# GMD Submission by Coxon et al

# DECIPHeR v1: Dynamic fluxEs and ConnectIvity for Predictions of HydRology

## General Response

We thank the reviewers for taking the time to review the paper and their comments, which have greatly helped to improve this manuscript and the quality and clarity of the research.

The main comments from the reviewers focused on (1) better defining the novelty and originality of the proposed modelling framework, (2) clarity of the model structure and equations and (3) the model evaluation.

In response to these reviewer comments, we have substantially re-written sections of the abstract, introduction and key concepts to highlight the unique features of the modelling framework and how it compares to other modelling frameworks in hydrology. We have produced flow timeseries plots for a new figure in the paper that demonstrates the ability of the model to reproduce the observed flow timeseries for six catchments.

Due to the requested extra time the GMD editor kindly agreed to, we have also been able to implement a new, more computationally efficient and stable, analytical solution to the subsurface flow equations and have detailed this solution in a new appendix. The result of this has a) increased the novelty of the model because the solutions are a departure from those implemented in Dynamic TOPMODEL (Beven and Freer, 2001), b) resulted in improvements in the national results overall, thus all figures/results now reflect these new equations, c) sped up simulation times as the solution is no longer iterative and d) addressed the comments of Reviewer 2 who requested more detail about the implementation of our flux equations and our approach has now been fully derived.

Detailed responses to all comments are provided below. Author responses are in **bold** and any modifications to the manuscript are in *italic* below each of the reviewer's comments.

Gemma Coxon, March 2019

# Reviewer #1

This paper describes the development of the Dynamic fluxEs and Connectivity for Predictiosn of HydRology (DECIPHeR) framework for simulation of hydrology (especially river flow) at catchment to continental scales. The model is tested across the Great Britain at 1,366 gauges in the current study but the authors intend to expand the model domain and suggest that it can be applied at the continental scales. The framework appears to be efficient computationally but there are a number of issues that authors need to address before the manuscript can be considered for publication. I provide my specific comments below.

**We thank Reviewer #1 for taking the time to review the paper. We appreciate their comments and provide our responses below.**

(1) The authors should revise the introduction to clearly highlight the motivation behind and the need for such a framework in relation to numerous other ongoing model development efforts. For example, how does the proposed study advance hydrological modeling compared to the model presented by Chaney et al. (2016)? Further, there are a number of large-scale models that have the capability to simulate far more number of processes (e.g., groundwater dynamics, pumping, flood dynamics, human impacts) than those presented in the current framework (for example: Hanasaki et al. 2008; Ozdogan et al. 2010; Pokhrel et al. 2015; Wada et al. 2014). Certainly these models are intended for global/regional applications but there have been ongoing efforts to increase the spatial resolution (i.e., hyper-resolution models) for application of these models at smaller scales. Extensive review of these models is available in recent literature (Nazemi and Wheater 2015; Pokhrel et al. 2016; Wada et al. 2017). I suggest that the authors thoroughly revise the introduction including a discussion on these past/ongoing efforts. Note that most of these models use TOPMODEL to simulate some of the surface/sub-surface hydrologic processes.

**We agree that we needed to make this clearer and have revised the abstract and introduction to clearly highlight the motivation behind the framework and it's unique features in response to comments from Reviewer 1 and 2.**

(2) Since the framework is currently designed to primarily simulate river flow, it is also important to note studies on streamflow/flood simulations at local to continental scales (Bates et al. 2010; Miguez-Macho and Fan 2012; Yamazaki et al. 2011; Zhao et al. 2017). What is the rationale for having the new framework?

**We do not believe the studies suggested by the reviewer are relevant to this study. Bates et al (2010) presents a new set of equations for floodplain inundation, Miguez-Macho and Fan (2012) are focused on the role of groundwater and Yamazaki et al (2011) and Zhao et al (2017) are primarily investigating flow routing schemes. While these are interesting studies, they are not focused on modelling frameworks that simulate the key hydrological processes (e.g. infiltration, runoff generation, subsurface flows etc.) at catchment-based scales in the generation of river flow and thus are fundamentally different to the modelling framework presented here.**

(3) The above two issues are important because the authors' intent is to provide a framework for large-scale application.

**We agree that this is an important point and address this comment in the two responses above.**

(4) P4, L16-40: Why did the authors use HRUs instead of doing a fully-distributed model? Is it just the run time minimization? Is there a compromise in terms of adding new features such as groundwater flows and human water use? Again, I suggest adding a note on how this framework advances our capability to simulate the hydrology in comparison to numerous existing framework (see comments above)?

