Flood and Drought Hydrologic Monitoring: The Role of Model Parameter Uncertainty

*Nathaniel W. Chaney[1], Jonathan D. Herman[2], Patrick M. Reed[2] and Eric F. Wood[1]

Submitted to Hydrology and Earth System Sciences

21 May 2015

[1]Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ, USA

[2]School of Civil and Environmental Engineering, Cornell University, Ithaca, NY, USA

*Corresponding Author:
Nathaniel W. Chaney
Department of Civil and Environmental Engineering
Princeton University
Princeton, NJ 08544
nchaney@princeton.edu

Key Points:

- Identifying model parameters at coarse time scales impacts the predictability of extreme hydrologic events
- Model parameter sensitivity varies as a function of time scale and region
- Drought and flood monitoring systems must account for model parameter uncertainty.

Abstract

Land surface modeling, in conjunction with numerical weather forecasting and satellite remote sensing, is playing an increasing role in global monitoring and prediction of extreme hydrologic events (i.e., floods and droughts). However, uncertainties in the meteorological forcings, model structure, and parameter identifiability limit the reliability of model predictions. This study focuses on the latter by assessing two potential weaknesses that emerge due to limitations in our global runoff observations: (1) the limits of identifying model parameters at coarser time scales than those at which the extreme events occur, and (2) the negative impacts of not properly accounting for model parameter equifinality in the predictions of extreme events. To address these challenges, petascale parallel computing is used to perform the first global-scale, 10,000 member ensemble-based evaluation of plausible model parameters using the VIC (Variable Infiltration Capacity) land surface model, aiming to characterize the impact of parameter identifiability on the uncertainty in flood and drought predictions. Additionally, VIC's global-scale parametric sensitivities are assessed at the annual, monthly, and daily timescales to determine whether coarse-timescale observations can properly constrain extreme events. Global and climate type results indicate that parameter uncertainty remains an important concern for predicting extreme events even after applying monthly and annual constraints to the ensemble, suggesting a need for improved prior distributions of the model parameters as well as improved observations. This study contributes a comprehensive evaluation of land surface modeling for global flood and drought monitoring and suggests paths forward to overcome the challenges posed by parameter uncertainty.

1. Introduction

Droughts and floods can have devastating consequences on ecosystems, food supply, and economies (Easterling et al., 2000). Providing real-time information and predictions to decision makers can be a valuable tool to mitigate their effects. This is an especially challenging task over data sparse regions, where unreliable monitoring networks and generally low institutional capacity limits the spread of timely information (Sheffield et al., 2013). State-of-the-art land surface models, in conjunction with numerical weather forecasting and satellite remote sensing, pose a plausible solution to supplement local observation networks. Given the accessibility of these data sources, multiple systems have arisen over the past decade that aim to provide continental and global monitoring and predictions of the hydrologic cycle (Sheffield et al., 2013;Vogt et al., 2011;Svoboda et al., 2002;Verdin et al., 2005).

The land surface model component of a monitoring system is useful to understand the impact of flood and drought on the energy, carbon, and hydrologic cycles. This is possible with the current generation of LSMs that include the main physical, biological, and chemical processes at the land surface (Niu et al., 2011). The increasing complexity and sophistication of land surface models can provide a more complete assessment of the state of the land surface but also requires an increase in the number of process parameterizations and model parameters. In the past, parameter estimation in land surface models consisted of using look-up tables to assign model parameters based on similarity between sites as a function of soil and vegetation. However, sensitivity analysis of macroscale land surface models suggests that this is overly simplistic and can lead to significant uncertainty (Rosero et al., 2010;Hou et al., 2012).

Parameter calibration, a common practice in hydrology, can help reduce model bias, understand model deficiencies, and increase the model's reliability (Harding et al., 2014;Cibin et al., 2010;Döll et al., 2003;Sheffield et al., 2013). However, optimizing model performance to a limited set of observations does not ensure the model is getting the right answer for the right reasons (Kirchner, 2006). Instead, there tend to be multiple parameter sets that satisfy the observations; in hydrology this is known as model parameter equifinality (Beven, 2006). Although the performance might be similar for a given calibration metric, the results can vary significantly when comparing other metrics, time scales, or variables (Gupta et al., 2008;Herman et al., 2013;Wagener and Gupta, 2005;Reusser and Zehe, 2011;Reusser et al., 2009;Clark and Vrugt, 2006).

The model equifinality hypothesis is especially relevant in global land surface modeling where the sparsity of observations in space and time and the increasing number of model parameters leads to heavily underconstrained parameter estimation. In this study, we use an ensemble of behavioral parameter sets to capture the spread in simulated energy and water cycles. This improves model evaluation by enabling a comprehensive assessment of the model parameter and model structure deficiencies (Pappenberger and Beven, 2006). A growing number of hydrologic monitoring systems already include the impact of uncertainty in meteorological forcing (Cloke and Pappenberger, 2009); this should be extended to include model parameter uncertainty.

Given the significant number of model parameters in existing global land surface models, carefully designed sensitivity analysis can help minimize the number of uncertain parameters that must be explored for effective model evaluations while reducing computational demands. Up to now, there have only been a limited number of sensitivity analyses of macroscale land surface models. These studies have shown that parameter sensitivity varies with climate, soil, and vegetation properties (Liang and Guo, 2003;Rosero et al., 2010). In the hydrologic cycle, evidence suggests that the runoff partitioning (i.e., between baseflow and surface runoff) plays a dominant role in daily flow estimates over a number of climates (Demaria et al., 2007). The baseflow generation model parameters can also play an important role in the seasonality of the land surface fluxes (Hou et al., 2012). However, questions remain regarding the applicability of these studies globally, suggesting the need for similar analyses over all global land area.

In this study, we accomplish this goal by performing a comprehensive sensitivity analysis of the global VIC (Variable Infiltration Capacity, Liang et al. (1996)) macroscale land surface model. A Latin Hypercube Sample of 10,000 parameter sets is used to run the model from 1948-2010 per 1.0 degree land grid cell over the globe. The GRDC (Global Runoff Data Centre) monthly climatology of gridded runoff observations (Fekete et al., 2002) is used to isolate the behavioral parameter sets. The constrained ensemble is then used to understand: first, the consequence of identifying model parameters at coarser time scales than those at which the extreme events occur, second, the impact of not properly accounting for model parameter equifinality in the estimates of extreme events, and third, the model parameters that control the hydrologic processes at the annual, monthly, and daily timescales. Finally, the results are used to propose paths to provide reliable uncertainty estimates and suggest processes and parameters that require improved observations and parameterizations.

2. Data

2.1 Meteorology: Princeton Global Forcing Dataset

The meteorological forcing dataset consists of 3-hourly, 1.0-degree resolution fields of near-surface meteorology for global land areas for 1948-2010 (PGF; Sheffield et al. (2006)). The dataset merges data from the NCEP-NCAR reanalysis (National Center for Environmental Prediction and National Center for Atmospheric Research; Kalnay et al. (1996)) with the GPCP

(Global Precipitation Climatology Project; Adler et al. (2003)) and TMPA (TRMM Multi-Satellite Precipitation Analysis; Huffman et al. (2007)) observation-based datasets of precipitation, temperature from CRU (Climatic Research Unit; New et al. (2000); Harris et al. (2013)), and radiation from SRB (Surface Radiation Budget; Stackhouse et al. (2004)). For the simulations, we use precipitation, temperature, pressure, downward shortwave and longwave radiation, specific humidity, and wind speed.

