University of Dundee

Influence of alternative representations of land use and geology on distributed hydrological modelling results: Eddleston, Scotland

Ruman, Stanislav; Ball, Tom; Black, Andrew; Thompson, Julian R.

Published in:
Hydrological Sciences Journal

DOI:
[10.1080/02626667.2020.1862851](https://doi.org/10.1080/02626667.2020.1862851)

Publication date:
2021

Document Version
Peer reviewed version

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):

Ruman, S., Ball, T., Black, A., & Thompson, J. R. (2021). Influence of alternative representations of land use and geology on distributed hydrological modelling results: Eddleston, Scotland. *Hydrological Sciences Journal*, 66(3), 488-502. <https://doi.org/10.1080/02626667.2020.1862851>

General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Publisher: Taylor & Francis & IAHS

Journal: *Hydrological Sciences Journal*

DOI: 10.1080/02626667.2020.1862851

Influence of alternative representations of land use and geology on distributed hydrological modelling results: Eddleston, Scotland

Stanislav Ruman¹, Tom Ball², Andrew R. Black³ and Julian R. Thompson⁴

¹ *Department of Physical Geography and Geoecology, Faculty of Sciences, University of Ostrava, Chittussiho 10, 710 00, Ostrava, Czech Republic, <https://orcid.org/0000-0002-2607-2454>, stanislav.ruman@osu.cz*

² *Department of Archaeology, Anthropology and Geography, Faculty of Humanities and Social Sciences, University of Winchester, Sparkford Road, Winchester, SO22 4NR, UK, <https://orcid.org/0000-0002-7324-4300>*

³ *Geography, School of Social Sciences, University of Dundee, Nethergate, Dundee DD1 4HN, Scotland, UK, <https://orcid.org/0000-0001-9292-1146>*

⁴ *Wetland Research Unit, UCL Department of Geography, University College London, Gower Street, London WC1E 6BT, UK, <https://orcid.org/0000-0002-8927-6462>*

Abstract

A distributed hydrological model was applied to a 69km² experimental catchment, Eddleston Water, Scotland, UK. The impact on model outputs of applying progressively simpler representations of spatial variability in land use and superficial geology was assessed. Alternative representations of the spatial distribution of superficial geology and land use

produced differences in model outputs. These differences were generally small with the exception of the maximum absolute error ($E_{\max}$). Inter-model differences were most sensitive to the largest precipitation events. Although variations in superficial geology dominated over those for land use, exceptions were seen in two sub-catchments. These were connected with particularly large variations in land use and/or the small spatial extent of superficial geology. Lower resolution spatial data produced superior model performance in the majority of sub-catchments. This has implications for modelling other catchments especially in situations where the high resolution data employed herein are not available.

Keywords: Distributed hydrological model, Land Use, Superficial Geology, spatial datasets

1. Introduction

Physically based, distributed modelling can improve understanding of the processes underlying catchment hydrological functioning and the effects of alterations that may include land-use or climate change as well as management interventions, water abstraction and in-stream modifications (Uhlenbrook and Sieber 2005; Fatichi et al. 2016; Yira et al. 2016; Thompson et al. 2017). Typically, such models employ the highest resolution spatial data that are available based on the assumption that they provide the most accurate representation of catchment characteristics. However, use of the finest scale data does not necessarily provide the best agreement with observations (Vázquez and Hampel 2014; Alvarenga et al. 2016; Camargos et al. 2018). The reasons for this lie in multiple sources of uncertainty (e.g. over-parametrization, variable calibration procedures and model structure) that have been described as these models have developed (Freeze and Harlan 1969; Beven 1993; Walker et al. 2003; Fatichi et al. 2016).

Although some studies have concluded that variations in the spatial distribution of rainfall have a dominant effect on model outputs (Bell and Moore 2000; Boyle et al. 2001; Andréassian et al. 2004; Masih et al. 2011; Lobligeios et al. 2014; Terink et al. 2018), several studies have attempted to quantify the impacts of using other alternative spatial datasets including digital elevation models (Yen et al. 2015), soils (Camargos et al. 2018) and, most notably, land use (Li and Wang 2009; Huisman et al. 2009; Bormann and Elfert 2010; Bingqing et al. 2015). The focus on land use reflects recognition that land cover change can significantly affect catchment hydrological processes. Expansion of urban areas can, for example, increase runoff and decrease evapotranspiration (Yang et al. 2017), while reforestation can reduce peak discharges during flood events (Kalantari et al. 2014). Furthermore, changes associated with flood risk and channel/floodplain restoration have also been reported (Thompson et al. 2017; Marhaento et al. 2018). Application of hydrological modelling has permitted investigations of the hydrological impacts of land use or land cover change by comparing current conditions with results of historical scenarios (e.g. Zhu and Li 2014) and/or predicted future changes (e.g. Bormann and Elfert 2010). These studies have revealed similarities between results across different environments. Land use changes have been reported to modify annual average discharges by 10% (Li and Wang 2009; Huisman et al. 2009; Bormann and Elfert 2010). Larger changes have been established for individual flood events (Bingqing et al. 2015; Kalantari et al. 2014).

In many cases, sensitivity analyses have demonstrated that the most sensitive parameters within distributed hydrological models are those associated with the saturated zone (Ross et al. 2009; Doummar et al. 2012; Thompson 2012; Vázquez and Hampel 2014). Representing groundwater - surface water interactions is critical for the simulation of catchment runoff and water quality responses (Scheliga et al. 2018; Ó Dochartaigh et al. 2019). These results emphasise the often-crucial role of surface hydrogeology and the

requirement to accurately characterise catchment geology within hydrological models. The studies of Vannier et al. (2016) and Carlier et al. (2018) both showed how changes in the representation of catchment geology within hydrological models can strongly influence simulation results. Inaccurate data, for example due to limited geological surveys, have the potential to impact accurate representation of catchment hydrological responses with implications for using models to predict future change in factors that include land use as well as climate. However, there are a limited number of studies that have assessed these influences using alternative spatial representations of subsurface geology within hydrological models.

Thus, this study investigates the implications of employing alternative spatial datasets upon simulation results from a hydrological model of an experimental catchment, Eddleston Water in the Scottish Borders, UK. Initial catchment modelling employed the highest resolution (and so assumed most accurate) spatial data for land use and superficial geology. Subsequently the model was run with coarser-resolution national and international spatial datasets. The aim was to quantify the effects of changing the spatial representation of catchment characteristics on (i) model performance and (ii) water balance components within different sub-catchments of the Eddleston Water.

2. Study Area

The Eddleston Water is a 69 km² catchment in the UK's Scottish Borders. It has an elevation range of 150–600 m and annual precipitation of 980 mm. The river main stem flows generally north to south in an alluvium-infilled valley, joining the River Tweed at Peebles (Fig 1). Soils within the catchment are dominantly sandy-loams (53%) followed by loamy-sands (20%), peats (17%) and loam (10%). Peat dominates the Middle Burn and Shiplaw Burn sub-catchments whilst sandy loams are concentrated in the Middle Longcote sub-catchment where

they cover 99% of the area (JHI 2014). Superficial geology of both glacial and glaciofluvial quaternary origins covers most of the catchment (65%) and has a maximum depth of 14 m. Till is the dominant sediment (45%), followed by gravel and sand (11%), gravel (5%) and peat (4%) (Ó Dochartaigh et al. 2012; Fig 2). The geological structure of the eastern part of the catchment is mainly bedrock and only small bands of quaternary sediments are found adjacent to stream channels (Fig 2). The underlying bedrock is mostly Silurian greywacke that is extensively weathered in the upper layers (Pearce et al. 2014). Catchment land use is mostly grass (67%) followed by coniferous forest (10%, predominantly in the Middle Burn sub-catchment). Marsh, shrubs (concentrated in the Middle Longcote sub-catchment) and mixed forest each account for 5% of the total catchment area. Urban development, strips of arable land, water, and broadleaf forest cover the remaining area (Fig 2).

Research undertaken under the auspices of the Eddleston Water Project (Tweed Forum, 2016) has investigated the effects of catchment management and restoration measures that have been implemented for both ecological improvements and as a means of alleviating local flood risk (Archer et al. 2013; Tweed Forum 2016). The catchment response to forest and grass land cover was examined in the studies of Archer (2013) and Peskett et al. (2020). Since 2011 the catchment has been equipped with four rain gauges and six stream gauging stations (Fig 1). Allied research has included habitat mapping, based on an aerial survey in 2009 (Medcalf and Williams 2010), and a survey of superficial geology in 2011 (Auton 2011; Ó Dochartaigh et al. 2012; Ó Dochartaigh et al. 2019). Following a baseline monitoring period from early 2011 to summer 2013, catchment restoration measures were initiated (Gilvear et al. 2013; Tweed Forum 2016). These include re-meandering the main channel, creating offline storage ponds, implementing large woody debris restrictors in tributaries, and planting new woodland (Tweed Forum 2016).

3. Methods

3.1. Hydrological model setup

A coupled hydrological/hydraulic model of the Eddleston Water catchment was developed using MIKE SHE/MIKE 11. This established modelling system has been employed in a wide range of situations from small catchments ($<25 \text{ km}^2$) or parts of catchments (e.g. Sahoo et al. 2006; Thompson et al. 2004; Thompson 2012) to large (several thousand km^2) river basins (e.g. Andersen et al. 2001; Singh et al. 2011; Thompson et al. 2013, 2014; Hudson and Thompson 2019). A $50 \text{ m} \times 50 \text{ m}$ computational grid was employed, resulting in the 69 km^2 catchment being discretized into 28,000 cells. A Digital Terrain Model (DTM) at 10m grid resolution was created from Ordnance Survey contours (OS, 2014) applying the “Topo to raster” tool in ArcGIS. The Richards equation approach was used for the unsaturated zone that was represented as two layers for soils (JHI 2014) and bedrock or superficial geology (Hughes 1996). A similar two-layer (upper zone: soils; lower zone: bedrock) approach was used in the finite difference saturated zone setup with superficial geology represented as lenses. Hydraulic parameters for the unsaturated (saturated hydraulic conductivity, Van Genuchten parameters) and saturated (horizontal and vertical hydraulic conductivity, specific storage and specific yield) zones were initially taken from the literature (Morris and Johnson 1967; Ó Dochartaigh et al. 2012; MacDonald et al. 2012; Thompson 2012; Foster and Allen 2015) and were subject to calibration within the bounds defined in this literature (Table 1).

