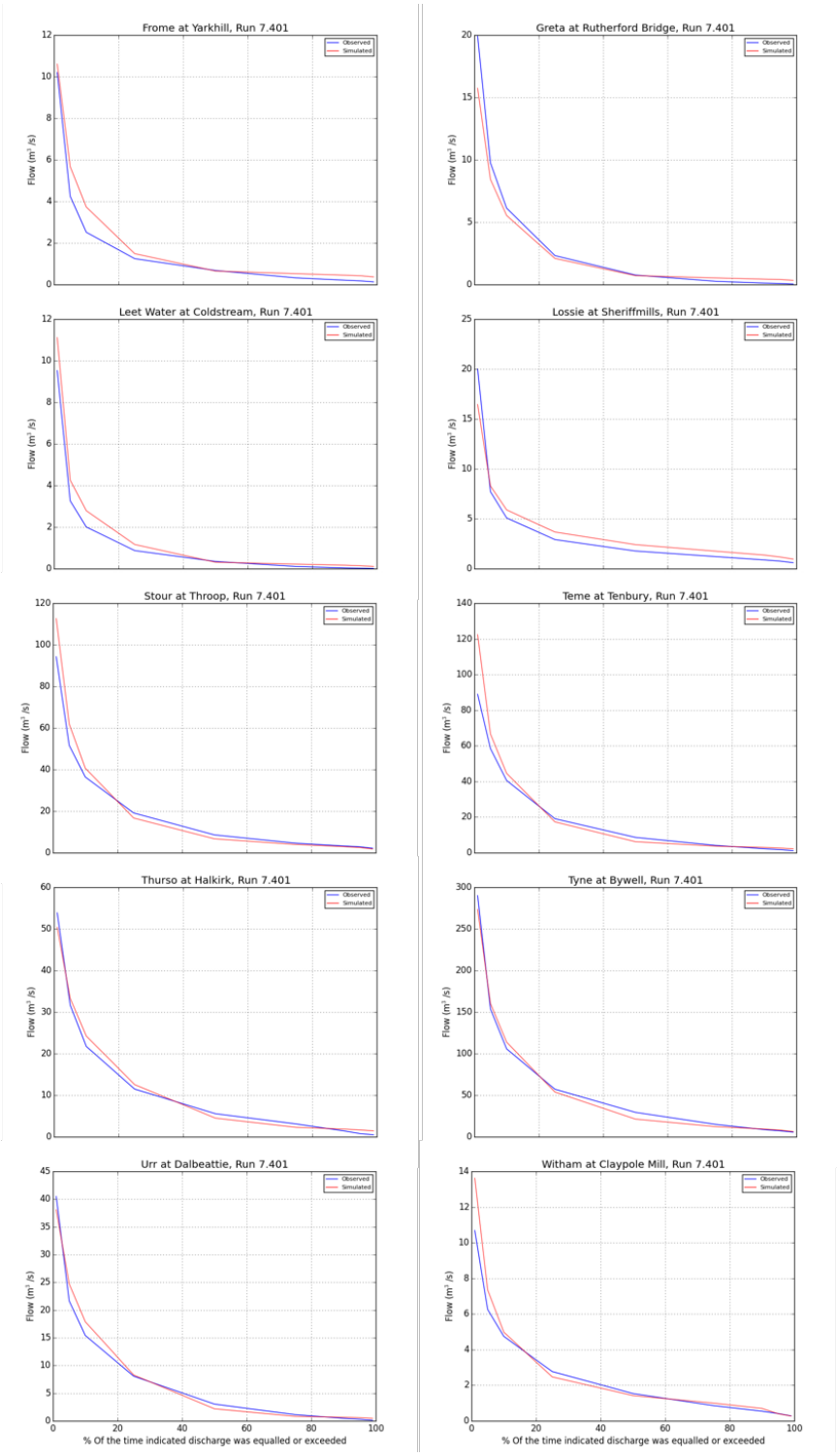



*Figure B.2: Flow duration curves of observed and simulated flows for catchments used in the climate change impact study in Chapter 6.*