2.2 Land Data

The default model soil and vegetation parameters are the same as those described in Sheffield and Wood (2007). The global soil texture comes from the 5-min FAO–UNESCO (Food and Agricultural Organization–United Nations Educational, Scientific, and Cultural Organization) digital soil map of the world and the World Inventory of Soil Emission Potentials (WISE) pedon database (Batjes, 1995). Land cover information is given by the University of Maryland land cover type dataset (Defries et al., 2000). The parameters for each land cover type are assigned using the sources described in Nijssen et al. (2001). The monthly climatology of leaf area index is based on Myneni et al. (1997). The baseline parameters for the land surface model come from these datasets.

2.3 Gridded Runoff Observations: GRDC Climatology

The observations of global gridded runoff come from the GRDC global runoff climatology (Fekete et al., 2002). The dataset provides the interstation observations at 663 stream gauges. To minimize river routing uncertainty, stream gauges are only used when the interstation area between two gauges is below 1 million squared kilometers and less than 10% of the grid cells have a travel time to the gauge above 10 days (assuming a fixed flow velocity of 1 m/s). The gridded estimates are obtained by spatially disaggregating the observed interstation area runoff using the VIC model ensemble. Following the work of Fekete et al. (2002), we assume that the simulations of the land surface model provide the true spatial heterogeneity at the monthly scale. The observed monthly climatology is then used to bias-correct each cell's ensemble mean of simulated monthly flow. Uncertainty in the observed monthly flow is assumed to be negligible relative to the impact of parameter uncertainty. Further details on the model ensemble will be given in section 3.2.

2.4 Köppen-Geiger Climate

The Köppen-Geiger climate classification is used to assess how model parameter sensitivity varies across climates. This dataset divides the world into five different climates based on five vegetation groups. The second and third categories consider precipitation and air temperature. The most recent version of this dataset was updated in 2006 using the CRU (Climatic Research Unit) and GPCC (Global Precipitation Climatology Centre) datasets. These updates make the dataset suitable for the second half of the 20th century (Kottek et al., 2006). In this study only the 5 general climate groups are used: Tropical, Arid, Temperate, Continental, and Polar.

3. Methodology

3.1 VIC: Land Surface Hydrologic Model

The macroscale VIC land surface hydrologic model (Liang et al., 1996) simulates the land surface hydrologic and energy cycles. The model's sub-grid heterogeneity is parameterized using the variable infiltration capacity curve and tiling of land cover classes. Baseflow is modeled as a nonlinear recession from the lowest soil layer (Dumenil and Todini, 1992) and evapotranspiration is calculated using Penman-Monteith (Monteith, 1964). The subsurface is discretized into multiple soil layers; gravity drainage models the movement of moisture between the soil layers. The model captures cold land processes through snow pack storage, frozen soils, and sub-grid distribution of snow based on elevation banding. For further details see Sheffield and Wood (2007).

3.2 Model Parameter Uncertainty: Latin Hypercube Sample

Samples of the model parameter space are obtained using a Latin Hypercube Sample of size 10,000. LHS is used due to its strength to properly sample the parameters by dividing the parameter space into regions of equal probability (McKay et al., 1979). Since this study focuses on the hydrologic cycle, we focus on sampling parameters that contribute to runoff generation. Seven of the nine chosen parameters come from Troy et al. (2008). A multiplier of the tabular minimum stomatal resistance values is added due to its potential impact on the partitioning of runoff and evaporation. Table 1 shows each parameter's name, description, units, and range. Each parameter is drawn from a uniform distribution; parameters that cover 2 or more orders of magnitude are sampled in $\log_{10}$ space. For each LHS parameter set, the model is run at a 3-hour time step between January, 1948 and December, 2010 with a 10 year spin up period. Parameter values are assumed to be uncorrelated in space. The 10,000 ensemble members are run for all 1.0 degree land grid cells over the globe excluding Greenland and Antarctica (15836 grid cells in total).

To assess how well the model can reproduce observed runoff, a set of annual and monthly thresholds are used to obtain each grid cell's behavioral parameter sets. The 10,000 LHS ensemble is constrained using the 1.0-degree observed gridded runoff. The relative error of the simulated annual runoff is used as a first constraint. For each grid cell, all parameter sets that lead to a relative error in annual mean runoff above 10% are discarded. This threshold is set relatively high due to measurement uncertainties in the observation dataset and the spatial disaggregation method described in section 2.3. The second constraint attempts to find all ensemble members that also follow the observations' seasonality. The simulated and observed monthly runoff climatologies are normalized (to remove remaining annual biases) and the Pearson correlation between the observations and simulations is computed. The correlation threshold is set to 0.75. This threshold is set relatively low due to incomplete accounting of the effects of river routing in the observations and simulations. Ensemble members satisfying both the annual and monthly constraints are deemed behavioral, and the posterior distributions of behavioral parameter values are used to assess parameter sensitivity.

3.2 Model Parameter Sensitivity

Quantifying the role of each model parameter at different time scales can help discern the parameters (and processes) that can be constrained using coarse time scale observations (e.g., annual and monthly flows). It can also inform us about which parameters play an important role at finer time scales (e.g., daily flows) and are minimally impacted by coarse timescale constraints.

3.2.1 Parameter Space Reduction: Annual and Monthly Flows

Beyond quantifying how many parameter sets of the 10,000 member ensemble satisfy the monthly and annual constraints, we aim to understand how the reduction in bias and increase in monthly skill is related to a location's climate. To accomplish this goal, the annual flows are analyzed by determining the change in runoff ensemble mean after applying the constraints. Furthermore, since the monthly constraint attempts to improve the simulation's unbiased seasonality, it effectively aims to capture the temporal smoothness of the observed climatology. This effect is quantified by analyzing the change in the 1-month lag autocorrelation.

Our computation of parameter sensitivities after applying the annual and monthly constraints – summarized in Figure 1 – follows the work of Fenwick et al. (2014). For each grid cell, the area between each parameter's prior cumulative distribution function and the posterior cumulative distribution function is computed as:

$$D_{CDF} = \int_{x_l}^{x_u} |F(x) - G(x)|\, dx$$

Where x_l and x_u are the lower and upper bounds of the parameter in question, which are normalized to [0,1] to improve interpretability of the result. The integrals are computed numerically using the trapezoid rule with $\Delta x = 0.01$. The calculated area serves as a robust sensitivity metric indicating the change in the distribution of each parameter caused by applying the performance constraints. Because the prior parameter distributions in this study are uniform, the maximum value of this metric is 0.5 (i.e., if only a single ensemble member satisfies the performance constraints and remains in the posterior distribution). This "CDF Distance" sensitivity method bridges the classical Regional Sensitivity Analysis framework (Spear and Hornberger, 1980) and the Delta Moment-Independent Measure (Plischke et al., 2013;Borgonovo, 2007). Regional Sensitivity Analysis employs the maximum difference between cumulative distributions as a sensitivity measure. The Delta Moment-Independent Measure (Plischke et al., 2013;Borgonovo, 2007) uses the area between prior and posterior PDFs rather than CDFs. We compute two CDF distances: first, between the original uniform distribution and the posterior after applying the annual constraint (below 10% absolute error), and second, between the posterior after the annual constraint and the posterior after applying the additional monthly constraint ($r > 0.75$). The advantages of the CDF Distance method for this study are (1) it does not require special statistical sampling and will work for the given data, and (2) it ties parameter sensitivity to a model performance threshold to identify parameters responsible for a particular outcome rather than overall changes in the output.

3.2.2 Parameter Uncertainty: Daily Flows

Reducing the annual and monthly model parameter uncertainty using the GRDC monthly climatology does not ensure a similar reduction in the uncertainty of daily flows. This is especially relevant to drought and flood monitoring systems that attempt to capture the sub-monthly hydrologic extremes over data sparse regions. If the most sensitive parameters at the daily scale are also the most sensitive parameters at the annual and monthly time scales, then there should be a substantial decrease in uncertainty. However, if the parameter sensitivity at different time scales is orthogonal, then the reduction in uncertainty at the daily scale will be negligible. To address this question, for each grid cell, the daily flow duration curves of the full ensemble (10,000 members) and behavioral parameter sets are calculated. The changes in the spread at different sections (low, median, and high flows) of the flow duration curve are analyzed.