Spatial distributions of superficial geology and land use were varied to investigate the impacts of alternative datasets upon model results (see Section 3.2). While a flashy response to rainfall was evident in the majority of sub-catchments, the Middle Longcote sub-catchment

on the eastern side (Fig 1) was an exception. Comparison between observed rainfall and stream discharge time series suggest that rainfall here tends to be stored and then slowly released. This is likely the result of the extensively weathered upper layers of greywacke (Pearce et al. 2014). This hydrological behaviour was represented using the MIKE SHE drainage option (DHI 2014) with drains located at 0.3 m below the surface and the drainage time constant varied during calibration.

The model used the finite difference approach for overland flow and the Kristensen-Jensen method for evapotranspiration. For the latter, a daily time series of reference evapotranspiration (FAO Penman-Monteith equation) was established using data from the central Darnhall automatic meteorological station (Fig 1 - the only station providing all of the necessary data) and distributed over the catchment within 16 zones following the approach of Foster and Allen (2015). Hourly precipitation from four rain gauges within the catchment was distributed using Thiessen polygons (Fig 1).

The MIKE 11 1D hydraulic model used the dynamic wave approximation of the St. Venant equations. Six branches were defined using the “Arc Hydro” extension of ArcGIS (Maidment 2002) and the DTM (Fig 1). A total of 280 cross-sections were specified throughout the river model based on topographic surveys (June 2013) of the six gauging stations. The river model included all six gauging station locations which provided hourly discharge observations against which model results were compared (see Section 3.4). The maximum allowed time step for all components of the MIKE SHE model was 1 hour, while a fixed time step of 5 seconds was applied in the MIKE 11 hydraulic model.

3.2. Alternative representations of land use and superficial geology

Four alternative versions of the MIKE SHE / MIKE 11 model were developed (Table 2). Whilst each had the same basic structure as described above, alternative data for land use and superficial geology were employed in each.

The first, referred to as the “Reference Model”, employed a 1:25,000 map of superficial geology (Auton 2011) and land use mapped from detailed aerial photography analysis (Medcalf and Williams 2010). Some aggregation of the categories within both maps was undertaken in order to obtain a manageable number of calibration parameters. The eight categories of the original superficial geology map were reduced to four based on particle size characteristics. Similarly, the original 39 land use categories were reduced to eight with the aggregation criteria based on roughness, leaf area index and root depth since these parameters are used in the land use parameterisation within the model (Fig 2). The second version of the model (“Land Use”) used a land cover map derived from the public CORINE Land Cover 2012 dataset (EEA 2012). A reduction in the number of categories from 13 to eight (consistent with the number of categories used in the Reference Model, see Table S1 in the Supplementary Material for differences) was undertaken. A third version of the model (“Superficial Geology”) employed a coarser 1:50,000 scale map of the superficial geology (Hughes 1996). The original four categories were reduced to two, these being the most common categories with the original 1:25,000 map after its aggregation (Fig 2, Fig 4, Table S2 in the Supplementary Material). The final version of the model (“Combination”) employed the CORINE-based land cover, again aggregated to eight categories, and the 1:50,000 superficial geology map, as before, aggregated to two categories (Fig 2).

The largest differences in land cover between the input data used in the Reference and Land Use versions of the model are in the Middle Burn and Middle Longcote sub-catchments (Fig 3 and Table S1 in the Supplementary Material). In Middle Burn, the differences result from an increase in the area of shrubs (from 0% to 56.1% of total area) and decrease of

coniferous forest (from 96% to 43.9%). In Middle Longcote the largest changes are due to an increase in wetland area (from 11.6% to 74.5%) and a decrease in the extent of grass (from 51.2% to 0%). The largest differences between these two versions of the model in the remaining sub-catchments are associated with an increase in the areas of shrubs.

The largest differences in the extent of superficial geology between the Reference and Superficial Geology models are associated with the presence of till (Fig. 4 and Table 2; Table S1 in the Supplementary Material). An increase in the extent of this class is evident in the Shiplaw Burn (from 50.4% to 74%), Early Vale (56.9% to 82.6%), Darnhall Mains (55.2% to 82.9%) and Kidston Mill (44.4% to 68.6%) sub-catchments, with consequent reductions in the extent of the other categories. In the remaining sub-catchments differences in superficial geology are smaller, with an increase of till (from 76.1% to 88.3%) in Middle Burn and a decrease in gravel (from 7% to 0%) in Middle Longcote.

3.3. Calibration and validation of the Reference Model

The simulation period (10/03/2011–31/08/2013) was limited by the initiation of catchment restoration measures. The first part of the period (10/03/2011–10/03/2012) was used for calibration and the latter part (10/03/2012–31/08/2013) for validation. Due to the relatively short length of the simulation period and the limited availability of hydrological data to drive the model, two years (10/03/2011–10/03/2013) of rainfall and reference evapotranspiration inputs were repeated for a warm-up period immediately prior to the simulation period in order to establish initial conditions.

A trial and error calibration approach (Vázquez and Feyen, 2003) was applied to the Reference Model and comprised several hundred model runs. The Nash-Sutcliffe Efficiency coefficient (NSE; Nash and Sutcliffe 1970) was systematically recorded for every gauging

station and an objective function defined as a mean value of the six NSE values (NSE_{overall}). The 20 parameter sets with the highest values of NSE_{overall} were used in the subsequent analyses. Additional performance metrics were applied to each of these model runs using the 20 parameter sets to further assess model behaviour: percent bias (PBIAS; equation 1) and maximum absolute error (E_{max} ; equation 2).

$$PBIAS = 100 \frac{\sum_{i=1}^N (S_i - O_i)}{\sum_{i=1}^N O_i} \quad (1)$$

$$E_{\text{max}} = \text{Max } |(S_i - O_i)| \quad (2)$$

where: S_i = simulated discharge at time i , O_i = observed discharge at time i . Cumulative values of water balance errors were also computed after each model run.

Subsequently, the Reference Model was validated using the same 20 parameter sets. The coefficient of variation (cv) was computed (absolute values were applied for PBIAS) to describe the variance of these simulations. Pairwise comparisons of the results of the calibration and validation were then conducted using Mann–Whitney U tests at a $P_{\text{crit}} = 0.005$ to assess model consistency. Observed and simulated discharges from each sub-catchment and 20 simulations were visualized using flow duration curves (FDC) and hydrographs.

3.4. Comparison of alternative representations of land use and superficial geology

The three models with alternative representations of land use and superficial geology were not independently calibrated/validated. Instead, the 20 alternative parameter sets obtained using the Reference Model were specified within each of these models. In applying common sets of parameters to all alternate models, our study followed the approach of Heisterman and Kneis (2011), although applied the data to a distributed rather than a lumped model and restricted the number of parameters sets to 20 to reflect the range within the parameter space whilst

maintaining logistically feasible computational times. In total, therefore, 80 model runs (20 parameter sets $\times$ 4 models) were undertaken. Differences between the Reference Model and the three alternative models were computed for the 2011/2012 hydrological year in every sub-catchment. A quantitative assessment of all 20 simulations was carried out on two levels: (i) the same model performance statistics employed in calibration (NSE, PBIAS, $E_{\max}$) and (ii) sub-catchment water balance components. For each sub-catchment, two major components of the water balance were assessed: drainage to the river (DR) and actual evapotranspiration (AE) in the Middle Longcote case and overland flow to the river (OF) and AE in remaining sub-catchments. These components were selected as the largest two terms of the simulated water balance in the respective sub-catchments although no direct observations are available. Kruskal–Wallis tests ($p=0.001$) were conducted to determine if the model performance statistics and water balance components were significantly different across the four versions of the model. Pairwise comparisons of the four models were then conducted using Mann–Whitney U tests using $P_{\text{crit}} = 0.005$ for each of the three levels of comparison to assess, firstly, the significance of any differences between alternative models and, secondly, whether the effect was predominantly the result of alternative Land Use or Superficial Geology datasets. The mean values across the 20 parameter sets of each model performance statistic and water balance component were also computed for each model using alternative representations of land use and superficial geology and compared with those from the Reference Model.

4. Results

4.1. Calibration and validation of the Reference Model

Superior values of the objective functions (NSE_{overall}) were obtained for the calibration (0.50-0.69) compared to the validation period (0.50-0.60). Generally, in the majority of sub-catchments, model performance based on both flow dynamics (NSE and E_{max}) and water balance ($PBIAS$) was superior for the calibration period compared to the validation period. Overall fits between the observed and simulated discharges were (with one exception) qualitatively good. Simulated peak discharges were characteristically advanced by on average 1.5 hours compared with the observed flows. Although the model tended to underestimate low and mean flows, it was able to reproduce very well the catchment dynamics of hourly flows during events of largest rainfall (Fig 5).

Differences in performance statistics between the sub-catchments are shown in Fig 6. Higher values of $PBIAS$ indicate underestimation (negative values) with the exception of the Shiplaw Burn where overestimation was indicated (see also Fig 5 and Fig 6). The Mann-Whitney U test established that differences in NSE across the 20 parameter sets for the calibration and validation periods were significant for the Early Vale and Middle Longcote gauging stations (Fig 6). Statistically significant differences in $PBIAS$ and E_{max} values for the calibration and validation periods were identified for all six stations. In general, lower values of the coefficient of variation were calculated for the calibration period compared to the validation period. However, in several cases, performance measures were not uniform in this conclusion (see Fig 6 for detail) and provide contradictory results in a single sub-catchment. The absolute mean cumulative water balance error was 0.2%, while the largest error was 1.3%.