**One clear benefit of using HRUs is minimising the run times of the model. However, the key benefit is the flexibility it gives you to modify the spatial complexity/scale of how spatial variability and hydrologic connectivity are represented. This flexibility means you can (1) run the model as a fully distributed, semi-distributed or lumped model, (2) have more/less spatial/process complexity where needed in the landscape and (3) represent point scale features in the landscape whilst still maintaining modelling efficiencies elsewhere. These features are hugely beneficial to having a pre-defined fully distributed model which cannot handle such occurrences. We have modified section 2.1 to clarify this point.**

**There isn't any compromise in terms of adding new features as each HRU is treated as a separate store in the model which can have different process conceptualisations and parameterisations. This means that more process complexity can be incorporated where needed to better suit local conditions e.g. to account for 'point-source' human influences or more complex hydrological processes such as surface-groundwater exchanges.**

(5) P5, L15: "must contain no sinks": What if there are real inland sinks? There are too many across continents.

**Sink filling is very common in digital terrain analyses when generating river and catchment layers (for example the SRTM DEM used for HydroSheds undergoes a sink filling process before it can be used to derive catchment basins). We agree that this will mean any real inland sinks in the digital elevation model will be filled. Currently the modelling framework is unable to account for these features (such as lakes), however, this is a feature we will be looking to include soon.**

(6) Section 2.2.3: What is the routing scheme used? I find some description later in another section. Please consolidate the text and provide more details.

**In Section 2.2.3 (now Section 2.2.2) we are describing the river routing data that is generated by the digital terrain analysis. These data (such as the river network connectivity and routing tables) provide the information for several different routing schemes. The river routing scheme is then described fully in the model structure (Section 2.3.4) as this is the current routing scheme implemented in the model.**

**We have modified section 2.2.2 to guide the reader better:**

*"From the river network and gauge locations, the river network connectivity is derived with each river section labelled with a unique river ID and a suite of routing tables so that each ID knows it's downstream connections and to allow multiple routing schemes to be*

*configured (see section 2.3.4 for a description of the current routing scheme implemented in the modelling framework)."*

(7) P7, L24: "potential evapotranspiration": first, this term is used here and then abbreviated several times later. Second, why is PET required for rainfall-runoff modeling? Is it to calculate the actual ET? If yes, where is such description provided?

**We have removed all abbreviations of potential evapotranspiration in the document. Potential evapotranspiration is a common input for hydrological models and is used to calculate actual ET. The description of how this is calculated is given by Equation 2 in Section 2.3.4, but the model could also include other conceptualisations depending on the users requirements.**

(8) P7, L40: why and how was the 1mm/day set?

**The model needs a starting flow to initialise the storage deficits. Typically we take this from an observed flow time series but in some cases, particularly for ungauged flow points, there may be no flow time series available. In this case we define a starting flow of 1mm/day as a representative starting flow for most catchments. The choice of this initial starting flow only affects model flows during the initialisation period and has no effect once the flows are fully initialised. The model is always started with a 'spin up' period as would be normal standard practice.**

(9) P7, L43: what are the "internal states"? Some examples should be provided.

**We have modified this to** *"model stores and fluxes"* **to better clarify this point. These are described in full in Section 2.3.4.**

(10) P7, L45: How are runoff generation, infiltration, and soil moisture movement modeled? Are they done in the same manner as in the original TOPMODEL?

**These processes are described fully in Section 2.3.5 focused on the model structure. We have made clearer the differences between Dynamic TOPMODEL and DECIPHeR in Section 2.1 in response to Reviewer 2.**

(11) P8, L15: What does the "multiple different" refer to?

**In this case it refers to the model structures i.e. you can implement many different types of model structure within the model framework. We have removed the word 'different' to better clarify this point.**

(12) P8, L24: How is SRmax determined?

**SRmax is a parameter within the model that determines the soil root zone. The user can either set this to a default value or it can be sampled from parameter bounds as explained in section 2.3.3.**

(13) P9, L6: "kinematic wave" formulation: is this sufficient when applying the model over large continents where backwater flow and other river-flood dynamics are important (see: Bates et al. 2010; Miguez-Macho and Fan 2012; Yamazaki et al. 2011; Zhao et al. 2017).