Given that uncertainty will persist in the daily flows after applying the constraints, the question remains about which parameters control the remaining ensemble spread and need to be more heavily constrained. This is done by analyzing the spread in daily flow extremes on both sides of the distribution (1^{st} and 99^{th} percentiles) for the strictest annual and monthly constraints (relative error below 10% and monthly correlation above 0.75). For each percentile, the Spearman rank correlation between all behavioral parameters and their associated flow is computed. The Spearman correlation was chosen here because (1) observations of daily flows are not available, so behavioral parameters cannot be identified as with the CDF distance measure described in Section 3.2.1, and (2) in general the relationship between parameter values and daily extreme flows will be nonlinear. The Spearman correlation provides a metric describing how a given parameter controls the spread in daily flows, which may have been underconstrained by the annual and monthly performance requirements imposed in the previous step. This is done for each of the 9 parameters.

4. Results

4.1 VIC Latin Hypercube Sample & Behavioral Parameter Space Reduction

For each land 1.0 degree grid cell (15,836) the VIC (Variable Infiltration Capacity) land surface model is run between 1948 and 2010 at a 3-hour temporal resolution for 10,000 parameter sets obtained from a Latin Hypercube Sample. This was possible due to the Blue Waters supercomputer (http://www.ncsa.illinois.edu/enabling/bluewaters); the simulations required more than 2 million computing hours (> 200 years) and resulted in an output of over 1.5 petabytes. The data was then summarized into daily, monthly, and yearly datasets. Each grid cell's 10,000 LHS ensemble VIC simulations are constrained using the observed gridded runoff fields described in section 2.3.

Figure 2 shows global maps of the fraction of parameter sets that fulfill each error criterion. In the northern hemisphere, a considerable number of grid cells have a large fraction of ensemble members that are below 10 and 20 percent relative error, suggesting a small annual bias in the input meteorological forcing and a diminished sensitivity to the parameters that impact the annual mean runoff. In many places, there is a sharp decrease in performance when constraining the ensemble with the normalized monthly climatology. This can most likely be attributed to the role that the parameter space plays in controlling runoff partitioning and the challenges when attempting to spatially disaggregate point runoff observations. However, the most prominent feature is the lack of runoff observations (grey areas) and behavioral parameter sets (pink areas) over arid regions and countries with limited adaptation capacity throughout the globe.

Figure 3 further summarizes these results as a function of climate classification. Although most of the regions with observations meet the annual constraints (10 and 20 percent relative error), there are distinct differences between climates. Tropical and dry climates see the smallest proportion of behavioral parameter sets while continental, polar, and temperate regions experience the largest. The number of behavioral parameter sets decreases even further for all climate types when applying the monthly constraint (Pearson correlation between the simulated and observed normalized monthly climatology). In the case of arid regions, the number of acceptable parameter sets is significantly smaller, especially for the North American mountain west, the Sahel, and most of Australia.

Figure 3 also shows how the change in behavioral parameter sets affects the climate averaged runoff ensemble mean and 1-month lag autocorrelation. The first annual constraint (20 percent relative error) leads to a decrease in annual runoff (increase in evaporation) in tropical, dry, temperate, and continental climates; there is an increase in annual runoff in polar climates. The changes in annual flows are negligible when applying the monthly constraints (explained by the normalization of the monthly runoff). The 1-month lag correlation is used as a smoothness metric to assess the impact of the chosen constraints on the simulated seasonality; a higher autocorrelation indicates smoother monthly flows. In all cases, the constraints increase smoothness. As expected, the largest changes occur when using the Pearson correlation as a constraint (increase in accuracy of seasonality of monthly runoff).

In the context of drought and flood monitoring, these results may have key implications. These include: 1) the large fraction of landmass without observations limits our ability to constrain the model parameter space over the globe; 2) a limited number of behavioral parameter sets over arid and regions with limited adaptation capacity - focus areas for monitoring systems -

suggests considerable limitations in monitoring systems as well as the potential for significant model structural errors; 3) regions with a high fraction of behavioral parameter sets will be susceptible to the impact of model parameter equifinality.

4.2 Model Parameter Sensitivity

4.2.1 Parameter Space Reduction: Annual and Monthly Flows

We formalize the sensitivity analysis by examining the cumulative distribution function (CDF) distance between each parameter's prior and posterior distributions. Figure 4 shows the global maps of the CDF distance metric for each parameter after applying the annual and monthly constraints. The color scale of Figure 4 ranges from 0.0, where the prior and posterior distributions match exactly, to 0.5, the maximum possible value of the CDF distance metric when the posterior distribution contains only a single ensemble member. In general, B, Ds_{max}, Exp, and C_{Rsmin} are the most sensitive parameters to the annual constraint (left panel). However, the sensitivity of C_{Rsmin} dominates the other parameters. Since C_{Rsmin} constrains the maximum transpiration rate in the model, these results suggest that the partitioning of evaporation and runoff dominates the model performance at the annual scale. Similarly, Figure 5 shows the mean CDF distance metric within each climate classification, with the interquartile range denoted by error bars. For the annual constraint, the sensitivities of B, Ds_{max}, and Exp are highest in regions with less defined seasonal cycles (e.g. Tropical). As will be discussed in the next section, this can likely be attributed to these parameters playing a distinct role in runoff seasonality.

When applying the monthly constraint, the sensitivity of most parameters changes. In Figures 4 and 5, the negligible sensitivity of C_{Rsmin} suggests that although it plays a fundamental role in ensuring the annual runoff ratio, it does not play an important role in the seasonality; the same applies to Exp. Instead, the most sensitive parameters are B and Ds_{max} since they control the partitioning of runoff into baseflow and surface runoff. As shown in Figure 5, this is especially true over regions with a characteristic seasonal cycle (e.g., continental climates). Regions that lack a distinct seasonality (e.g., tropical climates) are only sensitive to these parameters at annual time scales. When there exists a strong seasonality in runoff, these parameters can impact the seasonality at the monthly timescale. However, a weaker seasonality leads these parameters to act at an annual scale by controlling the soil water storage and therefore the partitioning of annual evaporation and runoff.

The contrast of the annual and monthly results brings to light the role that time scales can have on the sensitivity of model parameters (and, by extension, processes). The results suggest that the annual scale constraint does not play a large role in the partitioning of monthly baseflow and surface runoff. As will be discussed in the following section, these timescale dependent changes in parameter sensitivity can have large implications on the ability to simulate daily flows without daily observations to further constrain the ensemble.

4.2.2 Parameter Uncertainty: Daily Flows

The annual and monthly performance constraints allow us to explore the role of the remaining parameter uncertainty on daily runoff estimates. The runoff percentiles are calculated for each ensemble member of each grid cell. Figure 6 shows the climate-averaged spread of the flow duration curves of the 10,000 ensemble members and the most heavily constrained ensemble (annual and monthly). The change in spread provides insight into how constraining (or tuning) at coarser time scales can reduce uncertainty at the daily scale.

