Response to Reviewer's Comments : Anonymous Referee #1

Review of ms no. hess-2015-196 "Effectiveness of a regional model calibrated to different parts of a flow regime in regionalization" by H. S. Kim

In this manuscript, the author proposes a refined calibration approach to reduce the parametric uncertainty of the IHACRES rainfall–runoff model. The approach follows by calibrating the model separately to different parts of flow regime in 11 Korean river basins. Following this the author establishes potential linkages between catchment characteristics and the calibrated parameters to form regional models, concluding that "the regional models from the refined calibration approach clearly enhanced the hydrological behavior . . .". While the presented work clearly falls within the scope of this journal, there are some serious concerns (mentioned below) and unfortunately, I can't recommend the publication of this manuscript in its current form.

1. **(Referee Comment, RC)** Lack of novelty and advancement from the previous works: After reading the authors previous two papers published in the HP Journal (Kim and Lee; 2014a,b), I do not see any novelty in here presented current work. Both the seasonal calibration and regionalization of the IHACRES model parameters were already detailed in the previous study. Besides, there are many (significant) overlap between the authors previous and the current work - some figures are almost same. What is the point of presenting them again and again? One can just refer them to past publications. I, however, do not have any idea on how much overlap is allowed in the HESS journal considering that the work is already published previously in another journal (HP).

(Author Comment, AC)

This study investigates the possibility to improve the hydrological model parameter regionalization approach through rainfall-runoff model calibration for the different parts of flow regime (i.e., a seasonal calibration), particularly focusing on model behavior during dry periods.

The current paper is part of the same line of works by Kim and Lee (2014b, a); such as rainfall-runoff model, study sites and multi-objective optimization algorithm, but it has sufficient difference from such works in terms of research objectives. The work presented by Kim and Lee (2014b) addresses the use of regional regression of model parameters (through multi-objective calibration), but the seasonal calibration is not adopted; whereas Kim and Lee (2014a) proposes the use of a seasonal calibration (similar to the focus on the dry season applied to the present study), but does not carry out a regional regression of the model parameters (but only applies calibration of single catchments).

Kim and Lee (2014b) assessed the efficiency of the regionalisation approach based on a multiple objective calibration technique to rainfall–runoff modelling under the catchment conditions of the Republic of Korea. They concluded that the regional model with the multiple objective approach led to improved hydrological simulations in ungauged catchments over a single-objective approach, but there was still a flow-dependent bias (i.e. a tendency towards the behaviour of underestimating and overestimating high and low flows for the wet and dry periods, respectively) in the runoff simulations. The present paper builds on their analysis and examines the relative performance of regionalisation methods. The work presented in this paper is a methodological advancement of the work by Kim and Lee (2014b) in regionalisation studies. The present paper also goes beyond Kim and Lee (2014b) in three important aspects.

First, the variability (or consistency) of hydrologic response in model performance and parameter values and the NIREs (non-parametric impulse response estimates) have been implemented to investigate problems associated with the model accuracy under the different flow regime, which might be accumulated through regionalisation processes (i.e., leading to influences on the identification of relationships between the calibrated parameter values and the catchment characteristics as well as runoff predictions in ungauged catchments). Second, the calibration strategy (splitting the historical time series into wet and dry periods and calibrating data segments only for the dry period separately on the LRM (linear routing module) parameters because of a sufficiently high degree of similarity between the LRM parameter values for the wet period and those for the whole period) was established to improve the identification of relationships, according to an identification of deficiencies on model structure suitability for the different flow regime. The idea on seasonal calibration (Kim and Lee, 2014a) was accepted here. Third, Kim and Lee (2014b) concluded that a flow-dependent bias (i.e. underestimation and overestimation for the wet and dry periods, respectively) in the regional model might be reduced through modifying the model structure. However, the present study demonstrated the flow-dependent bias for the dry period in the regional model can be improved with an advanced calibration approach through improving the identification of the relationships between the catchment characteristics and the calibrated model parameters 'without modification of the model structure'.