4.2. Implications of alternative representations of land use and superficial geology

The Kruskal-Wallis tests confirmed significant differences at $P_{crit}=0.001$ for the majority of model performance statistics and sub-catchments when comparing the four different versions of the model, with the exception of NSE for Middle Burn ($p=0.015$) and E_{max} for Middle Longcote ($p=0.033$). The values of E_{max} were more sensitive to changes in spatial datasets compared to those for NSE and PBIAS (Fig 7 and Table 3).

Although the largest differences between the Reference Model and other versions of the model were in the majority cases of associated with the Combination Model (for both levels of comparison), there are two exceptions and only these are reported below. The largest differences (based on the Mann-Whitney U test) in the Middle Burn and Middle Longcote sub-catchments were produced by the Land Use Model. However, in several sub-catchments, ranking of performance for the different hydrological model varied when values of the coefficient variation for the calibration and validation period were compared in single sub-catchment (see Fig 7 for detail). Absolute differences between the mean values of each performance statistic (derived from the 20 parameter sets and using the Reference Model and the three alternative models) are shown in Table 3. In the Middle Longcote ($PBIAS_{mean}$) and Middle Burn ($E_{max\ mean}$) sub-catchments, larger differences were produced by the Land Use and Superficial Geology models, respectively. In the majority of other sub-catchments, the Superficial Geology Model produced the lowest variance (cv) for most performance statistics across the 20 parameter sets compared to the other models. Elsewhere, the Reference Model produced the lowest values of cv. In the Shiplaw Burn and Middle Burn sub-catchments this was the case for both NSE and E_{max} whilst in Kidston Mill it only applied to NSE (Fig 7). This again confirms the non-uniformity of performance measures when differences caused by alternate models were compared in single sub-catchment.

By comparing differences in the NSE values for the Reference Model with those calculated for the alternative models in all sub-catchments and for all 20 parameter sets (3

alternative models, 6 sub-catchments, 20 parameter sets = 360 values) it is possible to assess the impact of parameter transfer. In 50.6% of cases, all of the alternative models produced larger NSE values (in the remaining cases values of NSE were better for the Reference Model). This percentage was higher for PBIAS (68.1%), but lower for $E_{\max}$ (41.7%). For the Middle Burn, Early Vale and Darnhall Mains sub-catchments, all three alternative models produced superior values of all three performance statistics compared with the Reference Model (Fig 7). The opposite was the case for Shiplaw Burn with superior values of the three statistics being produced by the Reference Model. For the remaining catchments, non-uniformity in the relative magnitude of the performance measures was evident as shown in detail in Fig 7.

The results of the quantitative assessment based on the two major water balance components of the different sub-catchments are shown in Fig 8. Kruskal-Wallis tests computed for these components confirmed significant differences between the four different model versions for the majority of sub-catchments. There are two exceptions: the insignificant differences for AE ($p=0.004$) in Early Vale and DR ($p=0.027$) in Middle Longcote.

The largest differences from the Reference Model for the OR component in the Middle Burn were produced by Land Use, while in Kidston Mill they were associated with the Superficial Geology Model (see Fig 8). According to the test statistics for AE, the largest differences from the Reference Model were produced by the Land Use Model in Middle Longcote and Early Vale.

The absolute differences between the Reference Model and each of the alternative models in the mean values of each individual water balance component across the 20 parameter sets are shown in Table 4. The largest differences for AE in the Middle Burn and Middle Longcote sub-catchment are again associated with the Land Use Model. In the majority of cases, the Superficial Geology Model produces the lowest values of coefficient of

variation (Fig 8). The direction of change in AE between the Reference Model and the alternative models varies between model versions. In contrast, consistent increases in OR for all sub-catchments are produced by the three alternative models compared with the Reference Model. However, since, as previously noted, no measurements for these specific water balance components are available, it is not possible to assess whether an alternative model produces better results compared to the Reference Model.

An overall synthesis of the two-level comparison (performance statistics and water balance items) is shown in Table 5. When comparing just the Superficial Geology and the Land Use models, the differences from the Reference Model were largest for Land Use in just two sub-catchments: Middle Burn and Middle Longcote. Changes in the spatial representation of superficial geology were responsible for the largest differences from the Reference Model in the remaining sub-catchments. In the majority of cases, the Superficial Geology Model produced the lowest values of coefficient of variation across the 20 different parameter sets.

5. Discussion

A manual calibration strategy was used to calibrate the Reference Model to hourly observed discharges at the six gauging stations. Although manual calibration has some advantages (e.g. Vaze et al. 2012) it remains a subjective approach and a series of drawbacks have been reported (Boyle et al. 2001; Vázquez and Hampel 2014). In the study of Vázquez and Hampel (2014) significant difficulty in the rigorous inspection of the model parameter space was caused by manual calibration. This led to biased conclusions, particularly related to the sensitivity of the parameters within the MIKE SHE evapotranspiration module. It would, therefore, merit future investigation using the model of the Eddleston Water catchment. In our study, calibration and validation were carried out for the Reference Model rather than

repeating the calibration for each alternative model (see Thompson et al. 2014). In this way, the specific effects of specifying different spatial datasets can be more clearly identified since the re-calibration to each input data could neutralize the influence of the effective parameters as discussed in studies of Heistermann and Kneis (2011) and Yen et al. (2015). In the study of Heisterman and Kneis (2011), alternative spatial representations were not assessed, but rather various methods of gap filling of precipitation time series were investigated. It was suggested that the relative performance of gap-filling methods might be obscured if a model were calibrated for each particular method. Furthermore it was noted that (i) applying a single one-parameter set for the rainfall-runoff model might favour particular methods over others, and (ii) the application of two different parameter sets, each optimized to one of the input data sets, is not a satisfactory solution because possible errors in each input data and the extent to which these errors are compensated by the choice of parameters remains unknown. The study concluded that it might be possible to produce equivalent runoff simulations from precipitation data of very different quality by ‘tuning’ the model, as also noted by Yen et al. (2015). In applying common sets of parameters to all input data, our study followed the approach of Heisterman and Kneis (2011). We further checked if our number of parameters sets (20) was appropriate (again following approach of Heisterman and Kneis, 2011) by applying split sample to subsets of 10 parameters sets and computing performance statistics from these subsets (see Table S4 in the Supplementary material). Results of this split shown the ranking based on performance statistics of two subsets does not change, demonstrating that the ranking of alternate models is appropriate.

The Reference Model provided satisfactory simulations of observed stream discharges based upon a comparison of the NSE objective function for each of the 20 parameter sets in both the calibration and validation periods with the critical threshold of ≥ 0.5 defined by Moriati et al. (2007). The majority of model simulations did not, however, yield acceptable

results based on the critical value of PBIAS ($\pm 25\%$) defined by Moriasi et al. (2007) during the validation period (see percentages in Fig 4). As noted by U.S. EPA (2009), unsatisfactory model simulations can be attributable firstly, to significantly different conditions in the calibration and validation periods, secondly to inadequate or improper model calibration, thirdly to inaccurate observed data, fourthly to the spatial accuracy of input data and finally to the inability of the model to represent catchment hydrological processes. In order to assess whether differences in the prevailing hydrometeorological conditions in the calibration and validation periods could impact model performance, the cumulative precipitation in six sub-catchments, as well as mean hourly observed and mean hourly simulated discharges at each gauging station were compared for the calibration period and the corresponding time frame within the validation period (i.e. 3/10/2012-3/10/2013). The differences between two periods are shown in Table 6. In every single case both precipitation and observed discharges were higher in the validation period. Whilst the difference in observed discharges between two periods in Early Vale and Middle Longcote were relatively large compared to those in the other sub-catchments, they could not be caused by the differences in observed precipitation over these sub-catchments (Table 6). This finding is also supported by the Mann-Whitney U test applied to compare calibration and validation performance measures (see test statistics in Fig 6). This may point to issues of data quality in the discharge records for these sub-catchments which could, in turn, be responsible for the largest differences between observed and simulated discharges. Furthermore, relatively large differences between observed and simulated discharges were evident for low flows in the Shiplaw Burn and Middle Burn sub-catchments (Fig 3). Longer simulation periods would be essential to investigate these issues further although, as previously noted, the simulation period in this study was limited to the period before initiation of catchment restoration measures.

The impacts of the alternative models were assessed using one hydrological characteristic of sub-catchment behaviour, hourly stream discharge for which observational datasets were available. Whilst this integrates the impacts of hydrological processes across each sub-catchment, if more comprehensive observations (including, for example, actual evapotranspiration, soil moisture, and groundwater level) had been available and employed in the calibration process, uncertainty in other hydrological processes could have been reduced (Beven 2006) and impacts on other components of catchment hydrological functioning fully investigated (Yen et al. 2015).

Although results of the Kruskal-Wallis and Mann-Whitney U tests indicate that alternative spatial data generated significant differences in model output in most sub-catchments for both levels of comparison, these differences were generally small for practical modelling purposes. This agrees with the study of Yen et al. (2015) who demonstrated relative insensitivity of flow prediction to land use mapping. An exception was the differences for $E_{\max}$, which were constantly high. This could indicate that flood peaks are more affected by the different representations of input data compared to the water balance (Huisman et al. 2009; Bingqing et al. 2015). This conclusion is also supported by the work of Kalantari et al. (2014) who reported a 60% increase in peak discharge caused by clear-cutting of 30% of their catchment area. We suppose that these differences could potentially be neutralized, or at least reduced, through re-calibration, obviously at the expense of increased uncertainty, as noted in the studies of Vázquez and Hampel (2014), Yen et al. (2015) and Alvarenga et al. (2016).