As expected from Figure 3, the annual and monthly constraints lead to a reduction in the daily mean runoff for all climates (except polar). However, the constraints' ability to tighten the ensemble spread varies significantly among climates. The most substantial decrease occurs over continental and polar climates even though these regions experience the lowest decrease in the number of parameter sets (see Figure 3). This decrease is most likely connected to the results from the monthly sensitivity analysis (see section 4.2.1): over regions that have a distinct seasonal cycle, the monthly climatology is able to heavily constrain the B and Ds_{max} parameters; this then helps constrain runoff at daily time scales. This also explains the small decrease in spread over tropical climates seen in Figure 6; since the monthly constraints are not able to constrain the B and Ds_{max} parameters, their uncertainty drives the runoff at daily time scales. While predictions in tropical climates are not well constrained with this approach, the results are encouraging for monitoring the hydrologic cycle with properly-constrained land surface models in continental and polar climates.

Figure 6 also illustrates differences in the tightening of the flow duration curve spread at different percentiles. For example, in continental climates the percentiles close to the center experience a substantial decrease in spread; the change in the ensemble spread of the tails (hydrologic extremes) is less significant. This result holds to a varying degree for all climates. The most likely physical explanation is that the annual and monthly constraints focus on the percentiles that produce most of the runoff; this leads to a minimal impact on low flows and a reduced impact on high flows. The non-negligible role that high flows play in runoff production helps explain the larger decrease in spread when compared to low flows.

Given that considerable uncertainty remains in the daily flows after applying the annual and monthly constraints, we aim to understand what parameters (and, by extension, processes) control the spread. Figure 7 shows the global Spearman correlations between the daily flow extremes (1^{st} and 99^{th} percentile, in the left and right panels, respectively) and the behavioral parameters. Red indicates a negative correlation, blue indicates a positive correlation, and white indicates no observed correlation. The results in Figure 7 suggest that *B*, Ds_{max}, *Exp*, and C_{Rsmin} control the daily flow extremes, evidenced by a mix of strong positive and negative correlations. The negative correlation between the *B* parameter and low flows occurs because a decrease in *B* leads to an increase in infiltration. This results in a dampened response and an increase in available storage for low flow periods; the opposite is true for high flows. The negative correlation between low flows and Ds_{max} occurs because a decrease in Ds_{max} delays the release of water from storage allowing for a thicker recession curve and higher low flows. Finally, the positive correlation between C_{Rsmin} and high flows is because an increase in C_{Rsmin} leads to a decrease in evaporation; an increase in storage leads to an increase in baseflow and surface runoff (increase in soil saturation). By controlling how quickly the hydraulic conductivity decreases as a function of soil moisture, *Exp* controls water movement between soil layers during dry down periods. This parameter is negatively correlated with low flows since it controls the supply to the lowest soil layer where baseflow is created.

5. Discussion

5.1 Global Flood and Drought Monitoring: Ensemble Simulations

The results from this study are relevant to drought and flood monitoring systems that rely on land surface models to monitor and predict hydrologic extremes at daily time scales (Sheffield et al., 2013;Xia et al., 2012). When the land surface model parameters are not tuned, significant uncertainties exist in the estimated runoff. This is especially true over data sparse regions where the prior estimates of the model parameters are inadequate. Furthermore, when the parameters are tuned, a scale mismatch (space and time) between the observations and the intended application leads to limited improvement. As shown in section 4.2.2, although using annual and monthly observations does constrain the daily estimates near the median, considerable uncertainties remain in the simulated hydrologic extremes (low and high flows) over all Köppen-Geiger climates.

One obvious path forward is to use daily streamflow observations to further constrain the land surface model. This solution is practical over dense stream gauge networks but presents considerable challenges over data sparse regions and ungauged basins. A plausible solution is to use a more sophisticated technique to spatially disaggregate streamflow observations (e.g., Pan and Wood (2013)) to obtain daily gridded runoff fields. However, these methods will continue to struggle over sparse networks (e.g. Congo basin), areas that are heavily managed (e.g., southeast USA), and basins that experience substantial reinfiltration and stream evaporation (e.g., Colorado basin). Another option would be to use satellite based altimetry measurements (e.g., SWOT; Durand et al. (2014)). These observations could be combined with the spatially disaggregated runoff fields to provide the observed daily estimates of gridded runoff.

In any case, even if high quality daily runoff observations existed over the globe, a non-negligible spread will remain after applying the constraints due to the effects of model parameter equifinality. For this reason, we suggest that flood and drought monitoring systems that aim to capture hydrologic extremes move towards model parameter ensemble frameworks to provide not only predictions but also uncertainty estimates. To make this feasible for operational use, further work will need to determine how to cluster the behavioral parameter sets to below 100 per grid cell to minimize the increase in computation and storage requirements.

5.2 Model Parameters: Improve Prior Distributions

A common practice when tuning land surface model parameters at continental scales (e.g. Troy et al. (2008)) is to use the same prior distribution for each model parameter at each modeled grid cell or catchment; this uniform distribution is usually set to cover the entire span of physically plausible parameter values. This approach is one of the main drivers of the large spread in flow duration curves shown in Figure 6. Given the need to rely on monthly and annual observations to constrain the model parameter uncertainty, local prior distributions should be informed by spatial land surface characteristics to constrain the initial ensemble spread and the flow duration curves. Spatially distributed information could also be used to refine the distribution family and shape of the priors in addition to their ranges.

One option would be to use the uncertainty estimates available in remote sensing and in-situ datasets to define the local prior distributions. An example of this framework would be to use the gridded soil survey geographic (gSSURGO) continental soil's product (Soil Survey Staff, 2014) that provides detailed three-dimensional texture and hydraulic soil properties (and uncertainties) over the contiguous United States (CONUS). This would be simple to test for soil parameters that are used in land surface models and are generally reported in soil datasets (i.e., porosity). However, for parameters that are model specific (e.g., Ds_{max} and B in the VIC model), derived functional relationships will need to relate the model parameters to the observed parameters to assemble reliable prior distributions. However, as long as the uncertainties in the functional relationship (e.g., linear regression) inform the derived local prior distribution, the benefits should outweigh additional uncertainties.

A similar option would be to estimate model parameter prior distributions using local information (parameter covariates). The procedure used in this study (Latin Hypercube Sample) could be used over catchments with rich databases to constrain the uniform parameter values using available high spatial and temporal resolution observed data. The resulting behavioral parameter sets could then be related to the local information using machine learning algorithms (e.g., random forests; Liaw and Wiener (2002)) to provide catchment specific prior distributions.

In theory, available or upcoming high-resolution global datasets could then provide the covariates to estimate a parameter's prior distribution at each catchment or grid cell. These datasets could include HydroSHEDS DEM (Lehner et al., 2008), MODIS derived products (e.g., NDVI, albedo, and land cover type), TMPA satellite precipitation (Huffman et al., 2007), and the upcoming GlobalSoilMap (Arrouays et al., 2014), among others. Although the challenges in parameter regionalization (Hrachowitz et al., 2013) in catchment hydrology will also most likely apply to macroscale land surface models, we view it as a path that should be explored.

5.3 Model Structure: Next Generation Land Surface Modeling

Ultimately, more sophisticated parameter estimation techniques cannot fix model structure deficiencies. As the results of section 4.2.1 indicate, if the observed flow is not contained in the constrained ensemble then the problem can be traced to model structure deficiencies (assuming error free observations and input meteorology). This problem is apparent over arid regions (see Figure 3), arguably one of the main regions of focus for drought and flood monitoring systems. A lack of irrigation, reservoirs, river evaporation and reinfiltration, and groundwater in this version of VIC are most likely the drivers of model deficiency. Furthermore, parameterizations that play an important role in watershed dynamics and are highly sensitive to their parameter values (e.g., B in the variable infiltration curve) should be replaced with updated schemes that can effectively use available local high-resolution information (e.g., topography, soils, geology, and land cover) to more accurately represent the local physical processes while reducing reliance on parameter estimation.