Such issues were already discussed with the editors (handling editor: Dr. Elena Toth; handling executive editor: Prof. Dr. Erwin Zehe) through the editor decision (26 May 2015). The editors concluded that the present manuscript has sufficient difference from the previous works to allow being considered for publication and required to add an addition explanation on that in the manuscript. There was a minor revision on the introduction to make explicit what is in common with the two previous studies and what is novel in the submission after the discussion with the handling editors. Furthermore, more details on these issues (discussed above) will be added in the introduction and the conclusion sections in a next draft in order to clarify the advancements from the previous works.

2. **(RC)** Throughout the manuscript the author makes the case of model inadequacy to capture the hydrological effects of non-stationary catchment response dynamics under different climate conditions so to propose a refined calibration approach in which the model separately to different parts of flow regime (specifically to wet and dry periods). In context of the presented work, I have a fundamental concern on the usages of words like non-stationary catchment response and climate conditions. What are different climatic conditions used in this study? The seasonal course of the wet and dry periods of flow regimes is the part of hydro-climatic conditions, governed mainly by the seasonal precipitation and other forcing variables patterns. Also it is not clear to me what does the author mean by the non-stationary catchment response? Is the seasonal pattern of discharge time series classified as non-stationary response? If such is the case then the daily discharge time series which are highly variable could also be regarded as non-stationary catchment response. I would ask the author carefully consider these words that meant to be used for climatic time scale behavior and not for seasonal or daily time scale responses.

(AC)

This manuscript investigates the modelling limitations in the capacity to capture the variation in hydrological response of the catchment (often termed as non-stationarity, through this is not necessarily the correct use in the statistical meaning of this term, Croke and Shin (2015)) through time (i.e., the different climate conditions). The non-stationary catchment response dynamics (i.e., the non-stationarity in the hydrological model) is defined under the situation where hydrological model parameters vary in time and thus depend on the period of record used for their estimation. Such non-stationarity can lead to poor predictions, especially when the model is applied to a climatologically different period. In this study, the different climate conditions (hydro-climate conditions, as the reviewer said) were classified based on the wet (summers including September) and dry (the others) periods rather than the seasonal or daily time scale. 'The non-stationary catchment response dynamics (i.e., the variation in hydrological response of the catchment) ' and 'the different climate conditions (i.e., the different hydro-climatic conditions governed by the rainfall and streamflow distributions for the wet and dry periods) ' will be defined in the introduction to give clear meaning on these terms in the next draft.

3. **(RC)** Another major shortcoming of this paper is that the author does not consider evaluating the calibrated model parameters to the independent validation period. The whole available time series is used for the model calibration so how do we judge the model performance outside of the calibration

period? It is a standard practice in hydrological modeling to at-least retains a year of data outside of calibration period to evaluate the model performance.

(AC)

A model assessment scheme commonly used was proposed by Klemes (1986), involving a four-level test scheme (calibration and validation) involving simple split-sampling (stationary condition at the same catchment), differential split-sampling (non-stationary condition at the same catchment), simple proxy-catchment (stationary condition at different catchments), and differential proxy-catchment tests (non-stationary condition at different catchments). The term stationary was used to denote physical conditions that do not appreciably change with time (Klemes, 1986). The simple split-sampling test uses subsets of the time-series record for calibration and validation. On the other hand, the differential split-sampling test involves calibration and validation under different meteorological conditions (e.g. dry and wet periods). The simple proxy-catchment test can be used to show the general validity for different catchments. The differential proxy-catchment test is a combined approach, incorporating differential split-sampling and simple proxy-catchment tests. The simple split-sampling test is appropriate when the available record is sufficiently long to split into two equal parts. However, the other level tests should be considered if the available record cannot be meaningfully split (Klemes, 1986). The simple proxy-catchment test should be required when an available streamflow record in a catchment is not adequate for the simple split-sampling test and also be required as a basic test for the general model transposability such as ‘regionalization studies’ (Klemes, 1986).