In the majority of sub-catchments, alternative representations of land use, superficial geology and a combination of both, produced superior results in terms of simulated stream discharges compared with the Reference Model. Thus, the use of lower resolution spatial data provides superior results to the models using higher resolution data. This conclusion is in agreement with others studies (Bell and Moore 2000; Camargos et al. 2018). The study of

Bell and Moore (2000) investigated the sensitivity of catchment runoff models to rainfall data at different spatial scales. It was concluded that the best model performance was obtained using lower resolution rainfall data and that improvements in the distributed model structure was required to make better use of the higher resolution information. In the study of Camargos et al. (2018), the SWAT model was used to investigate the effect of various representations of land use, soils and elevation on predicted river discharges. A lower resolution soil map and DEM yielded higher NSE compared to use of higher resolution datasets. Similar variation in model performance have been identified using alternative temporal data (Vázquez and Hamle 2014; Ruman et al. 2020). Vázquez and Hampel (2014) investigated the effect of various potential evapotranspiration inputs applying a MIKE SHE model and the GLUE calibration strategy. They concluded that comparable, and in several cases better, model performance was obtained using lower resolution inputs. It was suggested that this was caused by the interaction and combination of parameters which compensate for the differences between alternative input data. Ruman et al. (2020) compared different methods of precipitation gap filling and found that use of a dataset in which gaps had been filled produced better model performance compared to the original rainfall data in 33.4% of all cases. However, all these findings are contradictory to those of Yen et al. (2015) who compared different quality land use maps. The reason for this conflict may lie firstly in the already mentioned calibration strategy (manual calibration, transfer of parameters from the Reference Model, Yen et al. (2015), secondly in the model structure (Bell and Moore (2000) and thirdly in other sources of uncertainty inherent in the process of hydrological modelling (Fatichi et al. 2016).

Based on the relatively low number of rain gauges (four, dictated by the distribution of stations) used to force the hydrological model and the conclusions of Bell and Moore (2000), we hypothesize that the available rainfall inputs data may be a reason why the potential of a

higher resolution model was not realised. This suggests that the results reported here may well be valid for this rainfall input but that different results could be obtained if higher spatial resolution data were available. As shown by several authors (Masih et al. 2011; Lobligeios et al. 2014; Terink et al. 2018), smaller catchments are more sensitive to the spatial variability of rainfall input compared with larger catchments. Terink et al. (2018) studied rain-gauge sampling uncertainty caused by rainfall-runoff model structure based on high-resolution radar data using a small (6.5 km^2) lowland catchment in Netherlands. They concluded that discharge uncertainty can be as large as 60% of average hourly discharge if data from a single gauge is applied as the model rainfall input. This reduces to 10% when data from 40 rain gauges were employed. In the study of Lobligeios et al. (2014), the sensitivity of large number (181) of French catchments was assessed in relation to spatial variability of rainfall. They concluded that improvements in model performance were particularly pronounced for those catchments with documented higher spatial variability in precipitation. A review of the data for the Eddleston Water catchment suggests that, in line with the prevailing oceanic climatic conditions, spatial variability between the different rain gauges was small for the majority of rainfall record, the exception being some of the high-intensity rainfall events. The hydrological model applied in this study showed good performance (see hydrographs of selected events in Fig. 5 and NSE in Fig. 6) in reproducing catchment dynamics, while its main issue remains the water balance (PBIAS) as discussed above. Further research is necessary to explore the consequences of alternative model structures and spatial resolution of driving meteorological data (Bell and Moore 2000; Boyle et al. 2001; Andréassian et al. 2004; Masih et al. 2011; Lobligeios et al. 2014; Terink et al. 2018).

The model response for the Middle Burn sub-catchment to the transition from coniferous forest to grass when the lower spatial resolution land use data is applied agrees with other studies undertaken in the Eddleston Water catchment (Archer et al. 2013; Peskett et

al. 2020). In the first of these studies, higher rainfall infiltration was reported under the broadleaved forest compared to grass. In the latter, a marginal impact of the forest cover on soil moisture storage was reported during wetter periods whilst a reduction in catchment response to rainfall events during drier periods was identified including lower river flows following summer storms.

A critically important step in spatial data preparation for hydrological modelling is the aggregation of categories. In general, there is a trade-off between spatial accuracy and parsimonious model structure (avoiding model over-parametrisation) when developing distributed hydrological models (Beven 2006; Alvarenga et al. 2016). Different aggregation of the land use and superficial geology categories to those employed herein could produce alternative results and would require future investigation.

Based on an overall synthesis of the two levels of comparison, the largest differences from the Reference Model were clearly associated with the Combination Model. The justification for this conclusion is based on Fig. 7 and Fig. 8, which show how the median of the performance statistics of hourly discharge (Fig. 7) and water balance components (Fig. 8) shifts from the Reference Model, as calculated using the Mann–Whitney U tests. Similar quantification based on the NSE was shown by Camargos et al. (2018). There are two exceptions, where the largest differences were not associated with the Combination Model: Middle Burn and Middle Longcote sub-catchments. Firstly, this occurred when the variation between alternative datasets were very large (51.9% decline in the extent of coniferous forest for the Land Use Model compared to the Reference Model in Middle Burn and a 51% increase in wetlands in Middle Longcote). Secondly, larger differences for the Land Use Model resulted when the extent of superficial geology was relatively small (in the case of the Middle Longcote). Within the remaining sub-catchments, variations in the spatial representation of superficial geology produced larger differences compared to alternative land

use representations. Therefore, in the case of the Eddleston Water catchment, the spatial representation of superficial geology is more important compared to land use. This could be a consequence of a greater potential for superficial geology to store water (Käser and Hunkeler, 2016), and hence modify runoff responses to precipitation, compared to vegetation. Only in a single case (Water Balance component - Actual Evapotranspiration) from the two-level comparison did the Land Use Model produce a larger difference from the Reference Model compared to the Superficial Geology Model in all sub-catchments. This probably reflects the influence of vegetation characteristics on transpiration. However, these findings are clearly specific to the Eddleston Water catchment. Although the results may vary for catchments of different sizes and environmental characteristics, the results of this study suggest that particular attention should be paid to the spatial representation of superficial geology.

We applied three different model performance statistics (NSE, PBIAS and $E_{\max}$) which in several cases provided contradictory indications of relative model performance when comparing the Reference Model with the three alternative models using different spatial data in a single sub-catchment. These contradictions are likely because each performance statistic assesses different model behavior (e.g. high flow period, water balance, maximum peak flows) and are likely to persist with the adoption of additional measures of model performance. To support this conclusion, we applied the mean annual discharge as an additional performance measure for the water balance and compared it with the PBIAS (see Table S3 in the Supplementary Material). The ranking of these mean annual discharges for the alternative models replicated that of PBIAS. This issue is also connected with the question of model complexity. Models with simple (conceptual) structures can perform better than models developed using sophisticated (physically-based) code (Jehn et al. 2019). Thus, the only way of obtaining a consistent message would be to ignore some performance measures which, we contend would not be appropriate. This result emphasizes the need to apply several

performance statistics to compare outputs of hydrological models in order to obtain a rigorous assessment of model behaviour (Moriati et al. 2007; Ritter and Muñoz-Carpena 2013).

The magnitude of the differences in mean annual discharges associated with the models using different land use data are within the range of those reported in other studies (Huisman et al. 2009; Li and Wang 2009; Zhu and Li 2014; Yira et al. 2016; Paule-Mercado et al. 2017; Ahiablame et al. 2017). The largest mean difference in PBIAS (7.6%) was for the comparison of the Reference Model and Land Use Model (Table S1 in the Supplementary Material). Huisman et al. (2009) reported differences of around 10% from a multi-model comparison exercise. According to the results of Bormann and Elfers (2010), afforestation caused a decrease in annual runoff of approximately 20% in several cases. Im et al. (2009) reported a 15% change in total runoff when applying MIKE SHE models to a catchment near the metropolitan area of Seoul in Korea and its use to simulate the impacts of urban expansion. Similarly, Yang et al. (2017) and Zhu and Li (2014) reported increases in mean annual runoff equal to 7.3% and 3%, respectively as a result of urban expansion.

In comparison to studies that have investigated the impacts of alternative land use on hydrological model simulation results, there is a relative scarcity of studies assessing the implications of alternative spatial representations of superficial geology or bedrock geology within hydrological models. The studies of Vannier et al. (2016) and Carlier et al. (2018) do, however, support the hypothesis of the dominant influence of (superficial) geology on simulated catchment discharge. In both studies significant changes in model performance occurred following changes in the representation of geological layers. However, results of this case study are catchment specific and further investigation in other catchments is essential.

6. Conclusion

The results of this study demonstrated significant differences for most sub-catchments in simulated discharges and model performance between the Reference Model and each of the alternative models employing different spatial datasets. However, in practical modelling these differences were generally small, except those associated with $E_{\max}$. The results for $E_{\max}$ suggests that catchment responses to relatively large events may be more sensitive to the representation of land use and superficial geology. The largest differences from the Reference Model in most sub-catchments were associated with the Combination Model that employed simpler representations of both superficial geology and land use. There were two sub-catchments where this was not a case and where larger differences were due to alternative land use data. This happened firstly when variations between alternative land use datasets were very large and/or when the extent of superficial geology was relatively small. Superficial geology in isolation had more of an impact on model results than alternative land use. This finding suggests that particular attention should be paid to the representation of superficial geology in hydrological models although the relative sensitivity of superficial geology versus land use is likely to be catchment specific. The magnitude of differences between the Reference and Land Use models is within the range of differences in mean annual discharge reported in other studies that have investigated the impacts of land cover change on simulated river flows. The application of lower resolution spatial data provided superior model results to the models employing spatially finer (and assumed more accurate) representations of land use and superficial geology. This is in agreement with a number of other studies. This suggests that the lower resolution datasets may be more appropriate for our modelling application. Although three different kinds of model performance statistic were applied (NSE, PBIAS and $E_{\max}$), they did not provide consistent responses when differences between the Reference

Model and the three alternative models were compared in a single sub-catchment. This emphasises the necessity of using several model performance statistics in hydrological modelling to obtain more robust conclusions on model performance.