**We believe the reviewer has mis-interpreted the routing used in the model. Channel river flow routing in the model is modelled using a set of time delay histogram. We**

**agree that using a set of time delay histograms may not be appropriate where backwater flow and other river-flood dynamics may be important. However, we would like to stress that the model is not intended to be a flood inundation model and is not trying to compute full hydrodynamics. Computation times for such models are significantly longer than here (hours rather than minutes for simulations over 30-40 years).**

**The model is flexible to accommodate other flow routing schemes (as discussed in Section 2.2.2) and allow for variability in channel routing at the reach scale to recognise changes in local routing velocities. This will certainly be an area of future research to improve the channel river flow routing. We have also recently coupled the model to LISFLOOD-FP to provide a better representation of river-flood dynamics in regions where this is important.**

(14) P9, L42: "evapotranspiration losses are highest . . ..": The figure shows PET, not the actual ET, and I believe high PET doesn't necessarily mean high ET (in water limited regimes). I think this argument is not supported unless the actual ET is shown. Could the authors clarify this?

**We have modified the text to clarify the data are potential evapotranspiration and not actual evapotranspiration.**

(15) P10, L32-L42: Is the river network map described consistent with the topography data described in the previous paragraph? Isn't it necessary to generate a river network map from the DEM used in the model?

**Yes, we ensure consistency between the river and the DEM by producing the river network used by the model from the DEM during the digital terrain analysis. As described in section 3.2, we extract headwater cells from an external river map (the Ordnance Survey MasterMap Water Network Layer) and then route these cells downstream via the steepest slope so that the DEM and the calculated stream network are consistent for flow accumulations based on surface slope. Consequently, it is generated from the DEM used in the model and thus consistent. We have better clarified this point in Section 3.2 to avoid confusion.**

(16) Section 3.3.1: Are the precip data used here same as those shown in Fig. 3.1?

**We believe the reviewer is referring to Figure 4a here. The data used to derive the hydro-climatic characteristics are the same as the model forcing data described in Section 3.3.1. We have made this point clearer in the text.**

(17) Section 3.3.2: What are the calibration and validation periods?

**As described in Section 3.3.1, daily data of precipitation, potential evapotranspiration and discharge for a 55-year period from 01/01/1961–31/12/2015 were used to run and assess the model. The year 1961 was used as a warm-up period for the model; therefore no model evaluation was quantified in this period and the model was evaluated from 01/01/1962 – 31/12/2015. In this study we don't use a split sample test with a calibration and validation period and instead choose to evaluate the model for the full time series available.**

(18) P11, L14-25: what is the use of PET here? In fact, it was not clear to me on what the forcing variables are. Typical hydrological models use Precip, Temp, Radiation, Humidity, Wind etc. If such variables are used, is the PET consistent with those forcing variables?

**Potential evapotranspiration is required as a forcing time-series for DECIPHeR to calculate actual evapo-transpiration. This is described in Section 2.3.2 which outlines the Data Pre-requisites of the model and Equation 2 in Section 2.3.5 which describes how potential evapotranspiration is used to calculate actual evapotranspiration.**

**We do not fully agree that 'typical' hydrological models use a range of data to construct PET internally. There are many hydrological models that use PET directly. Given there is considerable differences in how to construct PET then we often use more than one method to explore differences.**

**In this study (as described in Section 3.3.1), daily potential evapotranspiration (PET) data were obtained from the CEH Climate hydrology and ecology research support system potential evapotranspiration dataset for Great Britain (CHESS-PE) (Robinson et al., 2016). This dataset consists of 1km$^2$ gridded estimates of daily potential evapotranspiration for Great Britain from 1961 - 2015 calculated using the Penman-Monteith equation and data from the CHESS meteorology dataset (in this case air temperature, specific humidity, downward long- and shortwave radiation and surface air pressure). Consequently, potential evapotranspiration is calculated before being used as an input to DECIPHeR (as is common for many hydrological models).**

(19) Section 3.4.3 (P13, L33): The authors should present the actual time streamflow time series. Since this is the only the variable simulated/discussed, I was surprised that authors are not showing the time series plots. I suggest selecting certain representative gauging stations with varying catchment area and those located in different climatic regions for such analysis (it could be a 20 stations for example).

**Thank you for this suggestion. We have added a new figure and text to the manuscript that shows the flow time series results for six gauging stations.**

(20) Then, I also suggest showing the annual mean flow (rate or volume) as a scatter diagram for all gauging stations. Evaluation of high (Q5) and low (e.g., Q95) can also be presented similarly. Overall, the validation provided in the current version is not satisfactory/sufficient.