The level test similar to the simple proxy-catchment test was implemented ‘in this regionalization study’. The calibrated parameter values for the traditional and the refined approaches (in the nine calibration catchments) were evaluated in the validation catchments (the two catchments that were excluded from the correlation analysis and multiple regression) through the comparative assessment of the accuracy of predictions of the regional models derived from the two approaches: the traditional and the refined approaches. This will be discussed more in the next draft.

In addition, the appropriate length of record needed in model calibration was considered in order to obtain time-invariant parameter values and statistics with a minimum amount of data and reasonably high model performance. At least 8 years or more length of record was needed to obtain a reasonable catchment response in calibration (i.e., to stabilize performance statistics and detect parameter consistency in calibration, which is the important issue in regionalization studies) in the IHACRES rainfall-runoff model. Kim et al. (2008) investigated the convergence of the model parameters to time-invariant values with increasing length of calibration period using the mean, standard deviations (SDs) and coefficients of variation of the various model performance indicators and parameter values. Kim et al. (2008) proposed that an 8-year calibration period is appropriate for obtaining a reasonable catchment response, yielding stable and reasonably high model performance and reduced variation in parameter values over time for the IHACRES. In this study, an 8-year calibration period from 2003 to 2010 was used based on the previous study. However, available streamflow records are not sufficiently long enough to split into two equal parts at some catchments such as Hoengseong, Nam River, Miryang, Yongdam and Buan catchments (Table1). This is another reason for the use of the level test similar to the simple proxy-catchment test in this study. Details on an evaluation approach for the calibrated parameter values caused by data limitation will also be added more in the discussion part (6. Summary and conclusion) in the next draft.

4. **(RC)** The first line of the abstract stating the objective of this study as "to reduce the parameter uncertainty . . .", but there is no uncertainty analysis conducted in this study. I could not find any advanced Monte-Carlo or GLUE or Bayesian type of parameter sampling analysis in this manuscript. Besides there is also an equifinality issue in hydrologic modeling studies (i.e. there are many parameter sets that can provide acceptable results). How did you choose one among many optimal parameter sets for the regionalization?

(AC)

As the reviewer's comment, the uncertainty issues on hydrologic modeling can be supported by the approaches to quantify the uncertainty in hydrologic predictions such as Generalized Likelihood Uncertainty Estimation (GLUE), frequentist approaches, Standard Bayesian approaches, Bayesian Recursive Estimation, etc. However, the overall goal of this research is to assess the adequacy of the regional model for different parts of a flow regime. The scope of this study is limited to identifying the effectiveness of the regionalisation method based on regional models calibrated to different parts of a flow regime through comparison with that those calibrated to the whole period. There is an indirect link to parameter uncertainty issues, but a quantification of the parameter uncertainty is not the focus of this study. Therefore, the parts of parameter uncertainty in the text will be removed by the reviewer's comment, and the objective in abstract, introduction and conclusion sections will be re-organized based on the scope (discussed above) in the next draft.

The uncertainty in parameter estimation is significantly increased when the 'best' parameter set cannot be uniquely located through model calibration (i.e. the equifinality problem of Beven and Freer (2001)). Such problems with parameter identification can be minimized by limiting the number of parameters to be estimated and by using information about the system to limit the range of possible values (Dunn, 1999;Seibert and McDonnell, 2002;Koivusalo and Kokkonen, 2003;Uhlenbrook and Sieber, 2005). Parsimonious model structures have been proposed to overcome the limitation of the high number interactions between parameters by reducing model complexity (Jakeman and Hornberger, 1993;Wagener and Wheater, 2006;Hundecha et al., 2008). Multi-objective approaches have been proposed as an alternative to single-objective measures to improve the identifiability of model parameters by increasing the amount of information retrieved from the available data (thus reducing parameter uncertainty) (Gupta et al., 1998;Yapo et al., 1998;Vrugt et al., 2003).