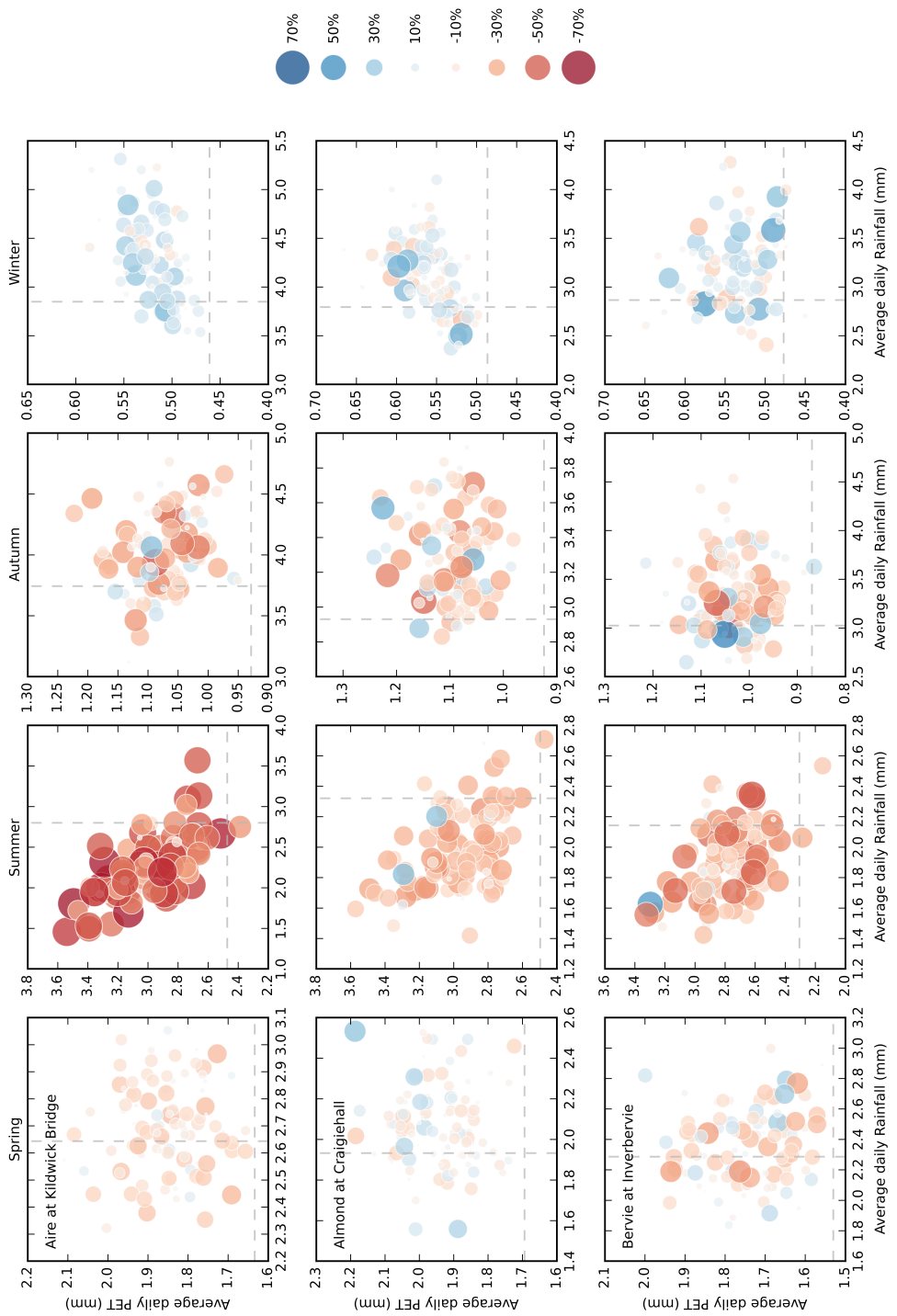



*Figure B.3: Change in mean flows for each future run from the mean control flow, for each catchment and season, plotted against average daily rainfall and average daily PET. The size and colour of the dot indicates the direction and magnitude of the percent change in mean flow. The grey dashed lines indicate the mean rainfall and PET of the control runs.*