The Reference Model provided unsatisfactory results in terms of PBIAS during the validation period in most sub-catchments. This could be a consequence of the different conditions in the calibration and validation periods and/or inaccurate observed data. The model also underestimated low flows. To explore these issues in greater depth and to evaluate their effect on model results, longer simulation periods would be required than those employed herein. Furthermore, more observations beyond hourly discharge would provide opportunities to investigate catchment hydrological processes and the performance of the model. Although aggregation of the spatial data was necessary to obtain a reasonable number of categories, further research would be required to assess the influence of this procedure. Additional research could also address the potential for the differences between the Reference Model results and those of the three alternative models to be neutralized, or at least reduced, through re-calibration. This would require the application of automatic calibration techniques.

Finally, results from this study are specific to catchments with a temperate climate. An expansion of the approaches adopted herein would be required to assess the significance of alternative representations of land use and superficial geology (and potentially other spatial datasets) on catchments of different size and environmental conditions.

Acknowledgements

We would like to thank Mr. Milan Trizna and Mr. Matej Horáček for providing support for this work and to the Tweed Forum for permission to use outputs of the gauging network set up under the project.

Conflicts of interest:

The authors declare that they have no conflict of interest.

References

- Ahiablame, L., et al., 2017. Streamflow response to potential land use and climate changes in the James River watershed, Upper Midwest United States, *Journal of Hydrology: Regional Studies*, 14, 150-166. doi.org/10.1016/j.ejrh.2017.11.004
- Alvarenga, L.A., et al., 2016. Assessment of land cover change on the hydrology of a Brazilian headwater watershed using the Distributed Hydrology-Soil-Vegetation Model, *CATENA*, 143, 2016, 7-17. doi.org/10.1016/j.catena.2016.04.001
- Andersen, J., Refsgaard, J. C., and Jensen, K. H., 2001. Distributed hydrological modelling of the Senegal river basin - model construction and validation, *Journal of Hydrology*, 247(3-4), 200-214. doi.org/10.1016/S0022-1694(01)00384-5
- Andréassian, V., et al., 2004. Impact of spatial aggregation of inputs and parameters on the efficiency of rainfall-runoff models: A theoretical study using chimera watersheds, *Water Resources Research*, 40. doi.org/10.1029/2003WR002854
- Archer, N.L., et al., 2013. Soil characteristics and landcover relationships on soil hydraulic conductivity at a hillslope scale: A view towards local flood management, *Journal of Hydrology*, 497, 208–222. doi.org/10.1016/j.jhydrol.2013.05.043

Auton, C A., 2011. *Eddleston Water Catchment, Superficial Geology*, 1: 25 000. Edinburgh: British Geological Survey.

Badoux, A., et al., 2006. Investigations on the runoff generation at the profile and plot scales, Swiss Emmental, *Hydrological Processes*, 20, 377-394. doi.org/10.1002/hyp.6056

Bell, V.A., and Moore, R.J., 2000. The sensitivity of catchment runoff models to rainfall data at different spatial scales. *Hydrology and Earth System Sciences* 4(4), 653–667. doi.org/10.5194/hess-4-653-2000

Beven K., 1993. Prophecy, reality and uncertainty in distributed hydrological modelling, *Advances in Water Resources*, 16, 41 -51. doi.org/10.1016/0309-1708(93)90028-E

Beven K., 2006. A manifesto for the equifinality thesis, *Journal of Hydrology*, 320(1), 18-36. doi.org/10.1016/j.jhydrol.2005.07.007

Bingqing, L., et al., 2015. Analyses of landuse change impacts on catchment runoff using different time indicators based on SWAT model, *Ecological Indicators*, 58, 55-63. doi.org/10.1016/j.ecolind.2015.05.031

Blasone, R.S., Madsen, H. and Rosbjerg, D., 2008. Uncertainty assessment of integrated distributed hydrological models using GLUE with Markov chain Monte Carlo sampling, *Journal of Hydrology*, 353(1-2), 18–32. doi.org/10.1016/j.jhydrol.2007.12.026

Bormann, D, and Elfert, S. 2010. Application of WaSiM-ETH model to Northern German lowland catchments: model performance in relation to catchment characteristics and sensitivity to and use change, *Advances in Geosciences*, 27, 1–10. doi.org/10.5194/adgeo-27-1-201

Boyle, D. P., et al., 2001. Toward improved streamflow forecasts: Value of semidistributed modeling, *Water Resources Research*, 37 (11), 2749–2759. doi.org/10.1029/2000WR000207

Camargos, C., Julich, S., Houska, T., Bach, M., Breuer, L. 2018. Effects of Input Data Content on the Uncertainty of Simulating Water Resources, *Water*, 10(5), 621, <https://doi.org/10.3390/w10050621>.

Carlier, C., et al., 2018. Geology controls streamflow dynamics, *Journal of Hydrology*, 566, 756-769. doi.org/10.1016/j.jhydrol.2018.08.069

Doummar, J., Sauter, M. and Geyer, T., 2012. Simulation of flow processes in a large scale karst system with an integrated catchment model (Mike She) – Identification of relevant parameters influencing spring discharge, *Journal of Hydrology*, 426-427, 112–123. doi.org/10.1016/j.jhydrol.2012.01.021

DHI, 2014. MIKE SHE User manual Volume 1: User guide, Mike by DHI, Copenhagen.

Ó Dochartaigh, B.E., et al., 2012. Eddleston Water Floodplain Project: Data Report, Nottingham, 95.

Ó Dochartaigh, B.E., Archer, N.A.L., Peskett, L. et al. 2019. Geological structure as a control on floodplain groundwater dynamics. *Hydrogeol J* 27, 703–716. <https://doi.org/10.1007/s10040-018-1885-0>

EEA, 2012. *CORINE Land-cover 2012 vector data, version v.18.5.1* [dataset], Available from: <https://land.copernicus.eu/pan-european/corine-land-cover> [Accessed 22 Dec 2015].

Fatichi, S., et al., 2016. An overview of current applications, challenges, and future trends in distributed process-based models in hydrology, *Journal of Hydrology*, 537, 45-60. doi.org/10.1016/j.jhydrol.2016.03.026

Foster, S.B. and Allen, D.M., 2015. Groundwater - Surface Water Interactions in a Mountain-to-Coast Watershed: Effects of Climate Change and Human Stressors, *Advances in Meteorology*, Vol 2015. Article ID 861805. doi.org/10.1155/2015/861805

Freeze, R.A. and Harlan, R. L., 1969. Blueprint for a physically-based, digitally-simulated hydrologic response model, *Journal of Hydrology*, 9, 237–258. [doi.org/10.1016/0022-1694\(69\)90020-1](https://doi.org/10.1016/0022-1694(69)90020-1)

Gilvear, D.J., Spray, Ch.J. and Casas-Mulet., R., 2013. River rehabilitation for the delivery of multiple ecosystem services at the river network scale, *Journal of Environmental Management*, 26, 30-43. doi.org/10.1016/j.jenvman.2013.03.026

Heistermann, M. and Kneis, D., 2011. Benchmarking quantitative precipitation estimation by conceptual rainfall-runoff modeling. *Water Resources Research*, 47 (6), 1–23.

doi.org/10.1029/2010WR009153

Hudson, C.E., Thompson, J.R. 2019. Hydrological modelling of climate change impacts on river flows in Siberia's Lena River Basin and implications for the Atlantic Meridional Overturning Circulation, *Hydrology Research*, 50 (6): 1577–159. doi.org/10.2166/nh.2019.151.

Hughes, R.A, 1996. *Geological map series, Sheet number: 24E, Sheet title: Peebles*, 1: 50 000, Edinburgh: Geological Survey of Scotland.

Huisman, J.A., et al., 2009. Assessing the impact of land use change on hydrology by ensemble modelling (LUCHEM) III: scenario analysis, *Advances in Water Resources*, 32, 159–170. doi.org/10.1016/j.advwatres.2008.06.009

Im, S., Kim, H., Kim, C. et al. 2009. Assessing the impacts of land use changes on watershed hydrology using MIKE SHE. *Environ Geol* 57, 231. doi.org/10.1007/s00254-008-1303-3

Jehn, F.U., Chamorro, A., Houska, T., Breuer, L. 2019. Trade-offs between parameter constraints and model realism: a case study. *Scientific Reports*, 9(1) 10729. doi:10.1038/s41598-019-46963-6.

JHI, 2014. James Hutton Institute, Soil Map (National Soil Map), 1:250,000 ,vector data [dataset].

Kalantari, Z., Lyon, S.W., Folkeson, L., French, H.K., Stolte, J., Jansson, P., Sassner, M.

2014. Quantifying the hydrological impact of simulated changes in land use on peak discharge in a small catchment, *Science of The Total Environment*, 466–467, 741-754. doi.org/10.1016/j.scitotenv.2013.07.047.