In this study, the parsimonious model structure, requiring minimal input data, was adopted to avoid adverse effects arising from model complexity, over-parameterization and data requirements. The IHACRES rainfall-runoff model (Jakeman and Hornberger, 1993) was applied to represent the dynamic response characteristics of the catchments. The 'best' model fits by calibration were selected by considering trade-offs among multiple statistics. A multi-criteria rejection algorithm (Kim et al., 2011) was adopted in this study. The multi-criteria rejection approach using a grid search provides an approximation of the Pareto optimal set. This algorithm minimized the loss of information for each objective function by avoiding a single aggregation of different multi-objective functions. Some of the issues related to the complexity of multi-objective optimization problems were circumvented by finding a reasonable solution without an evolutionary algorithm and a stochastic search scheme. Details on the multi-criteria rejection algorithm (discussed above) will be added in the modelling section (3. Modelling of the gauged catchments) in a next draft.

5. **(RC)** P7059: L7: The author should enhance the literature review by considering more recent work on hydrologic model parameter regionalization studies, for example, by Yadav et al 2007 (doi:10.1016/j.advwatres.2007.01.005), Pokrel et al 2008 (doi:10.1029/2007WR006615), Oudin et al 2008 (doi:10.1029/2007WR006240.), Samaniego et al 2010 (doi: 10.1029/2008WR007327), Kumar et al 2013 (doi:10.1029/2012WR012195) as also discussed in detail in the recent book by Blöschl, G.: Rainfall-runoff modeling of ungauged catchments, in: Encyclopedia of Hydrological Sciences, edited by: Anderson, M. G., John Wiley & Sons, Chichester, 2061–2080, 200

(AC)

A next draft will include additional literature review of rainfall-runoff model parameter regionalization research over recent years: a model independent approach to streamflow prediction in ungauged basins based on empirical evidence of relationships between watershed structure, climate and watershed response behavior (Yadav et al., 2007); an assessment of the effectiveness of hydrological model parameter regionalization approaches (Oudin et al., 2008;Merz and Blöschl, 2004;McIntyre et al., 2005;Parajka et al., 2005;Kay et al., 2006;Young, 2006;Wale et al., 2009;Zhang and Chiew, 2009); a multiscale parameter regionalization technique (Samaniego et al., 2010;Kumar et al., 2013).

Besides, literature review regarding studies on a sequential regionalisation (Lamb et al., 2000;Wagener and Wheater, 2006), a regional calibration approach in regionalization (Hundecha and Bardossy, 2004;Hundecha et al., 2008), a regionalization approach combining an iterative regional calibration (Parajka et al., 2007) will be presented in a next draft. However, a regularization technique (Pokhrel et al., 2008), which was suggested by the reviewer, to reduce problems related to over-parameterization in hydrological modelling is slightly out of the scope of this study.

6. **(RC)** P7061: L7: The author states “The objective of this study was to reduce the parameter uncertainty in regionalization studies . . .. without adding additional parameters or modification of the model structure” is not true. In your refined calibration approach you need to estimate two set of (same) model parameters – one for the wet period and one for the dry period – so essentially you have two times more the unknown parameters. Isn’t the case?

(AC)

The reviewer's comment is right. The words, ' without adding additional parameters', in the text will be removed in a next draft.

7. **(RC)** P7062: L24: As stated if all basins show a high degree of similarity then what do we learn from having so many catchments whose response to precipitation is same? Despite this the author should consider increasing the sample size (i.e. the number of study basins) to get more robust results from the regionalization analysis. Only eleven river basins is not enough – and it may be that you are catching more local behavior than regional ones as is the goal of regionalization studies.

(AC)

The sentence (from P7062 L23 to L25) was used to explain seasonal influences on the historical data such as a strong seasonality in the rainfall and streamflow distributions in Korea. These were inferred from the mean monthly rainfall and streamflow (Fig. 2) and the autocorrelation function for rainfall and streamflow (Fig. 3). In order to give clear meaning on the sentence (from P7062 L23 to L25), additional explanation on the autocorrelation function for rainfall and streamflow in the catchment (Fig. 3) will be added in the next draft (please see the first rejoinder for the RC #10). In addition, the sentence (from P7062 L23 to L25) will be revised as follows: "The catchments in Korea show generally similar seasonal distributions of rainfall and streamflow, with a strong seasonal correlation, where summers are wet and winters are dry."