The improved macroscale parameterizations to address these process deficiencies should capitalize on the increase in computation resources and available high resolution land data and meteorological data to more explicitly model the fine scale hydrologic processes. (Wood et al., 2011;Bierkens et al., 2014). This effort could provide solutions to improve the prediction of hydrologic extremes over the globe by including: 1) detailed hydrodynamic modeling to account for flash floods, irrigation, reservoirs, and urban flooding; 2) integrated river modeling to enable river evaporation and reinfiltration; 3) improved runoff generation processes. Although the addition of these processes will likely lead to additional parameter complexity and uncertainty, it is seen as a necessary next step to improve the reliability and utility of global drought and flood monitoring systems.

6. Conclusions

The Variable Infiltration Capacity model (VIC) has been run globally at a 1.0 degree spatial resolution between 1948 and 2010 using 10,000 parameter sets from a Latin Hypercube Sample to assess the role of parameter uncertainty in flood and drought monitoring. The 10,000 member ensemble is constrained using a spatially disaggregated version of the GRDC runoff climatology at annual and monthly time scales. A multi-time scale sensitivity analysis is then used to determine the role of each of the model's parameters and the overall model performance. The results vary according to Köppen-Geiger climate. While in arid and tropical regions few parameter sets fulfill the constraints, polar and continental climates maintain a large number of behavioral parameter sets. The annual constraints focus on reducing the annual bias by changing the annual evaporation; the monthly constraints alter the monthly autocorrelation of flow by partitioning the runoff into baseflow and surface runoff. The parameters that control the monthly runoff autocorrelation also play an important role at the daily time scale. For this reason, regions that have a distinct seasonality (continental and polar) see the largest decrease in the spread of their representative daily flow duration curves. These results illustrate the challenges in using current land surface models for global drought and flood monitoring. However, they also indicate a path forward which involves adopting ensemble frameworks to account for model parameter uncertainty, designing and implementing improved observation networks to better constrain land surface models, providing improved local prior distributions via emerging high resolution land data, and improving model structure to better account for the processes that dominate the hydrology over regions prone to droughts and floods.

7. Acknowledgements

This study was supported by funding from NOAA grant NA11OAR4310175 (Improving land evaporative processes and land-atmosphere interactions in the NCEP Global Forecast System (GFS) and Climate Forecast System (CFS), and NSF grant 1144217 (Petascale Design and Management of Satellite Assets to Advance Space Based Earth Science).

8. References

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D. T., Gruber, A., Susskind, J., and Arkin, P. A.: The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present), Journal of Hydrometeorology, 4, 1147-1167, 2003.

Arrouays, D., McKenzie, N., Hempel, J., Richer de Forges, A. R., and McBratney, A. B.: GlobalSoilMap: Basis of the global spatial soil information system, CRC Press, London, UK, 494 pp., 2014.

Batjes, N. H.: A Homogenized Soil Data File for Global Environmental Research: A Subset of FAO, ISRIC, and NRCS profiles (Version 1.0), Working Paper and Prepring 95/10b, International Soil Reference and Information Center, Wageningen, 1995.

Beven, K.: A manifesto for the equifinality thesis, Journal of Hydrology, 320, 18-36, 2006.

Bierkens, M. F. P., Bell, V., Burek, P., Chaney, N. W., Condon, L., Cédric, D., de Roo, A., Döll, P., Drost, N., Famiglietti, J. S., Flörke, M., Gochis, D., Houser, P., Hut, R. W., Keune, J., Kollet, S., Maxwell, R., Reager, J. T., Samaniego, L., Sudicky, E., Sutanudjaja, E. H., van de Giesen, N., Winsemius, H. C., and Wood, E. F.: Hyper-resolution global hydrological modeling: what's next, Hydrological Processes, 29, 310-320, 2014.

Borgonovo, E.: A new uncertainty importance measure, Reliability Engineering & System Safety, 92, 771-784, 10.1016/j.ress.2006.04.015, 2007.

Cibin, R., Sudheer, K. P., and Chaubey, I.: Sensitivity and identifiability of stream flow generation parameters of the SWAT model, Hydrological Processes, 24, 1133-1148, 2010.

Clark, M., and Vrugt, J.: Unraveling uncertainties in hydrologic model calibration: Addressing the problem of compensatory parameters, Geophysical Research Letters, 33, 10.1029/2005GL025604, 2006.

Cloke, H., and Pappenberger, F.: Ensemble flood forecasting: A review, Journal of Hydrology, 375, 613-626, 2009.

Defries, R. S., Hansen, M. C., Townshend, J. R. G., Janetos, A. C., and Lovelands, T. R.: A new global 1-km dataset of percentage tree cover derived from remote sensing, Global Change Biology, 6, 247-254, 2000.

Demaria, E. M., Nijssen, B., and Wagener, T.: Monte Carlo sensitivity analysis of land surface parameters using the Variable Infiltration Capacity model, Journal of Geophysical Research: Atmospheres (1984-2012), 112, 10.1029/2006JD007534, 2007.

Döll, P., Kaspar, F., and Lehner, B.: A global hydrological model for deriving water availability indicators: model tuning and validation, Journal of Hydrology, 270, 105-134, 2003.

Dumenil, L., and Todini, E.: A rainfall-runoff scheme for use in the Hamburg climate model in: Advances in theoretical hydrology, A tribute to James Dooge, edited by: O'kane, P., European Geophysical Society Series on Hydrological Sciences, Elsevier, Amsterdam, 1992.

Durand, M., Neal, J., Rodriguez, E., Andreadis, K. M., Smith, L. C., and Yoon, Y.: Estimating reach-averaged dicharge for the River Severn from measurements of river water surface elevation and slope, Journal of Hydrology, 511, 92-104, 2014.

Easterling, D. R., Meehl, G. A., Parmesan, C., Changnon, S. A., Karl, T. R., and Mearns, L. O.: Climate Extremes: Observations, Modeling, and Impacts, Science, 289, 2068-2074, 2000.

Fekete, B. M., Vörösmarty, C. J., and Grabs, W.: High-resolution fields of global runoff combining observed river discharge and simulated water balances, Global Biogeochemical Cycles, 16, 10.1029/1999GB001254, 2002.

Fenwick, D., Scheidt, C., and Caers, J.: Quantifying Assymetric Parameter Interactions in Sensitivity Analysis: Application to Reservoir Modeling, Mathematical Geosciences, 46, 493-511, 10.1007/s11004-014-9530-5, 2014.

Gupta, H., Wagener, T., and Liu, Y.: Reconciling theory with observations: elements of a diagnostic approach to model evaluation, Hydrological Processes, 22, 3802-3813, 2008.

Harding, R. J., Weedon, G. P., Van Lanen, H. A. J., and Clark, D. B.: The future for global water assessment, Journal of Hydrology, In Press, 2014.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset, International Journal of Climatology, 34, 623-642, 2013.

Herman, J., Kollat, J. B., Reed, P. M., and Wagener, T.: From maps to movies: High resolution time-varying sensitivity analysis for spatially distributed watershed models, Hydrol. Earth Syst. Sci., 17, 5109-5125, 2013.

Hou, Z., Huang, M., Leung, L. R., Lin, G., and Ricciuto, D. M.: Sensitivity of surface flux simulations to hydrologic parameters based on an uncertainty quantification framework applied to the Community Land Model, Journal of Geophysical Research, 117, 10.1029/2012JD017521, 2012.

Hrachowitz, M., Savenjie, H. H. G., Blöschl, G., McDonnel, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P. A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Woods, R. A., Zehe, E., and Cudennec, C.: A decade of Predictions in Ungauged Basins (PUB) - a review, Hydrological Sciences Journal, 58, 1198-1255, 2013.

Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., and Wolff, D. B.: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales, Journal of Hydrometeorology, 8, 38-55, 2007.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetma, A., Reynolds, R., Jeane, R., and Joseph, D.: The NCEP/NCAR 40-Year Reanalysis Project, Bulletin of the American Meteorological Society, 77, 437-471, 1996.

Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, Water Resources Research, 42, W03S04, 10.1029/2005WR004362, 2006.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F.: World Map of the Köppen-Geiger climate classification updated, Meteorologische Zeitschrift, 15, 259-263, 2006.

Lehner, B., Verdin, K., and Jarvis, A.: New global hydrography derived from spaceborne elevation data, EOS, transactions, AGU, 89, 93-94, 2008.

Liang, X., Wood, E. F., and Lettenmaier, D. P.: Surface soil moisture parameterization of the VIC-2L Evaluation and modification, Global and Planetary Change, 13 195-206 1996.

Liang, X., and Guo, J.: Intercomparison of land-surface parameterization schemes: sensitivity of surface energy and water fluxes to model parameters, Journal of Hydrology, 279, 182-209, 2003.

Liaw, A., and Wiener, M.: Classification and Regression by randomForest, R News, 2, 18-22, 2002.

McKay, M. D., Beckman, R. J., and Conover, W. J.: A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics, 21, 239-245, 1979.

Monteith, J. L.: Evaporation and environment. The state and movement of water in living organisms, Symposium of the society of experimental biology, 19, 205-234, 1964.

Myneni, R. B., Nemani, R. R., and Running, S. W.: Estimation of Global Leaf Area Index and Absorbed Par Using Radiative Transfer Models, IEEE Transactions on Geoscience and Remote Sensing, 35, 1380-1393 1997.

New, M., Hulme, M., and Jones, P.: Representing Twentieth-Century Space-Time Climate Variability. Part II: Development of 1901-96 Monthly Grids of Terrestrial Surface Climate, Journal of Climate, 13, 2217-2238, 2000.

Nijssen, B., O'Donnell, G. M., Lettenmaier, D. P., Lohmann, D., and Wood, E. F.: Predicting the Discharge of Global Rivers, Journal of Climate, 14 3307-3323, 2001.

Niu, G.-Y., Yang, Z.-L., Mitchel, K. E., Chen, F., Ek, M., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., and Xia, Y.: The community Noah land surface model with multiparameterization options (NOAH-MP): 1. Model description and evaluation with local-scale measurements, Journal of Geophysical Research, 116, 10.1029/2010JD015139, 2011.

Pan, M., and Wood, E. F.: Inverse streamflow routing, Hydrol. Earth Syst. Sci., 17, 4577-4588, 2013.

Pappenberger, F., and Beven, K. J.: Ignorance is bliss: Or seven reasons not to use uncertainty analysis, Water Resources Research, 42, 10.1029/2005WR004820, 2006.

Plischke, E., Borgnovo, E., and Smith, C. L.: Global sensitivity measures from given data, European Journal of Operational Research, 226, 536-550, 10.1016/j.ejor.2012.11.047, 2013.

Reusser, D., Blume, T., Schaefli, B., and Zehe, E.: Analyzing the temporal dynamics of model performance for hydrologic models, Hydrol. Earth Syst. Sci., 13, 999-1018, 2009.

Reusser, D., and Zehe, E.: Inferring model structural deficits by analyzing temporal dynamics of model performance and parameter sensitivity, Water Resources Research, 47, W07550, 10.1029/2010WR009946, 2011.

Rosero, E., Yang, Z., Wagener, T., Gulden, L. E., Yatheendradas, S., and Niu, G.: Quantifying parameter sensitivity, interaction, and transferability in hydrologically enhanced versions of the Noah land surface model over transition zones during the warm season, Journal of Geophysical Research: Atmospheres (1984-2012), 115, D03106, 10.1029/2009JD012035, 2010.

Sheffield, J., Goteti, G., and Wood, E. F.: Development of a 50-Year High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling, Journal of Climate, 19, 3088-3111, 2006.

Sheffield, J., and Wood, E. F.: Characteristics of global and regional drought, 1950-2000: Analysis of soil moisture data from off-line simulation of the terretrial hydrologic cycle, Journal of Geophysical Research, 112, 10.1029/2006JD008288, 2007.

Sheffield, J., Wood, E. F., Chaney, N., Sadri, S., Guan, K., Yuan, X., Olang, L., Amani, A., and Ali, A.: A Drought Monitoring and Forecasting System for Sub-Saharn African Water Resources and Food Security, Bull. Amer. Meteor. Soc., 95, 861-882, 2013.

Soil Survey Staff: Soil Survey Staff. Gridded Soil Survey Geographic (gSSURGO) Database for the Conterminous United States. United States Department of Agriculture, Natural Resources Conservation Service. Available online at http://datagateway.nrcs.usda.gov/. 8, 15, 2014 (FY2014 official release). 2014.

Spear, R., and Hornberger, G. M.: Eutrophication in peel inlet-II. Identification of critical uncertainties via generalized sensitivity analysis, Water Research, 14, 43-49, 10.1016/0043-1354(80)90040-8, 1980.

Stackhouse, P. W., Gupta, S. K., Cox, S. J., Mikowitz, J. C., Zhang, T., and M.Chiacchio: 12-year surface radiation budget dataset, GEWEX News, 14, 10-12, 2004.

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Tinker, R., Palecki, M., Stooksbury, D., Miskus, D., and Stephens, S.: The Drought Monitor, BAMS, 83, 1181-1190, 2002.

Troy, T. J., Wood, E. F., and Sheffield, J.: An efficient calibration method for continental-scale land surface modeling, Water Resources Research, 44, 10.1029/2007WR006513, 2008.

Verdin, J., Funk, C., Senay, G., and Choularton, R.: Climate science and famine early warning, Phil. Trans. R. Soc. B., 360, 2155-2168, 2005.

Vogt, J. V., Barbosa, P., Hofer, B., Magni, D., Jager, A. D., Singleton, A., Horion, S., Sepulcre, G., Micale, F., and Sokolova, E.: Developing a European drought observatory for monitoring, assessing and forecasting droughts across the European continent, AGU Fall Meeting Abstracts, San Francisco, CA, 2011, 2011.

Wagener, T., and Gupta, H. V.: Model Identification for hydrological forecasting under uncertainty, Stochastic Environmental Research and Risk Assessment, 19, 378-387, 2005.

Wood, E. F., Roundy, J. K., Troy, T. J., van Beek, L. P. H., Bierkens, M. F. P., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffe, P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., and Whitehead, P.: Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water, Water Resources Research, 47, 10.1029/2010WR010090, 2011.

Xia, Y., Mitchel, K. E., Ek, M., Cosgrove, B., Sheffield, J., Luo, L., Alonge, C., Wei, H., Meng, J., Livneh, B., Duan, J., and Lohmann, D.: Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow, Journal of Geophysical Research, 117, 10.1029/2011JD016051, 2012.

Table 1. Range of VIC parameters used in the 10,000 Latin Hypercube Sample. Each parameter is drawn from a uniform distribution; parameters that cover 2 or more orders of magnitude are sampled in $\log_{10}$ space.