Käser, D., and Hunkeler, D., 2016. Contribution of alluvial groundwater to the outflow of mountainous catchments, *Water Resources Research*, 52(2), 680-697. doi.org/10.1002/2014WR016730

Li, Y.K., and Wang, C.Z., 2009. Impacts of urbanization on surface runoff of the Dardenne Creek Watershed, St. Charles County, Missouri, *Physical Geography*, 30 (6), 556-573. doi.org/10.2747/0272-3646.30.6.556

Lobligeois, F., Andréassian, V., Perrin, C., Tabary, P., Loumagne, C. 2014. When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events, *Hydrol. Earth Syst. Sci.*, 18 (2), 575-594. doi.org/10.5194/hess-18-575-2014

Maidment, D. R., 2002. Arc Hydro: GIS for Water Resources, ESRI, California, 203

Masih, I., Shreedhar, M., Ulenbrook, S., Smakhtin, V. 2011. Assessing the Impact of Areal Precipitation Input on Streamflow Simulations Using the SWAT Model, *Journal of the American Water Resources Association (JAWRA)*, 47(1), 179-195. doi.org/10.1111/j.1752-1688.2010.00502.x

Medcalf, J., and Williams, H., 2010. Scottish Borders Council and Tweed Forum Consortium

Tweed Aerial Survey Phase 2: Aerial Photography Interpretation Land Cover Classification and Habitat Mapping. Final Report. Available from www.scotborders.gov.uk. Accessed [03.10.2018].

MacDonald, A.M., et al., 2012. Relating in situ hydraulic conductivity, particle size and relative density of superficial deposits in a heterogeneous catchment, *Journal of Hydrology*, 434–435, 130-141. doi.org/10.1016/j.jhydrol.2012.01.018

Marhaento, H., Booij, M.J., and Hoekstra, A.Y., 2018. Hydrological response to future land-use change and climate change in a tropical catchment, *Hydrological Sciences Journal*, 63(9), 1368 – 1385. doi.org/10.1080/02626667.2018.1511054

McMillan, H., and Clark, M., 2009. Rainfall-runoff model calibration using informal likelihood measures within a Markov Chain Monte Carlo sampling scheme. *Water Resources Research*, 45(4). doi.org/10.1029/2008WR007288

Moriasi, D.N., et al., 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Transactions of the ASABE*. 50 (3): 885–900. doi.org/10.13031/2013.23153

Morris, D.A., and Johnson, A.I., 1967. Summary of hydrologic and physical properties of rock and soil materials as analyzed by the Hydrologic Laboratory of the U.S. Geological Survey, U.S. Geological Survey Water-Supply Paper 1839-D, 42.

OS, 2017, Ordnance survey, *PROFILE Contours, NTF geospatial data, Scale 1:10000*,

Tiles: t13ne, nt14ne, nt14se, nt15ne, nt15se, nt23nw, nt23ne, nt24sw, nt24nw, nt24ne, nt24se, nt25sw, nt25nw, nt25ne, nt25se, Updated: November 2009, Ordnance Survey (GB), [dataset], EDINA Digimap Ordnance Survey Service, Available from: <http://edina.ac.uk/digimap>, [Accessed 17 Apr 2013].

Paule-Mercado, M.A., et al., 2017. Influence of land development on stormwater runoff from a mixed land use and land cover catchment, *Science of The Total Environment*, 599–600, 2142–2155. doi.org/10.1016/j.scitotenv.2017.05.081

Pearce, D., et al., 2014. Glacial geomorphology of the Tweedsmuir Hills, Central Southern Uplands, Scotland, *Journal of Maps*, 10:3, 457–465. doi.org/10.1080/17445647.2014.886492

Peskett, L., et al., 2020. The impact of across-slope forest strips on hillslope subsurface hydrological dynamics, *Journal of Hydrology*, 581, 124427. doi.org/10.1016/j.jhydrol.2019.124427

Rahim E.A., et al., 2012 Application of MIKE SHE modelling system to set up a detailed water balance computation, *Water and Environment Journal*, 26 (4), 490–503. doi.org/10.1111/j.1747-6593.2012.00309.x

Ritter, A., and Muñoz-Carpena, R., 2013 Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments, *Journal of Hydrology*, 480, 33–45. doi.org/10.1016/j.jhydrol.2012.12.004

Ross, J.L., Ozbek, M.M., and Pinder, G.F., 2009. Aleatoric and epistemic uncertainty in

groundwater flow and transport simulation, *Water Resources Research*, 45.
doi.org/10.1029/2007WR006799

Ruman, S., Krpec, P., Rusnok, P. Black, A.R., Trizna, M., Ball, T. 2020. Impact of missing precipitation values on hydrological model output: a case study from the Eddleston Water catchment, Scotland, *Acta Geophys.* 68, 565–576, doi.org/10.1007/s11600-020-00409-0

Sahoo, G.B., Ray, C., and De Carlo, E.H., 2006. Calibration and validation of a physically distributed hydrological model, MIKE SHE, to predict streamflow at high frequency in a flashy mountainous Hawaii stream, *Journal of Hydrology*, 327(1–2), 94–109.
doi.org/10.1016/j.jhydrol.2005.11.012

Scheliga B, Tetzlaff D, Nuetzmann G, Soulsby C . 2018 Groundwater dynamics at the hillslope-riparian interface in a year with extreme rainfall. *J Hydrol* 564, 509–528.
doi.org/10.1016/j.jhydrol.2018.06.082

Singh, C.R., et al., 2011. Modelling water-level options for ecosystem services and assessment of climate change: Loktak Lake, northeast India, *Hydrological Sciences Journal*, 56:8, 1518–1542. doi.org/10.1080/02626667.2011.631012

Terink, W., Leijnse, H., Eertwegh, G.V.D., Uijlenhoet, R. 2018. Spatial resolutions in areal rainfall estimation and their impact on hydrological simulations of a lowland catchment, *Journal of Hydrology*, 563, 319–335. doi.org/10.1016/j.jhydrol.2018.05.045

Thompson, J.R., et al., 2004. Application of the coupled MIKE SHE/MIKE 11 modelling

system to a lowland wet grassland in southeast England, *Journal of Hydrology*, 293, 1–4, 151-179. doi.org/10.1016/j.jhydrol.2004.01.017

Thompson, J.R. 2012. Modelling the impacts of climate change on upland catchments in southwest Scotland using MIKE SHE and the UKCP09 probabilistic projections, *Hydrology Research*, 43(4), 507. doi.org/10.2166/nh.2012.105

Thompson, J.R., et al., 2013. Assessment of uncertainty in river flow projections for the Mekong River using multiple GCMs and hydrological models, *Journal of Hydrology*, 486, 1-30. doi.org/10.1016/j.jhydrol.2013.01.029

Thompson, J.R., Green, A.J. and Kingston, D.G., 2014. Potential evapotranspiration-related uncertainty in climate change impacts on river flow: An assessment for the Mekong River basin, *Journal of Hydrology*, 510, 259-279. doi.org/10.1016/j.jhydrol.2013.12.010

Tweed Forum, 2016. The Eddleston Water Project Report, (online). Edited by Spray. C..J. Available from www.tweedforum.org. Accessed [03.10.18].

Uhlenbrook, S. and Sieber, A., 2005. On the value of experimental data to reduce the prediction uncertainty of a process-oriented catchment model. *Environmental Modelling and Software*, 20(1), 19–32. doi.org/10.1016/j.envsoft.2003.12.006

Uusitalo, L., et al., 2015. An overview of methods to evaluate uncertainty of deterministic models in decision support, *Environmental Modelling and Software*, 63, 24-31. doi.org/10.1016/j.envsoft.2014.09.017

U.S. EPA, 2009. Guidance on the Development, Evaluation, and Application of Environmental Models, Council for Regulatory Environmental Modeling U.S. Environmental Protection Agency Washington, DC 20460.

Vannier, O., Anquetin, S., and Braud, I., 2016. Investigating the role of geology in the hydrological response of Mediterranean catchments prone to flash-floods: Regional modelling study and process understanding, *Journal of Hydrology*, 541. doi.org/10.1016/j.jhydrol.2016.04.001

Vaze, J., et al., 2012. Towards best practice model application, eWater Cooperative Research Centre, Australia, 46.

Vázquez, R.F., Feyen, J. 2003. Effect of potential evapotranspiration estimates on effective parameters and performance of the MIKE SHE-code applied to a medium-size catchment, *Journal of Hydrology*, 270 (4), 309-327. doi.org/10.1016/S0022-1694(02)00308-6

Vázquez, R.F., and Hampel, H., 2014. Prediction limits of a catchment hydrological model using different estimates of ET, *Journal of Hydrology*, 513, 216-228. doi.org/10.1016/j.jhydrol.2014.03.051

Wang, H., Tetzlaff, D., and Soulsby, CH., 2018. Modelling the effects of land cover and climate change on soil water partitioning in a boreal headwater catchment, *Journal of Hydrology*, 558, 520-531. doi.org/10.1016/j.jhydrol.2018.02.002

Walker, W.E., et al., 2003. Defining uncertainty a conceptual basis for uncertainty

management in model-based decision support, *Integrated Assessment*, 4 (1), 5-17.
doi.org/10.1076/iaij.4.1.5.16466

Xingguo M., et al., 2009. Grid-size effects on estimation of evapotranspiration and gross primary production over a large Loess Plateau basin, China, 54(1), 160 – 173, *Hydrological Sciences Journal*. doi.org/10.1623/hysj.54.1.160

Yang, L, Feng, Q, Yin, Z, Wen, X, Si, J, Li, C, Deo, RC. 2017. Identifying separate impacts of climate and land use/cover change on hydrological processes in upper stream of Heihe River, Northwest China, *Hydrological Processes*, 31, 1100– 1112. doi.org/10.1002/hyp.11098

Yen, H., Sharifi, A., Kalin, L., Mirhosseini, G., Arnold, J.G. 2015. Assessment of model predictions and parameter transferability by alternative land use data on watershed modeling, *Journal of Hydrology*, 527, 458-470, doi.org/10.1016/j.jhydrol.2015.04.076.