The possible improvements for the dry period in regionalization were identified and evaluated using 11 catchments located in Korea (the data limitation caused by significant reservoir regulation was introduced in 'Section 2. Study site and data set'). The overall goal of this research is not to derive accurate and reliable regional model equations, which is unlikely to be robust with the small number of catchments used, but to investigate the effectiveness of the regional models derived from the refined calibration approach through comparison with those from the traditional approach. This study focuses on identifying the differences in the relationships between the catchment attributes and the model parameters (calibrated based on the traditional and the refined approaches) and evaluating the adequacy of the regional models (from the traditional and the refined calibration approaches) based on the comparison under the same conditions, rather than evaluating the accuracy of the regional models themselves.

I agree on a lake of robustness in the regional models due to the small number of catchments. The results are, however, sufficient to provide the effectiveness of the regional models from the refined calibration approach when it is compared with that of the regional models derived from the traditional calibration approach under the same conditions. As discussed above, this study is limited to investigating the differences caused by the calibration methods through a regionalisation procedure (i.e., relationships between the catchment attributes and the model parameters in terms of a hydrological perspective; regional models etc). This issue on the lake of robustness in the regional models due to the small number of catchments will be added more in the discussion part (6. Summary and conclusion) in the next draft.

8. **(RC)** P7066: L: 10: What is NIRE and how does this method work? Please provide detail on this?

(AC)

An estimate of the non-parametric impulse response curve (or total unit hydrograph including all flow components) is a useful means of comparing hydrologic response characteristics and their change over time to develop an understanding of the catchment characteristics. The method in this study used a technique for deriving a non-parametric estimate of the impulse response by obtaining an average event unit hydrograph (total unit hydrograph incorporating both quick and slow flow components) response curve directly from observed streamflow (Croke, 2006). The average event unit hydrograph is derived solely from streamflow data (i.e. without using rainfall data).

The method used for estimating the hydrologic response to an average flow event has three steps. Identification of separated flow events having reasonably clean recessions (which are not affected significantly by previous or following flow events) in the hydrograph is the first step. This can be done by scanning for flow peaks in the stream hydrograph series and choosing peaks further apart than some minimum separation from their nearest neighbours. The selected flow peaks are then set to time step zero and the event response profiles are summed after exponential interpolation to one tenth of a time step. The peak value of the summed response curves are normalized to 1. This produces the raw response curve over an interval before and after the peaks. Finally, the procedure is then iterated to give the final profile, with subsequent flow peaks removed using the current estimated.

The definition of an average event unit hygrograph is used instead of the instantaneous unit hydrograph. Also, the peak of derived event response profile is scaled to 1 for simplicity rather than constrained by a unit volume. The derived average event unit hydrograph represents the hydraulic properties of the catchment with the average effective rainfall pattern due to the intensity and duration over the rainfall events selected, not for a pulse of rainfall. Details on the NIREs (discussed above) will be added in a next draft.

9. **(RC)** P7070: L: 14: Why such a selective approach of verifying the regional models is used? What is unique about the selected two basins for verification, and why not to use the established Jacknife type of resampling approach for the verification purpose?

(AC)

The regional models were verified using a split-sampling (i.e., n-2/n development and 2/n validation where n is the number of catchments) in this regionalization study (Wagener and Wheater, 2006) rather than cross-validation such as a Jack-knife resampling.

The two validation catchments were selected based on the criteria on model performance in low flow (i.e., $NSE_{\ln}$) because this study is focused on potential improvements to model behavior during dry periods. The Nam River catchment exhibited the best model performance in $NSE_{\ln}$, while the Seomjin River catchment had the worst value in $NSE_{\ln}$ (see Table **I** below). These catchments were chosen to see how much the regional models with the two approaches (i.e., the traditional and the refined approaches) influenced runoff prediction in the paired catchments having good and bad model performance in low flows. In particular, the Nam River catchment (2,285 km^2) was selected for validation in order to assess whether the regional models derived from the catchment attributes (varying in size between 59 and 1,584 km^2, Table 3 in the manuscript) ensured predictive capability for the catchments outside the prescribed range of the catchment size. The criteria to select the validation catchments will be added in the next draft.