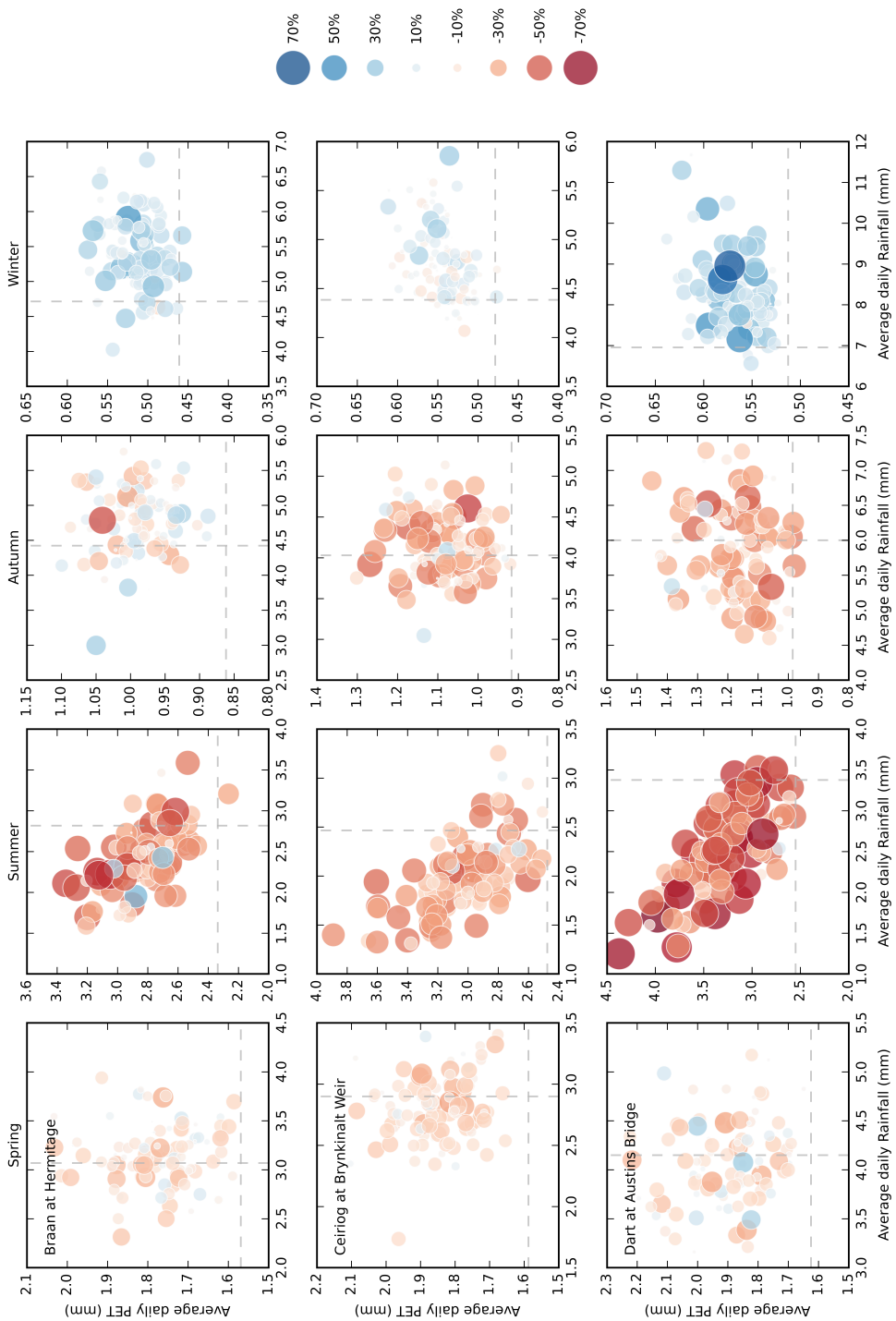



*Figure B.4: Change in mean flows for each future run from the mean control flow, for each catchment and season, plotted against average daily rainfall and average daily PET. The size and colour of the dot indicates the direction and magnitude of the percent change in mean flow. The grey dashed lines indicate the mean rainfall and PET of the control runs.*