Yira, Y., et al., 2016. Modeling land use change impacts on water resources in a tropical West African catchment (Dano, Burkina Faso), *Journal of Hydrology*, 537, 187-199. doi.org/10.1016/j.jhydrol.2016.03.052

Zhu CH. and Li Y., 2014. Long-Term Hydrological Impacts of Land Use/Land Cover Change From 1984 to 2010 in the Little River Watershed, Tennessee, *International Soil and Water Conservation Research*, 2 (2), 11-21. doi.org/10.1016/S2095-6339(15)30002-2

Table Titles

Table 1. Principal calibrated parameters of coupled MIKE SHE / MIKE 11 model. Highlighted parameters were varied according to the presented range in 20 parameter-sets.

Table 2. Spatial data input sources employed in alternative model versions.

Table. 3: Mean differences in performance statistics. Using the 20 values of performance statistic. The mean was computed from absolute differences between the reference model and alternate model versions. For the $E_{\max}$, relative values in percent was calculated to allow comparison. LU – Land Use, Geol. – Geology, Comb. - Combination.

Table. 4: Mean differences in major water balance components between the reference model and alternate model version. Mean was computed from absolute differences between the reference model and alternate model version. Drainage to river was applied only to Middle Longcote sub-catchment. LU – Land Use, Geol. – Geology, Comb. - Combination.

Table 5: Overall results of the two-level comparison. Asterix denote the insignificant results of Kruskal-Wallis test. Abbreviation (Geol and LU) mark the version leading to highest differences between the Land Use and Superficial Geology. Bolded text marks the version causing the highest biases among all alternate versions (if not bolded, the Combination version was producing highest biases). Superscripts denote the version with the lowest values of coefficient of variation. L – Land Use; G – Superficial Geology; WB – water balance.

Table. 6: Differences between the selected characteristics of part of the validation (2012) and calibration (2011) period. The length of the part of the validation period was equal to the length of the calibration (i.e. 10/3/2012 - 10/3/2013).

ACCEPTED MANUSCRIPT

Figure Titles

Fig 1: Geographical position of the study area. A – localization within western Europe; B – position within south-west Scotland; C- geographical situation of close surroundings of the Eddleston Water catchment. Total catchment area: 69 km², Middle Burn 2.1 km², Shiplaw Burn 3.6 km², Middle Longcote 3.7 km², Early Vale 26.1 km², Darnhall Mains 30.3 km², Kidston Mill 62.7 km². For the numerical values of annual rainfall totals for every single sub-catchment see Fig. 8. Data Sources: Geographic Coordinate System: GCS OSGB 1936.

Fig 2: Input data of alternative versions applied in the model set-up. A – D: Superficial geology. A – Original map, 1:25 000; B – Generalized 1:25 000 map; C – Original map 1:50 000; D – Generalized 1:50 000 map; E – F: Land Use; E – Generalized categories from Aerial Photographs (Original categories of Aerial Photographs not shown due to high number of categories – 39); F – original CORINE Land Cover; G – Generalized model CORINE. Data Sources: Hughes, 1996; Medcalf & Williams, 2010; Auton, 2011; EEA, 2016; Geographic Coordinate System: GCS OSGB 1936.

Fig 3. Comparison of land use categories applied in the Reference model and Land Use model. For the absolute sub-catchment area, see Fig. 1.

Fig 4. Comparison of superficial geology categories applied in the Reference model and Superficial geology model. For the absolute sub-catchment area, see Fig. 1.

Fig 5: FDC and hydrographs of observed and simulated discharges for sub-catchments of Eddleston Water. Legend: a – observed discharges, b – average from 20 simulations, c – range of 20 simulations.

Fig 6: Performance statistics for the calibration and validation. Whiskers are equal to 10 and 90 percentile ($n = 20$ for each box plot). Results of Mann–Whitney U test ($p=0.005$, significant values highlighted), as well as coefficient of variation (cv), are shown. Percentages indicated the cases when simulation met the criterion of satisfactory model.

Fig 7: Comparison of alternative representation of land use and superficial geology in sub-catchments of Eddleston Water based on performance statistics computed for the hydrological year 2011/2012. Whiskers are equal to 10. and 90. percentile ($n = 20$ for each box plot). Results of Mann–Whitney U test ($p=0.005$, significant values highlighted), as well as coefficient of variation (cv), are shown. Representation generating highest difference to the reference version is marked by the dotted rectangle. Percentage indicate superior results of performance statistics compared to the reference model from 20 simulations.

Fig 8: Comparison of alternative representation of land use and superficial geology of Eddleston Water based on simulation of major water balance components computed for the hydrological year 2011/2012. Whiskers represent 10th and 90th percentiles ($n = 20$ for every single box plot). Results of Mann–Whitney U test ($p=0.005$, significant values highlighted), as well as coefficient of variation (cv), are shown. The representation that generated the highest difference compared with the reference version is marked by the dotted rectangle.

	Item / parameter	Mannings M	Root Depth (mm)	LAI (max)	Crop Coefficient (max)		
OVERLAND FLOW	Coniferous Forest	1.25 - 2.5	1562	6.50	1		
	Mixed Forest	1.25 - 2.5	1526	5.50	0.7 - 1		
	Broadleaf Forest	1.25 - 2.5	1526	4 - 6.5	0.8 - 1		
	Urban	20 - 100	0	0.9 - 1.25	1		
	Grass	30	540	4	1		
	Water	100	0	1	1		
	Shrubs	10	1137	4.0	0.7 - 0.8		
	Wetlands	20	150	4.0	1		
	Arable Land	16.67 - 28.57	0 - 880	2 - 4	1		
		Detention Storage (mm)		Manning n			
	Uniform	0.6 - 5		0.045 - 0.055			
UNSATURATED ZONE	Item / parameter	Saturated moisture content	Residual Moisture Content	Alpha	n	Saturated Hydraulic Conductivity	Bulk Density (kg.m ³)
	Loamy Sand	0.415	0.041	0.029	1.496	1.46 ⁻⁵	1330
	Sandy Loam	0.387	0.039	0.038	1.47	2.77 ⁻⁵	1310
	Peat	0.4	0.02	0.12	1.7	1.00 ⁻⁷	1200
	Loam	0.41	0.074	0.01	1.49	7.98 ⁻⁷	1470
	Built-up	0.423	0.056	0.014	1.486	1.00 ⁻¹³	2400
	Alluvial	0.46	0.034	0.19	1.8	7.00 ⁻⁵	1345
	Wacke	0.387	0.039	0.383	1.47	1.00 ⁻⁵	2400
		Maximum bypass Fraction					
	Uniform	0.15 - 0.35					
SATURATED ZONE	Item / parameter	Horizontal Hydraulic Conductivity (m.s ⁻¹)		Vertical Hydraulic Conductivity (m.s ⁻¹)	Specific Yield	Specific Storage	
	Loamy Sand	1.46 ⁻⁷		1.46 ⁻⁶	0.2	0.1	
	Sandy Loam	1.26 ⁻⁶ - 6.28 ⁻⁷		1.26 ⁻⁵ - 6.28 ⁻⁷	0.2	0.1	
	Peat	1.00 ⁻⁶ - 4.00 ⁻⁶		1.00 ⁻⁷ - 4.00 ⁻⁷	0.2	0.1	
	Loam	1.28 ⁻⁷		1.28 ⁻⁶	0.2	0.1	
	Built up	2.22 ⁻¹³		2.22 ⁻¹²	0.2	1.00 ⁻³	
	Sand, Silt, Clay	4.42 ⁻⁶		4.42 ⁻⁵	0.1	1.00 ⁻⁴	
	Till	1.20 ⁻⁸ - 2.40 ⁻⁸		1.20 ⁻⁷ - 2.40 ⁻⁷	0.17	0.017	
	Sand and Gravel	1.58 ⁻⁶		1.58 ⁻⁵	0.24	5.33 ⁻⁴	
	Peat	1.00 ⁻⁶ - 4.00 ⁻⁶		1.00 ⁻⁷ - 4.00 ⁻⁷	0.2	1.00 ⁻³	
	Gravel	3.00 ⁻⁶ - 1.00 ⁻⁶		1.00 ⁻⁷ - 3.00 ⁻⁷	0.24	7.55 ⁻⁴	
	Wacke	1.00 ⁻⁶ - 3.00 ⁻⁷		6.00 ⁻⁶ - 1.00 ⁻⁷	0.21	5.25 ⁻⁴	
		Drainage Time Constant (sec)		Drainage Level (m)			
	Uniform	1.00 ⁻⁷		-0.3			

Table 1. Principal calibrated parameters of coupled MIKE SHE / MIKE 11 model. Highlighted parameters were varied according to the presented range in 20 parameter-sets.

ACCEPTED MANUSCRIPT

Model version	Type of input data with numbers of original and aggregated categories
Reference model	Superficial Geology Map 1:25,000 (8-4), Aerial Photograph (39 - 8)
Land Use	Superficial Geology Map 1:25,000 (8 - 4), Corine Land Cover (13 - 7)
Superficial Geology	Superficial Geology Map 1:50,000 (4 - 2), Aerial Photograph (39 - 8)
Combination	Superficial Geology Map 1:50,000 (4 - 2), Corine Land Cover (13 - 7)

Table 2. Spatial data input sources employed in alternative model versions.