Table **I** Model performance statistics during the calibration period (8yr, 2003~2010)

Catchment	***NSE***	***NSE*$_{ln}$**	***%bias***
Goesan	0.72	0.63	4.90
Hoengseong	0.82	0.40	0.36
Andong	0.78	0.48	8.43
Imha	0.78	0.48	5.49
Hapcheon	0.86	0.59	0.58
Nam River	0.82	**0.72**†	-0.82
Miryang	0.82	0.64	7.78
Yongdam	0.84	0.53	1.43
Buan	0.72	0.39	9.14
Seomjin River	0.80	**0.37**††	-0.02
Juam	0.80	0.57	-0.89

† the highest value in NSE_{ln}
†† the lowest value in NSE_{ln}

10. **(RC)** The presented Figure 3 is not discussed - either consider removing it or discuss it in the manuscript. Besides I find Figures 2 and 3 presenting as separate figure redundant (suggestion merge them together). Figure 7 is redundant considering that all information is given in Table 2. Likewise the information provided in Table 3 can be easily assimilated into Table 1, and then we do not need Table 3 as such.

(AC)

Figure 2 and 3 shows the variability of the monthly rainfalls and streamflows for the catchments. Figure 2 represents seasonal distributions of rainfall and streamflow, where summers are wet and winters are dry. Figure 3 shows the autocorrelation function for rainfall and streamflow in the catchment. The autocorrelation function for rainfall and streamflow was used to show seasonal influences on the historical data. A strong seasonal correlation (i.e., a strong seasonality in the rainfall and streamflow distributions in the catchment) was identified according to the periodic shape in Figure 3. In a next draft, an explanation on Figure 3 will be added and Figures 2 and 3 will also be merged together. Figure 7 will be removed in a next draft, in order to avoid a repetition of the same results. In a next draft, the catchment characteristics in Table 3 will be assimilated into information in Table 1.

References

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

Croke, B. F. W.: A technique for deriving an average event unit hydrograph from streamflow-only data for ephemeral quick-flow-dominant catchments, Advances in Water Resources, 29, 493-502, 2006.

Croke, B. F. W., and Shin, M. J.: Modifications to a rainfall-streamflow model to handle "non-stationarity", Proc. IAHS, 371, 29-33, 10.5194/piahs-371-29-2015, 2015.

Dunn, S. M.: Imposing constraints on parameter values of a conceptual hydrological model using baseflow response, Hydrology and Earth System Sciences, 3, 271-284, 1999.

Gupta, H. V., Sorooshian, S., and Yapo, P. O.: Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information, Water Resources Research, 34, 10.1029/97wr03495, 1998.

Hundecha, Y., and Bardossy, A.: Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model, Journal of Hydrology, 292, 281-295, 2004.

Hundecha, Y., Ouarda, T. B. M. J., and Bardossy, A.: Regional estimation of parameters of a rainfall-runoff model at ungauged watersheds using the spatial structures of the parameters within a canonical physiographic-climatic space, Water Resources Research, 44, W01427, 10.1029/2006wr005439, 2008.

Jakeman, A. J., and Hornberger, G. M.: How much complexity is warranted in a rainfall-runoff model?, Water Resources Research, 29, 2637-2649, 1993.

Kay, A. L., Jones, D. A., Crooks, S. M., Calver, A., and Reynard, N. S.: A comparison of three approaches to spatial generalization of rainfall–runoff models, Hydrological Processes, 20, 3953-3973, 10.1002/hyp.6550, 2006.

Kim, H. S., Croke, B. F. W., and Jakeman, A. J.: Model-based analysis to identify the impacts of climate variability and land use changes on streamflow, American Geophysical Union, Fall Meeting 2008, San Francisco, USA, 2008, Kim, H. S., Croke, B. F. W., Jakeman, A. J., and Chiew, F. H. S.: An assessment of modelling capacity to identify the impacts of climate variability on catchment hydrology, Mathematics and Computers in Simulation, 81, 1419-1429, 2011.