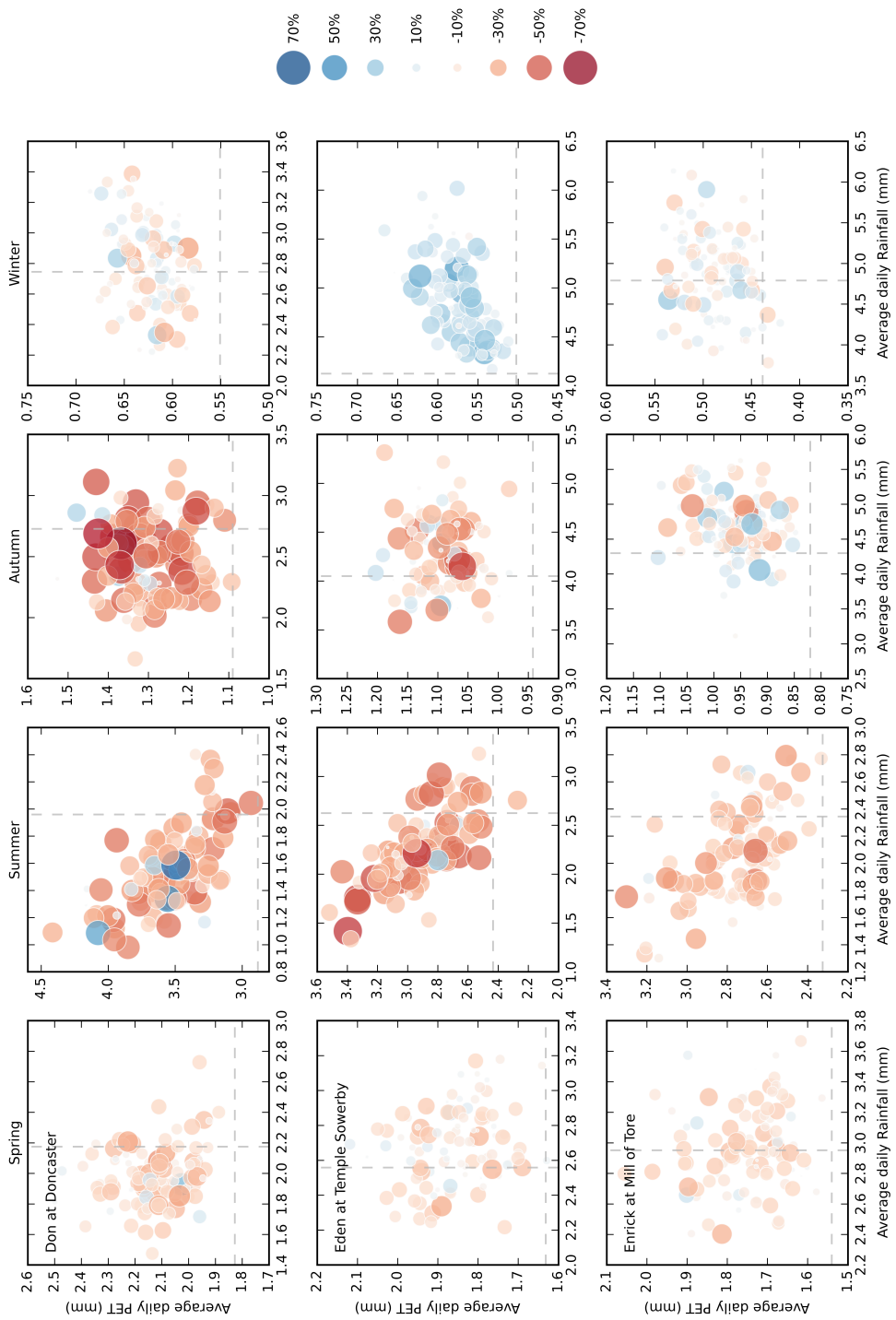



*Figure B.5: Change in mean flows for each future run from the mean control flow, for each catchment and season, plotted against average daily rainfall and average daily PET. The size and colour of the dot indicates the direction and magnitude of the percent change in mean flow. The grey dashed lines indicate the mean rainfall and PET of the control runs.*