Sub-catchment / Performance statistics	NSEmean			PBIASmean			Emax mean (%)		
	LU	Geol.	Comb.	LU	Geol.	Comb.	LU	Geol.	Comb.
Shiplaw Burn	0.08	0.11	0.20	5.50	9.70	15.50	20.30	20.40	41.70
Middle Burn	0.04	0.05	0.05	6.10	3.00	8.80	20.80	30.40	24.40
Middle Longcote	0.18	0.04	0.17	7.60	4.30	6.10	5.50	5.10	19.90
Earlyvale	0.04	0.05	0.06	3.80	8.50	11.10	4.50	7.20	9.60
Darnhall Mains	0.02	0.04	0.04	5.50	12.90	14.00	13.70	18.70	24.10
Kidston Mill	0.06	0.07	0.15	7.20	13.40	17.40	23.90	33.30	54.10

Table. 3: Mean absolute differences in performance statistics. Using the 20 values of performance statistic. The mean was computed from absolute differences between the reference model and alternate model versions. For the Emax, relative values in percent was calculated to allow comparison. LU – Land Use, Geol. – Geology, Comb. - Combination

Sub-catchment / Water Balance component	Overland Flow to River _{mean} / Drainage To river _{mean} (%)			Actual Evapotranspiration _{mean} (%)		
	LU	Geol.	Comb.	LU	Geol.	Comb.
Shiplaw Burn	1.8	5.5	7.1	1.9	0.6	1.9
Middle Burn	4.4	2.0	6.7	5.6	0.8	5.1
Middle Longcote	3.3	2.0	5.7	4.1	0.8	4.0
Earlyvale	1.0	8.0	9.3	0.5	0.5	0.6
Darnhall Mains	3.1	12.7	15.1	3.0	1.0	3.3
Kidston Mill	4.8	6.6	12.5	4.0	0.5	4.2

Table. 4: Mean absolute differences in major water balance components between the reference model and alternate model version. Drainage to river was applied only to Middle Longcote sub-catchment. LU – Land Use, Geol. – Geology, Comb. - Combination.

Sub-catchment	Performance statistics			WB components	
	NSE	E _{max}	PBIAS	OR/DR	AE
Shiplaw Burn	Geol ^R	Geol ^R	Geol	Geol ^R	LU ^G
Middle Burn	Geol ^{*R}	Geol ^R	LU	LU ^R	LU ^L
Middle Longcote	LU ^L	Geol ^{*L}	LU	LU*	LU ^G
Earlyvale	Geol ^G	Geol ^G	Geol	Geol ^G	LU ^{*G}
Darnhall Mains	Geol ^L	Geol ^G	Geol	Geol ^G	LU ^G
Kidston Mill	Geol ^R	Geol ^G	Geol	Geol ^L	LU ^G

Table 5: Overall results of the two-level comparison. Asterix denote the insignificant results of Kruskal-Wallis test. Abbreviation (Geol and LU) mark the version leading to highest differences between the Land Use and Superficial Geology. Bolded text marks the version causing the highest biases among all alternate versions (if not bolded, the Combination version was producing highest biases). Superscripts denote the version with the lowest values of coefficient of variation. L – Land Use; G – Superficial Geology; WB – water balance.

Sub-catchment	Rainfall (mm/year)	Difference 2012 - 2011 %	
		Mean discharge (m ³ .s ⁻¹) Observed	Simulated
Middle Burn	12.84	62.53	29.54
Shiplaw Burn	12.93	40.30	30.82
Middle Longcotte	29.08	148.64	94.62
Earlyvale	15.22	130.82	40.82
Darnhall Mains	14.81	66.20	44.33
Kidston Mill	27.95	38.92	57.34

Table. 6: Differences between the selected characteristics of part of the validation (2012) and calibration (2011) period. The length of the part of the validation period was equal to the length of the calibration (i.e. 10/3/2012 - 10/3/2013)

Subcatchment / Land Use category	Shiplaw Burn		Middle Burn		Middle Longcote		Earlyvale		Darnhall Mainz		Kidston Mill	
	Refr.	LU	Refr.	LU	Refr.	LU	Refr.	LU	Refr.	LU	Refr.	LU
Arable Land	12.3	0.0	0.4	0.0	1.1	0.0	4.9	8.4	4.4	9.3	3.6	4.5
Broadleaf Forest	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.5
Coniferous Forest	31.1	21.0	96.0	43.9	0.0	4.2	22.7	13.6	20.1	11.7	10.8	7.1
Grass	36.6	30.6	1.3	0.0	51.2	0.0	57.7	46.2	60.2	49.3	67.9	57.3
Mixed Forest	0.7	0.0	1.5	0.0	0.5	0.0	2.1	0.0	3.6	0.0	3.8	0.0
Shrubs	7.3	48.4	0.0	56.1	35.6	21.3	1.1	27.4	1.0	25.9	5.0	23.1
Urban	1.0	0.0	0.6	0.0	0.0	0.0	3.7	2.2	3.7	1.9	2.7	0.9
Water	0.1	0.0	0.2	0.0	0.0	0.0	0.2	0.0	0.2	0.0	0.2	0.0
Wetlands	10.9	0.0	0.0	0.0	11.6	74.5	7.6	2.2	6.8	1.9	5.6	6.6

Table S1: Absolute differences (%) in the area of land use categories used in Reference (Refr.) and Land Use (LU) version. For the absolute sub-catchment area, see Fig. 1.

Subcatchment / Superficial geology category	Shiplaw Burn		Middle Burn		Middle Longcote		Early Vale		Darnhall Mainz		Kidston Mill	
	Refr.	Geol	Refr.	Geol	Refr.	Geol	Refr.	Geol	Refr.	Geol	Refr.	Geol
Till	50.4	74.0	76.1	88.3	8.4	12.5	56.9	82.6	55.2	82.9	44.4	68.6
Sand and gravel	1.9	0.0	2.0	0.0	0.0	0.0	11.3	0.0	13.9	0.0	11.3	0.0
Peat	7.4	4.6	4.6	0.8	0.0	0.3	9.2	8.0	8.1	7.1	3.9	3.4
Gravel	10.1	0.0	0.1	0.0	7.0	0.0	3.3	0.0	4.5	0.0	5.8	0.0
Bedrock	30.3	21.4	17.2	11.0	84.6	87.2	19.3	9.4	18.3	10.0	34.6	28.0

Table S2: Absolute differences (%) between the area of superficial geology categories used in Reference (Refr.) and Superficial Geology (Geol) version. For the absolute sub-catchment area, see Fig. 1.

Model Version	Shiplaw Burn		Middle Burn		Middle Longcote		Early Vale		Darnhall Mainz		Kidston Mill	
	PBIAS	Mean	PBIAS	Mean	PBIAS	Mean	PBIAS	Mean	PBIAS	Mean	PBIAS	Mean
Land Use	2.75	2.51	5.25	8.46	3.65	5.01	0.50	0.88	1.40	2.09	3.10	4.59
Superficial Geology	8.90	8.15	1.45	2.31	1.60	2.28	6.85	12.13	11.10	36.48	11.15	16.52
Combination	30.05	9.24	6.75	10.80	1.95	2.59	7.20	12.75	11.75	17.51	11.50	20.24

Table S3: Absolute differences between Reference Model and alternate model versions for the Mean and PBIAS. Statistics were computed for the Hydrological year 2011/2012 as median values from 20 parameter sets.

Model Version	Sub-catchment	NSE			PBIAS			Emax		
		N=20	N=(first) 10	N=(last) 10	N=20	N=(first) 10	N=(last) 10	N=20	N=(first) 10	N=(last) 10
Reference	Shipley Burn	0.67	0.67	0.67	9.35	9.30	9.65	1.02	1.06	1.11
	Middle Burn	0.73	0.73	0.72	37.25	37.55	35.15	0.79	0.81	0.80
	Middle Longcotte	0.64	0.64	0.64	27.90	27.55	30.70	0.27	0.29	0.32
	Early Vale	0.68	0.68	0.68	43.35	43.60	43.15	40.43	40.33	40.01
	Darnhall Main	0.84	0.84	0.84	32.55	32.50	32.60	15.87	15.76	15.56
	Kidston Mill	0.71	0.71	0.71	32.55	32.40	32.70	17.17	16.34	16.14
and Use	Shipley Burn	0.64	0.63	0.65	12.10	11.60	13.40	1.14	1.18	1.23
	Middle Burn	0.74	0.74	0.74	32.00	32.25	31.30	0.69	0.76	0.74
	Middle Longcotte	0.43	0.43	0.40	31.55	31.45	33.10	0.28	0.28	0.32
	Early Vale	0.70	0.70	0.69	42.85	42.80	42.85	39.08	39.42	38.71
	Darnhall Main	0.85	0.85	0.85	31.15	30.95	31.35	14.09	14.60	13.96
	Kidston Mill	0.68	0.68	0.70	29.45	29.35	29.95	21.04	21.21	21.89
Natural Geology	Shipley Burn	0.56	0.56	0.60	18.25	18.05	18.75	1.21	1.18	1.22
	Middle Burn	0.75	0.75	0.75	35.80	36.15	33.80	0.58	0.74	0.60
	Middle Longcotte	0.67	0.67	0.58	26.30	26.20	27.90	0.27	0.27	0.31
	Early Vale	0.73	0.73	0.73	36.50	36.60	36.30	37.53	37.79	37.36
	Darnhall Main	0.87	0.87	0.87	21.45	21.50	21.45	12.86	12.90	13.05
	Kidston Mill	0.65	0.65	0.68	21.40	21.35	21.60	23.25	23.37	22.93
Contribution	Shipley Burn	0.51	0.51	0.51	19.40	19.15	23.70	1.37	1.40	1.52
	Middle Burn	0.76	0.76	0.76	30.50	30.80	28.70	0.65	0.77	0.72
	Middle Longcotte	0.42	0.43	0.40	29.85	29.80	31.70	0.27	0.28	0.39
	Early Vale	0.74	0.74	0.74	36.15	36.25	35.30	36.70	36.95	35.96
	Darnhall Main	0.86	0.86	0.86	20.80	20.80	20.70	11.78	12.40	11.90
	Kidston Mill	0.58	0.58	0.58	19.05	18.85	19.40	27.19	27.96	27.80

Table S4. Median values of performance statistics computed for the hydrological year 2011/2012 from all parameter sets (N=20) and subsets (first and last 10).