Parameter	Units	Range	Description
B	-	0.001-1.0	Variable Infiltration Curve parameter
Ds	-	0.001-1.0	Fraction of Ds_{max} where non-linear flow begins
Ds_{max}	mm/d	0.1-50.0	Maximum baseflow velocity
Ws	-	0.2-1.0	Fraction of Ws_{max} where non-linear flow begins
Layer 2	m	0.1-3.0	Depth of Layer 2
Layer 3	m	0.1-3.0	Depth of Layer 3
Exp	-	0.1-30.0	Characterizing the variation in K_{sat} with soil moisture
C_{Rsmin}	-	0.1-10.0	Multiplier of tabular minimum stomatal resistance values
K_{sat}	mm/d	100-10000	Saturated Hydraulic Conductivity

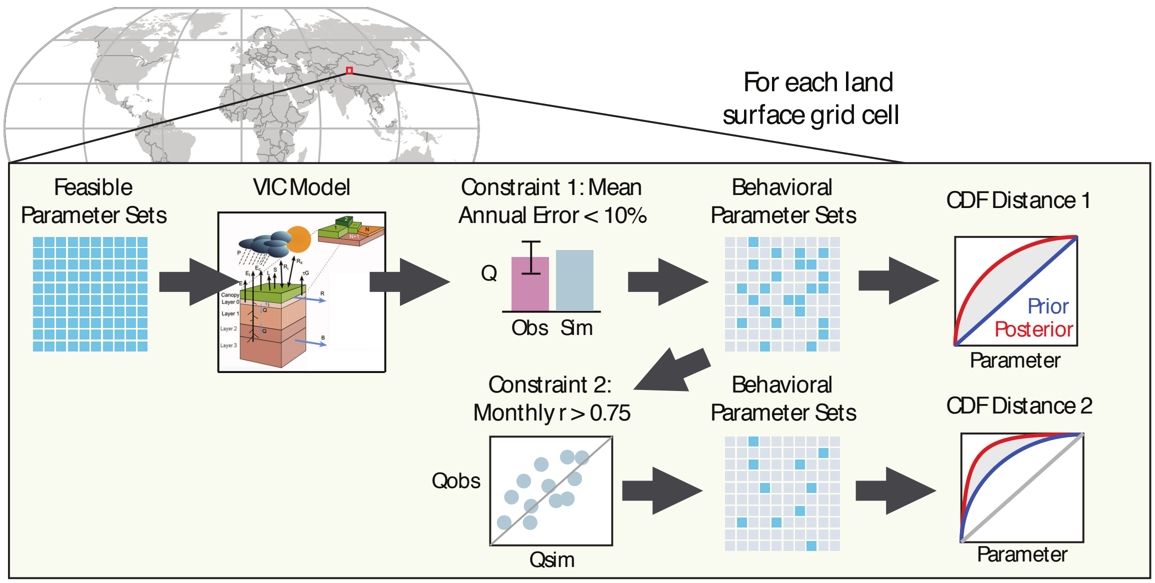

Figure 1. Steps used to build and constrain the 10,000 Latin Hypercube VIC ensemble. The CDF distance is calculated for each VIC parameter after applying the annual error constraint and again after applying the monthly correlation constraint.

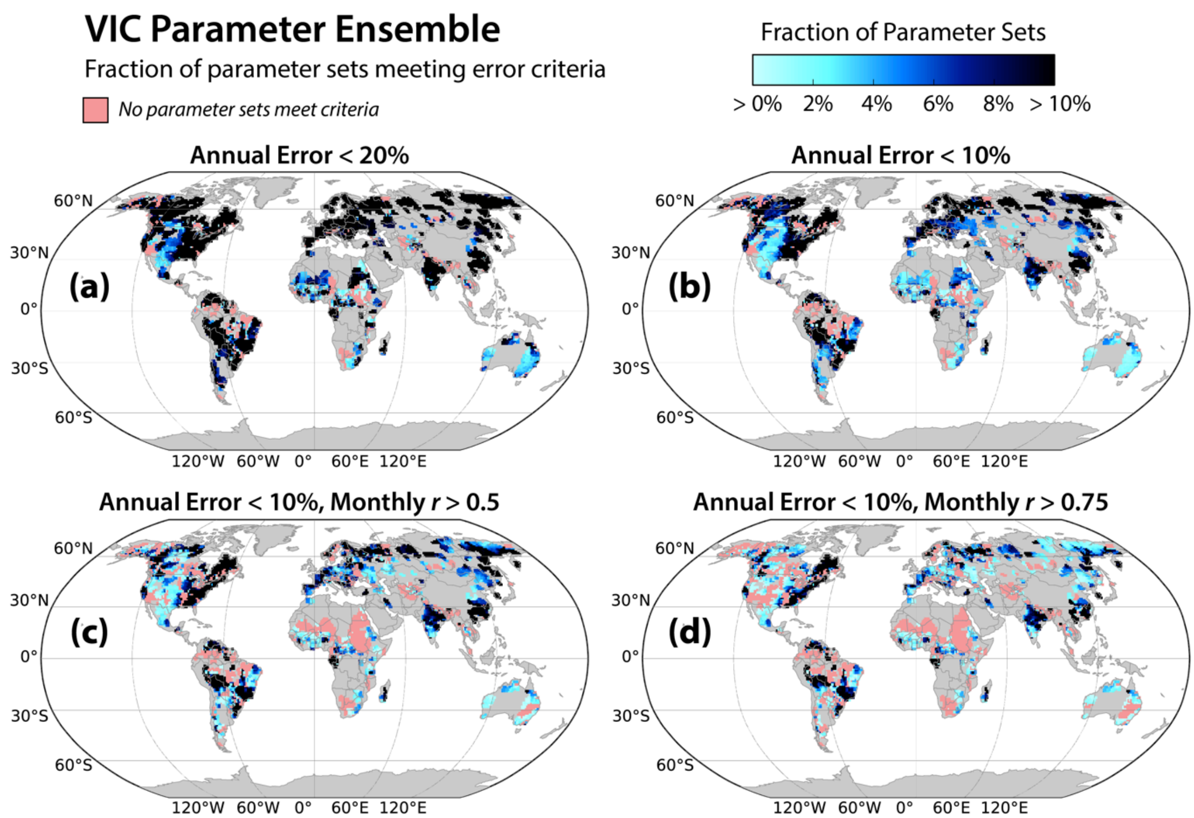

Figure 2. Fraction of parameter sets from the 10,000 Latin Hypercube VIC ensemble that fulfill a set of criteria. The comparison is between the annual and monthly climatology of simulated runoff and the GRDC database. The grey areas are regions that are not covered by the GRDC database.

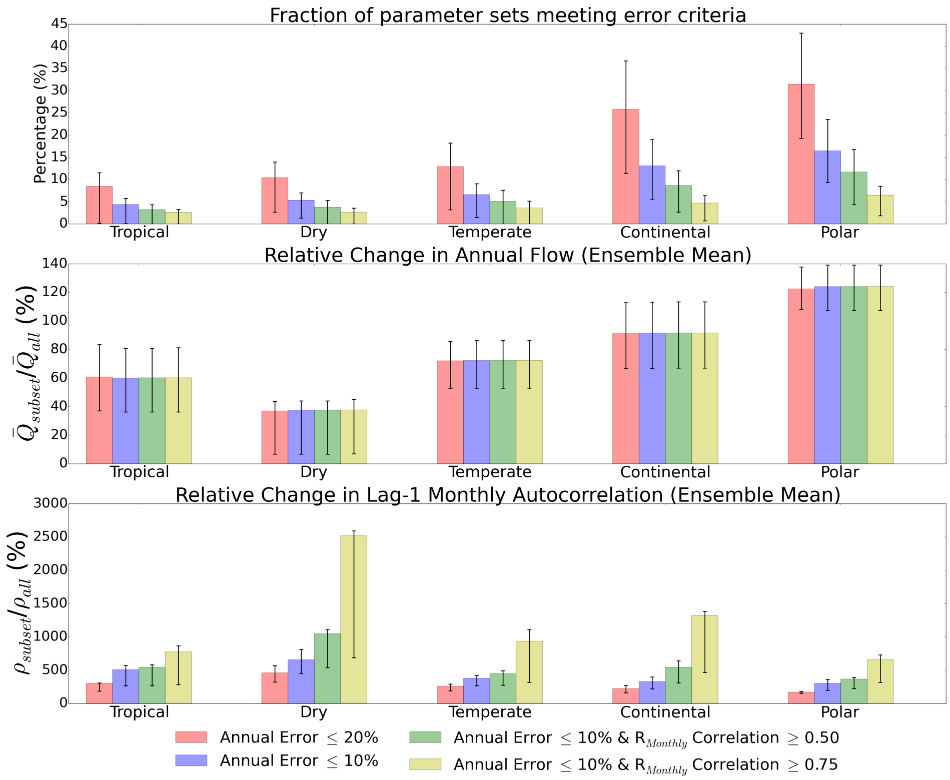

Figure 3. The grid cells with runoff observations are combined using the Köppen-Geiger climate classification to assess performance of the VIC ensemble as a function of climate type. The constraints define the fraction of parameters that meet the error criteria (top), the change in annual mean flow (center), and the change in 1-month lag correlation (bottom). The error bars quantify the variability within the climate type (25^{th} and 75^{th} percentile).