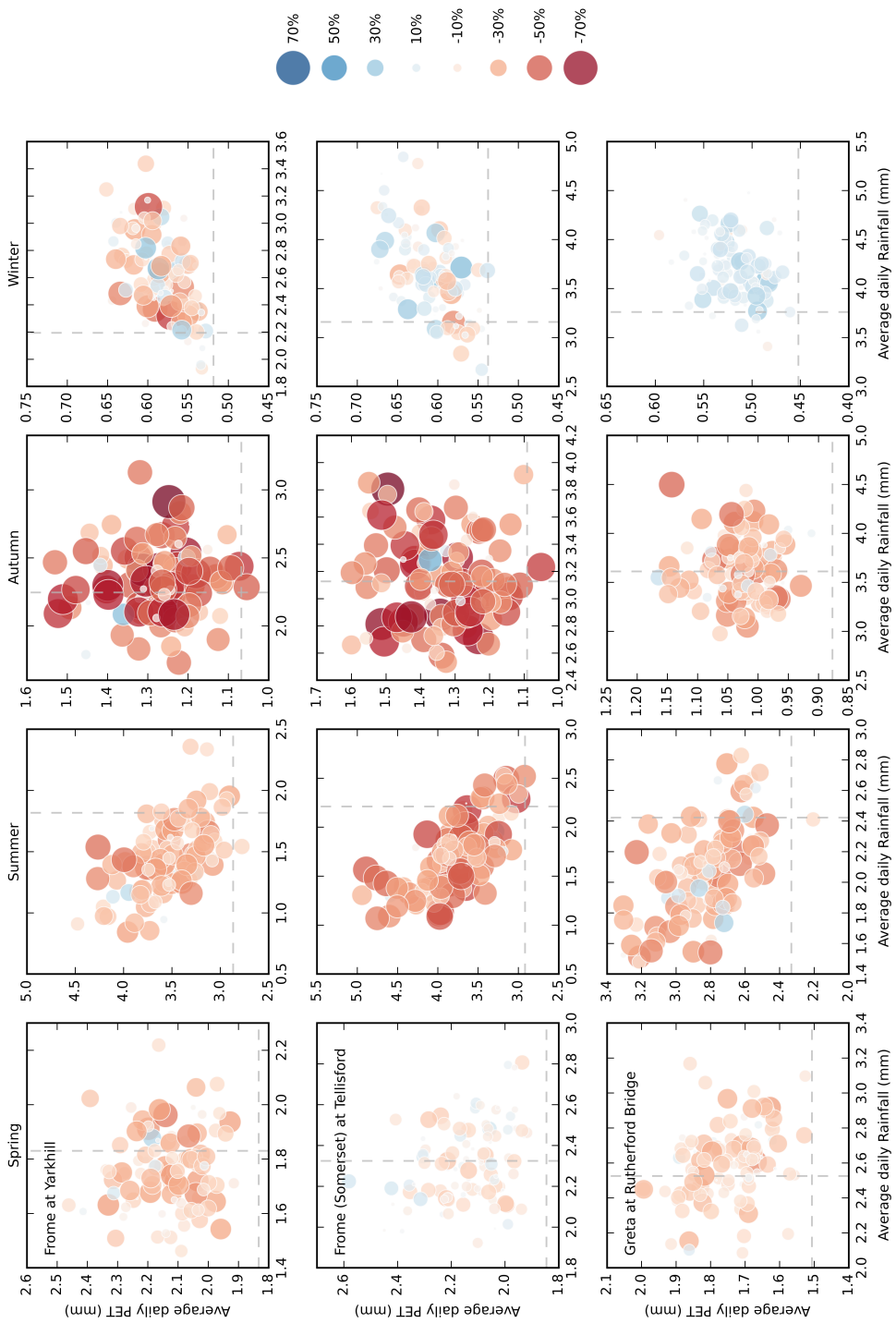



*Figure B.6: Change in mean flows for each future run from the mean control flow, for each catchment and season, plotted against average daily rainfall and average daily PET. The size and colour of the dot indicates the direction and magnitude of the percent change in mean flow. The grey dashed lines indicate the mean rainfall and PET of the control runs.*