**The model is evaluated against a large sample of catchments (1,366) for a number of different metrics capturing the annual flow rate (bias in runoff ratio), low flows (bias in low flow volume) and high flows (nash-sutcliffe efficiency). While we appreciate the reviewer's suggestions, we present results that already evaluate the model's ability to capture these aspects of the flow regime (see Figure 7). The main aim of the paper is to provide a description of DECIPHeR and more detailed model evaluation is outside the scope of this paper.**

(21) P14, L23: "time series": where is this shown?

**This is now shown in the new figure.**

(22) P14, L39-45: The authors could discuss the appropriateness of different performance measures by referring some recent studies that have used a wide range of such performance

measures (Veldkamp et al. 2018; Zaherpour et al. 2018). This comment is relevant to P12, L5-15 as well.

**Thanks for the suggestions. We have added these references in section 4.1.**

(23) P15, L23: “groundwater dynamics and human influences”: Is the HRU-based representation a suitable choice for the representation of these missing factors? Would a fully distributed be required? Please also see a related comment earlier.

**Please see response to comment 4 above.**

(24) Finally, the authors should provide caveats in the current framework and the challenges in upscaling the framework to continental and possibly to global scales. The discussion regarding advancements compared to the existing models/ongoing efforts (e.g., the National Water Model) also becomes relevant here. A note on the use on the use HRUs, and not distributed grids, should also be made.

**Section 4.2 and 4.3 discuss the limitations of the modelling framework as applied in this study and the areas for future research. These limitations are very relevant to applying the framework across continental scales and we have made this clearer in the discussion.**

**Minor/editorial issues:**

(25) P2, L2: impact on “what”?

**We have modified sentence**

(26) P2, L8: some refs contain first names/initials

**We have removed the first name and initial from this reference**

(27) P11, L8: PET is abbreviated here but already used before.

**We have modified to include only the abbreviation**

(28) P12, L32: the catchment details are redundant with the information in Section 3.

**We disagree and believe these catchment details are essential information.**

**References:**


Bates, P. D., M. S. Horritt, and T. J. Fewtrell, 2010: A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. Journal of Hydrology, 387, 33-45.

Chaney, N. W., P. Metcalfe, and E. F. Wood, 2016: HydroBlocks: a fieldscale resolving land surface model for application over continental extents. Hydrol. Process., 30, 3543-3559.

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

Miguez-Macho, G., and Y. Fan, 2012: The role of groundwater in the Amazon water cycle: 1. Influence on seasonal streamflow, flooding and wetlands. J. Geophys. Res. Atmos., 117, D15113.

Nazemi, A., and H. S. Wheater, 2015: On inclusion of water resource management in Earth system models – Part 1: Problem definition and representation of water demand. Hydrol. Earth Syst. Sci., 19, 33-61.

Ozdogan, M., M. Rodell, H. K. Beaudoing, and D. L. Toll, 2010: Simulating the Effects of Irrigation over the United States in a Land Surface Model Based on Satellite-Derived Agricultural Data. Journal of Hydrometeorology, 11, 171-184.

Pokhrel, Y., N. Hanasaki, Y. Wada, and H. Kim, 2016: Recent progresses in incorporating human land–water management into global land surface models toward their integration into Earth system models. WIREs Water, 3, 548-574.

Pokhrel, Y., S. Koirala, P. J. F. Yeh, N. Hanasaki, L. Longuevergne, S. Kanae, and T. C5 Oki, 2015: Incorporation of groundwater pumping in a global Land Surface Model with the representation of human impacts. Water Resources Research, 51, 78-96.

Veldkamp, T. I. E., F. Zhao, P. J. Ward, H. de Moel, J. C. Aerts, H. M. Schmied, F. T. Portmann, Y. Masaki, Y. Pokhrel, and X. Liu, 2018: Human impact parameterizations in global hydrological models improve estimates of monthly discharges and hydrological extremes: a multi-model validation study. Environmental Research Letters, 13, 055008.

Wada, Y., D. Wisser, and M. F. P. Bierkens, 2014: Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources. Earth Syst. Dynam., 5, 15-40.