Kim, H. S., and Lee, S.: Assessment of a seasonal calibration technique using multiple objectives in rainfall–runoff analysis, Hydrological Processes, 28, 2159-2173, 10.1002/hyp.9785, 2014a.

Kim, H. S., and Lee, S.: Assessment of the adequacy of the regional relationships between catchment attributes and catchment response dynamics, calibrated by a multi-objective approach, Hydrological Processes, 28, 4023-4041, 10.1002/hyp.9942, 2014b.

Klemes, V.: Operational testing of hydrological simulation models, Hydrological Sciences Journal 31, 13-24, 1986.

Koivusalo, H., and Kokkonen, T.: Modelling runoff generation in a forested catchment in southern Finland, Hydrological Processes, 17, 313-328, 2003.

Kumar, R., Samaniego, L., and Attinger, S.: Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations, Water Resources Research, 49, 360-379, 10.1029/2012wr012195, 2013.

Lamb, R., Crewett, J., and Calver, A.: Relating hydrological model parameters and catchment properties to estimate flood frequencies from simulated river flows, BHS 7th National Hydrology Symposium, Newcastle-upon-Tyne, U. K., 2000.

McIntyre, N., Lee, H., Wheater, H., Young, A., and Wagener, T.: Ensemble predictions of runoff in ungauged catchments, Water Resources Research, 41, W12434, 10.1029/2005wr004289, 2005.

Merz, R., and Blöschl, G.: Regionalisation of catchment model parameters, Journal of Hydrology, 287, 95-123, 2004.

Oudin, L., Andréassian, V., Perrin, C., Michel, C., and Le Moine, N.: Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments, Water Resources Research, 44, W03413, 10.1029/2007wr006240, 2008.

Parajka, J., Merz, R., and Blöschl, G.: A comparison of regionalisation methods for catchment model parameters, Hydrology and Earth System Sciences Discussions, 2, 509-542, 10.5194/hessd-2-509-2005, 2005.

Parajka, J., Blöschl, G., and Merz, R.: Regional calibration of catchment models: Potential for ungauged catchments, Water Resources Research, 43, W06406, 10.1029/2006wr005271, 2007.

Pokhrel, P., Gupta, H. V., and Wagener, T.: A spatial regularization approach to parameter estimation for a distributed watershed model, Water Resources Research, 44, W12419, 10.1029/2007wr006615, 2008.

Samaniego, L., Kumar, R., and Attinger, S. C. W.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, Water Resources Research, 46, n/a-n/a, 10.1029/2008wr007327, 2010.

Seibert, J., and McDonnell, J. J.: On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, Water Resources Research, 38, Art. no. -1241, 2002.

Uhlenbrook, S., and Sieber, A.: On the value of experimental data to reduce the prediction uncertainty of a process-oriented catchment model, Environmental Modelling and Software, 20, 19-32, 2005.

Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., and Sorooshian, S.: Effective and efficient algorithm for multiobjective optimization of hydrologic models, Water Resources Research, 39, 1214, 10.1029/2002wr001746, 2003.

Wagener, T., and Wheater, H. S.: Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty, Journal of Hydrology, 320, 132-154, 2006.

Wale, A., Rientjes, T. H. M., Gieske, A. S. M., and Getachew, H. A.: Ungauged catchment contributions to Lake Tana's water balance, Hydrological Processes, 23, 3682-3693, 10.1002/hyp.7284, 2009.

Yadav, M., Wagener, T., and Gupta, H.: Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins, Advances in Water Resources, 30, 1756-1774, 2007.

Yapo, P. O., Gupta, H. V., and Sorooshian, S.: Multi-objective global optimization for hydrologic models, Journal of Hydrology, 204, 83-97, 1998.

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

Zhang, Y., and Chiew, F. H. S.: Relative merits of different methods for runoff predictions in ungauged catchments, Water Resources Research, 45, W07412, 10.1029/2008wr007504, 2009.