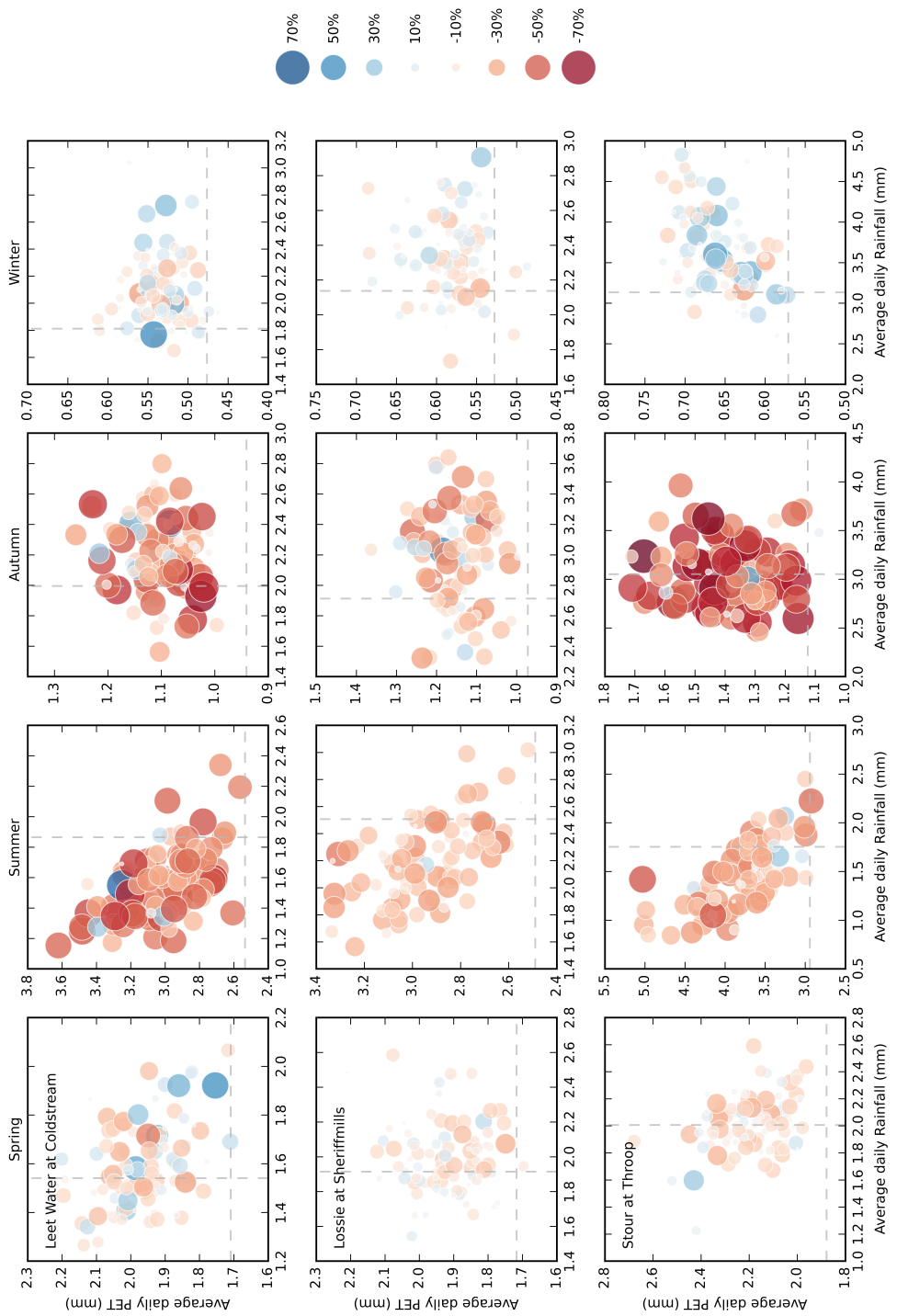



*Figure B.7: Change in mean flows for each future run from the mean control flow, for each catchment and season, plotted against average daily rainfall and average daily PET. The size and colour of the dot indicates the direction and magnitude of the percent change in mean flow. The grey dashed lines indicate the mean rainfall and PET of the control runs.*