Wada, Y., M. F. P. Bierkens, A. de Roo, P. A. Dirmeyer, J. S. Famiglietti, N. Hanasaki, M. Konar, J. Liu, H. Müller Schmied, T. Oki, Y. Pokhrel, M. Sivapalan, T. J. Troy, A. I. J. M. van Dijk, T. van Emmerik, M. H. J. Van Huijgevoort, H. A. J. Van Lanen, C. J. Vörösmarty, N. Wanders, and H. Wheater, 2017: Human–water interface in hydrological modelling: current status and future directions. Hydrol. Earth Syst. Sci., 21, 4169-4193.

Yamazaki, D., S. Kanae, H. Kim, and T. Oki, 2011: A physically based description of floodplain inundation dynamics in a global river routing model. Water Resources Research, 47, W04501.

Zaherpour, J., S. N. Gosling, N. Mount, H. M. Schmied, T. I. E. Veldkamp, R. Dankers, S. Eisner, D. Gerten, L. Gudmundsson, and I. Haddeland, 2018: Worldwide evaluation of mean and extreme runoff from six global-scale hydrological models that account for human impacts. Environmental Research Letters.

Zhao, F., T. I. Veldkamp, K. Frieler, J. Schewe, S. Ostberg, S. Willner, B. Schauberger, S. N. Gosling, H. M. Schmied, F. T. Portmann, G. Leng, M. Huang, X. Liu, Q. Tang, N. Hanasaki, H. Biemans, D. gerten, Y. Satoh, Y. Pokhrel, T. stacke, P. ciais, J. Chang, A. Ducharne, M. Guimberteau, Y. Wada, H. kim, and D. Yamazaki, 2017: The critical role of the routing scheme in simulating peak river discharge in global hydrological models. Environmental Research Letters, 12, 075003.

# Reviewer #2

General comments

The authors extended the concept and code of Dynamic TOPMODEL and developed an improved model termed DECIPHeR v1. They applied it to the entire Great Britain by calibrating and validating at 1366 gauges, and claimed that the performance was satisfactory. As a hydrological modeler who has developed open source code and applied it to an extensive study domain, I fully acknowledge the considerable efforts the authors made. The paper is overall readable for most parts but seems lacking some important statements particularly on the novelty and originality. The key characteristics and strengths of DECIPHeR should be clearly stated in comparison with existing catchment and global hydrological models (the current form of paper only compares DECIPHeR with the original Dynamic TOPMODEL). Also the value and significance of the model application to the entire Great Britain should be further discussed (the current form displays the performance scores without referring any earlier efforts).

**We thank Reviewer #2 for taking the time to review the paper. We appreciate their comments and provide responses below.**

Specific comments

Page 1 Line 15: “a new flexible model framework”: Make this part more specific. What is a flexible model framework (or what is an inflexible model)? Also, add the key strengths and characteristics of DECIPHeR compared to existing hydrological models.

**We have modified the abstract to be more specific on the flexibility of the model framework (see response to comment below).**

**The key strengths and characteristics of DECIPHeR compared to existing hydrological models is now better discussed in the introduction in response to the comments of Reviewer #1 – we feel this discussion is more appropriate in the introduction rather than the abstract.**

Page 1 Line 18: “modified to represent different levels of heterogeneity, connectivity and hydrological processes as needed”: Make this part more specific. All models can be “modified to represent” these in some extent. Add more concrete words in what sense DECIPHer is more adaptable compared with other models.

**We have modified the abstract to be more specific on the flexibility/adaptability of the model framework.**

*“This paper presents DECIPHeR (Dynamic fluxEs and ConnectIvity for Predictions of HydRology); a new model framework that simulates and predicts hydrologic flows from spatial scales of small headwater catchments to entire continents. DECIPHeR can be adapted to specific hydrologic settings and to different levels of data availability. It is a flexible model framework which has the capability to (1) change its representation of spatial variability and hydrologic connectivity by implementing hydrological response units in any configuration, and (2) test different hypotheses of catchment behaviour by altering the model equations and parameters in different parts of the landscape.”*

Page 2 Line 30 “the underlying model structures do not have the flexibility to represent different levels of complexity in different landscapes”: Quite unclear. Since this part is crucially important to identify the research needs/questions, discuss concretely what have been already achieved and what are still lacking by earlier models.

**We agree this could be made clearer and have removed this sentence. We have significantly rewritten the introduction in response to comments from both reviewers to clarify the novelty of DECIPHeR and it’s differences to other modelling frameworks.**

Page 2 Line 42: “This is despite significant development of various modeling tools . . .”: Again quite unclear. What have been already achieved and what are still lacking by earlier models?