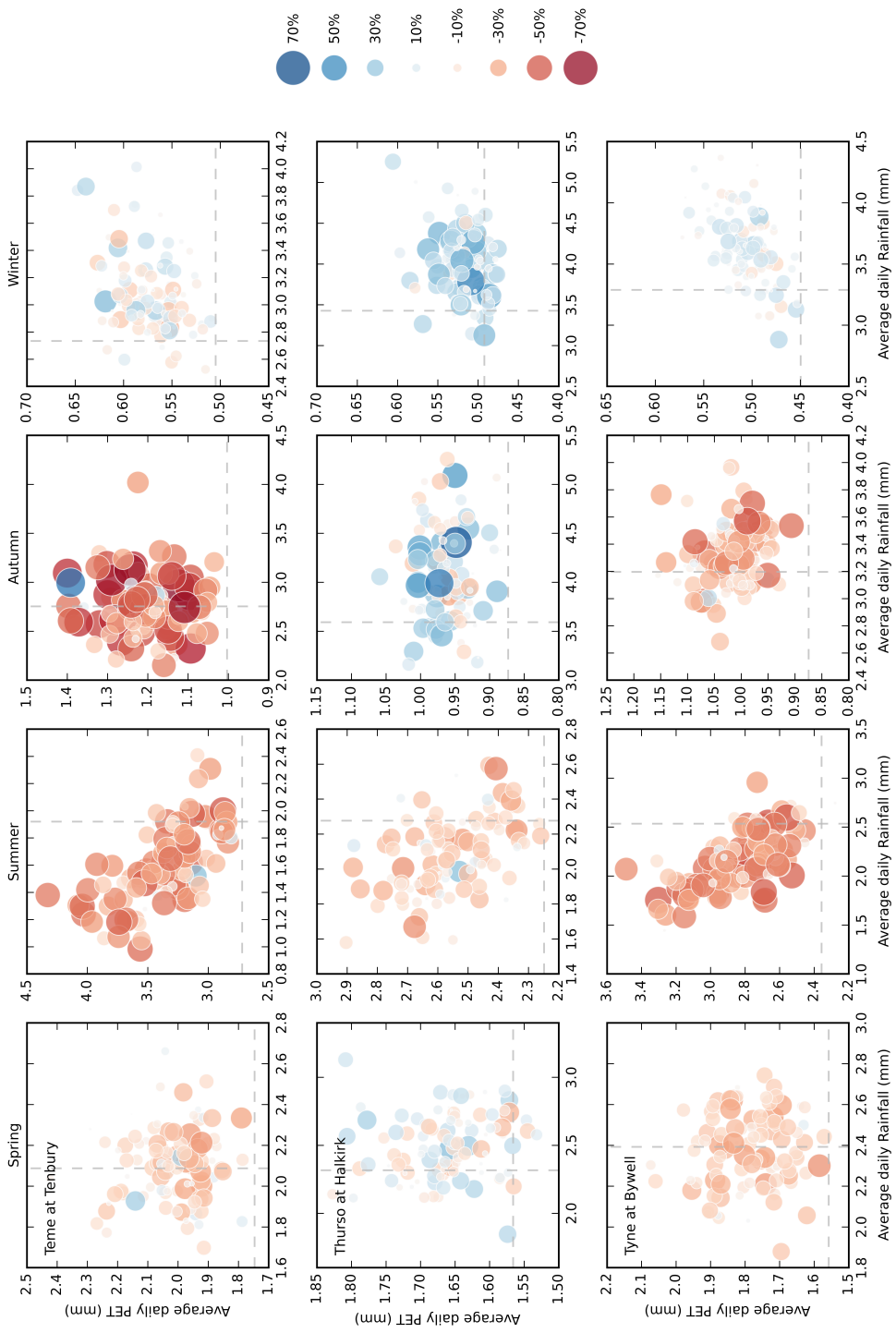



*Figure B.8: Change in mean flows for each future run from the mean control flow, for each catchment and season, plotted against average daily rainfall and average daily PET. The size and colour of the dot indicates the direction and magnitude of the percent change in mean flow. The grey dashed lines indicate the mean rainfall and PET of the control runs.*

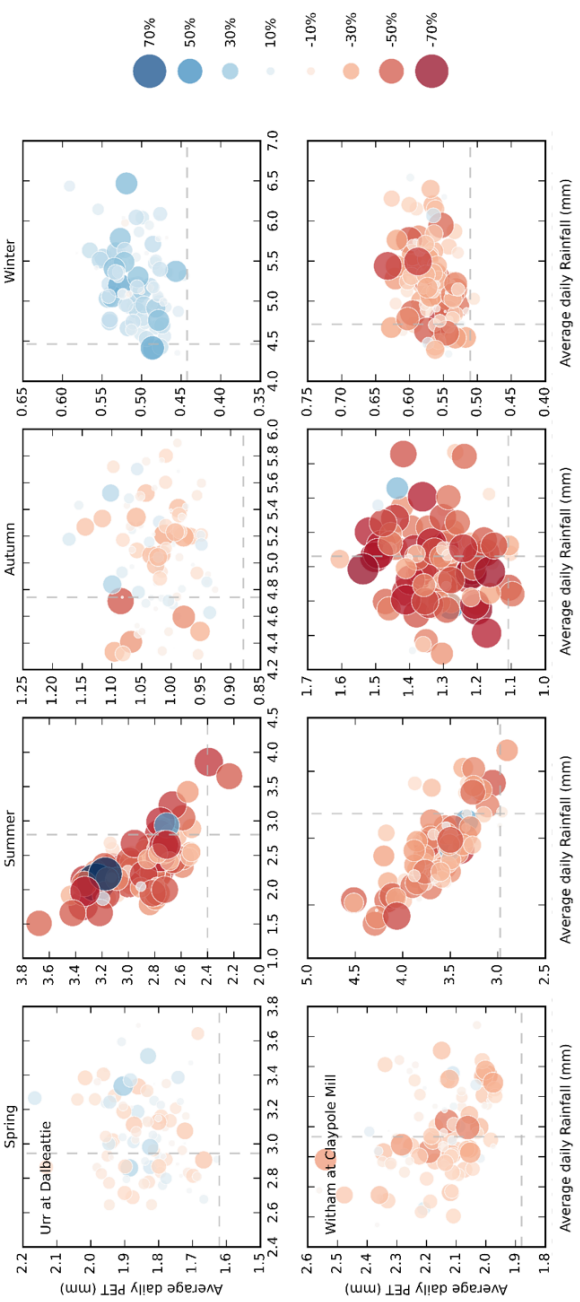



*Figure B.9: Change in mean flows for each future run from the mean control flow, for each catchment and season, plotted against average daily rainfall and average daily PET. The size and colour of the dot indicates the direction and magnitude of the percent change in mean flow. The grey dashed lines indicate the mean rainfall and PET of the control runs.*