**See response to comment above**

Page 3 Line 36 “builds on the code and key concepts of Dynamic TOPMODEL.”: This sounds that DECIPHeR is an upgrade of Dynamic TOPMODEL. If this is the case, it is more readable to introduce the concept and formulations of Dynamic TOPMODEL first, then show the new functions and characteristics of DECIPHeR. Actually, the present form is hard to know what are same or different between two models.

**We agree this could be made clearer. We have rewritten Section 2.1 to ensure this point is clarified. As suggested by the reviewer we now introduce the key concepts of Dynamic TOPMODEL first and then make clear the changes we have made to the model code as follows:**

Page 4 Line 18 “To realise this, DECIPHeR uses hydrological response units (HRUs)”: It is hard to know whether the HRV concept has been already included in Dynamic TOPMODEL or not. I was confused similarly by many parts in this section. As mentioned earlier, please make it clear what are same or different between two models more clearly.

**See response to comment above.**

Page 6 Line 9 “In DECIPHeR, they provide the basis for river routing . . .”. Ibid.

**As now made clear in Section 2.1, the river routing code is completely new so this is unique to DECIPHeR.**

Page 8 Line 12 “2.3.5 Model Structure”: Unfortunately, I could hardly understand the model structure. Please describe all the equations for the terms in Figure 3 and the parameters in Table 1. At least describe where such full description of equations is available.

**We have modified this section to provide a better description of all the key equations and parameters shown in Figure 3. We have also included the derivation of the new analytical solution for the subsurface zone in the appendix (see comments in general response).**

Page 9 Line 9 “The parameter, SZM, sets..”: This paragraph is particularly hard to follow. Please show the key equations how these parameters work.

**We have modified this section and included the key equations for these parameters (see response to previous comment for modifications made to the manuscript).**

Page 12 Line 44 “3.4.2 Overall model performance” and Figure 6: I am wondering why the parameters are so insensitive to the results (i.e. it is surprising that 90% of parameter sets yield NSE >0). I am also puzzled why the entire ensemble outperforms the behavioral ensemble (top 1% performance, if I understood correctly). Please elaborate these points.

**We believe that Reviewer #2 has misinterpreted parts of these results. Figure 6 shows the percentage of catchments that meet the weaker and stricter performance thresholds for each catchment. Consequently, 90% of catchments yield NSE > 0, not 90% of the parameter sets (the number of parameter sets that achieve a score of NSE > 0 varies significantly between catchments). We have modified text in section 3.4.2 to make this clearer.**

**The ‘best score’ from the entire ensemble for any given metric is likely to outperform the best score from the behavioural ensemble as the behavioural ensemble is the top 1% based on the combined score of the four metrics. When creating a combined score of the four metrics, you would expect some trade offs between the different metrics as any simulation is unlikely to have the best score for all four metrics.**

Page 14 Line 21 “We calculated four evaluation metrics for 10,000 model simulations for 1366 GB gauges. . .”: Is this the first study to apply a hydrological to the entire Great Britain? If it is the case, clearly state so. If it is not, clearly refer the earlier efforts and compare the performance of them with this study.

**This isn’t the first study to apply a hydrological model to the entirety of Great Britain. However, it is the first to have such a comprehensive model evaluation against 1,366 gauges. We have made this clearer in the discussion and included a comparison of our model performance against other GB model evaluations in Section 4.1.**

Technical comments

Page 7 Line 27 “a parameter file specifying set parameter bounds for Monte-Carlo sampling”: Is “set” needed?

**Agree. We will remove ‘set’.**

Page 7 Line 42 Q_SAT: I guess this term first appears. Define what this term is.

**We have modified this to “***used as the starting value for QSAT (subsurface flow)”*

Page 12 Line 37: “13,600,600” reads 13,660,000.

**Thanks for spotting this. We have modified the text.**

Page 13 Line 4“The vast majority of gauges (90% of the whole ensemble)”: 90% of the gauges or 90% of the ensemble (i.e. 9000 simulations)?

**Apologies that this was not clear. We meant 90% of the gauges and have modified the text (see response to Page 12 Line 44 above).**

Page 14 Line 27: “is” reads in.

**We have modified the text as suggested.**

Page 34 figure 6: The caption says “weaker and stricter” while the figure says “upper and lower”.

**Thanks for spotting this.  We have modified the legend in the figure so it says weaker and stricter thresholds.**