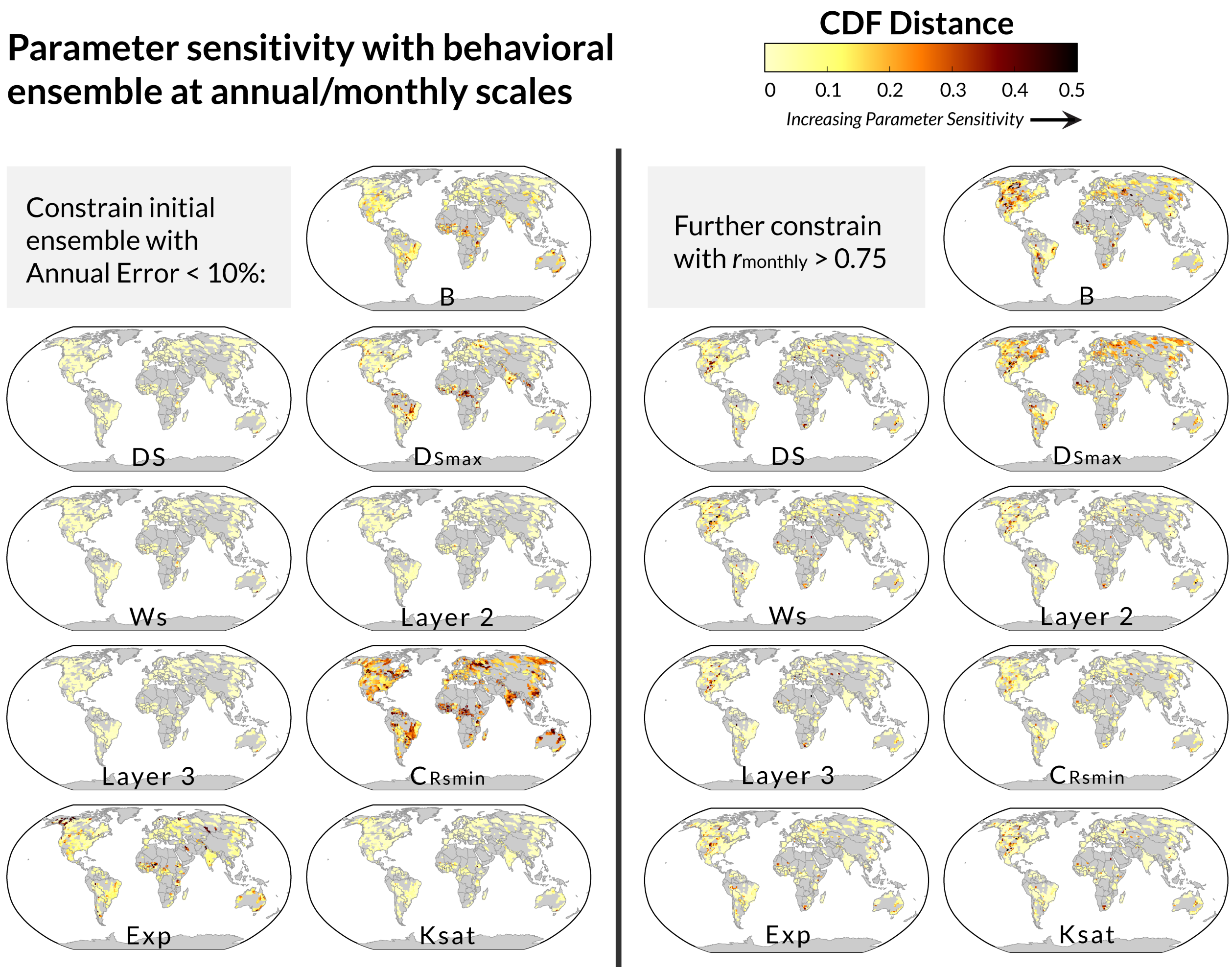

Figure 4. Global maps of the sensitivity of each VIC parameter used in the 10,000 Latin Hypercube Sample simulations. The CDF distance is calculated for each VIC parameter after applying the annual error constraint (left) and again after applying the monthly correlation constraint (right).

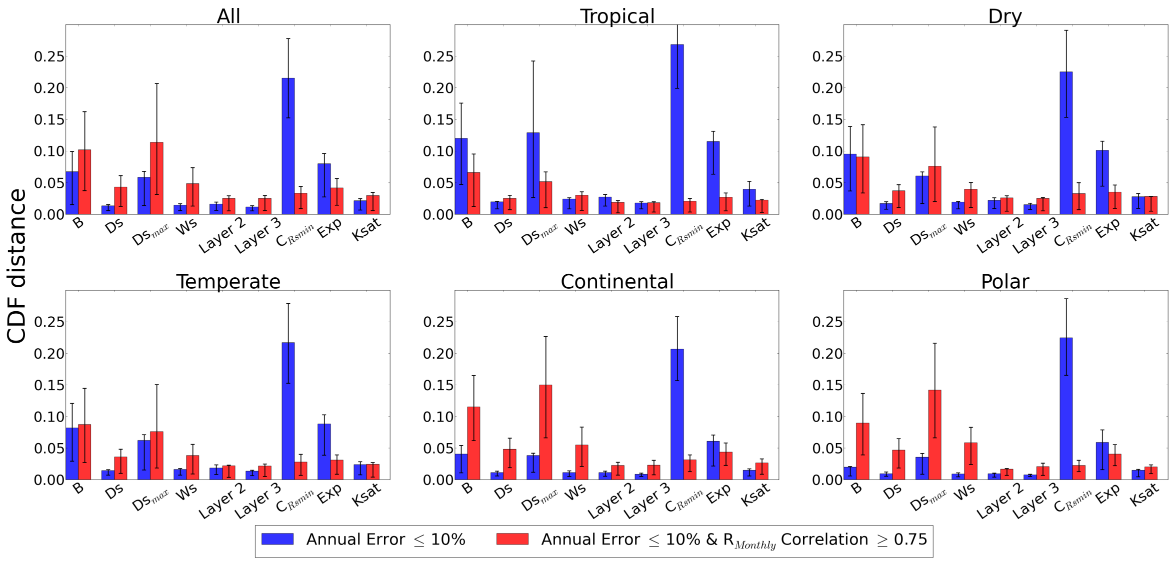

Figure 5. Climate average sensitivity of each VIC parameter used in the 10,000 Latin Hypercube Sample simulations. The CDF distance is calculated for each VIC parameter after applying the annual error constraint and again after applying the monthly correlation constraint. The error bars quantify the variability within the climate type (25th and 75th percentile).

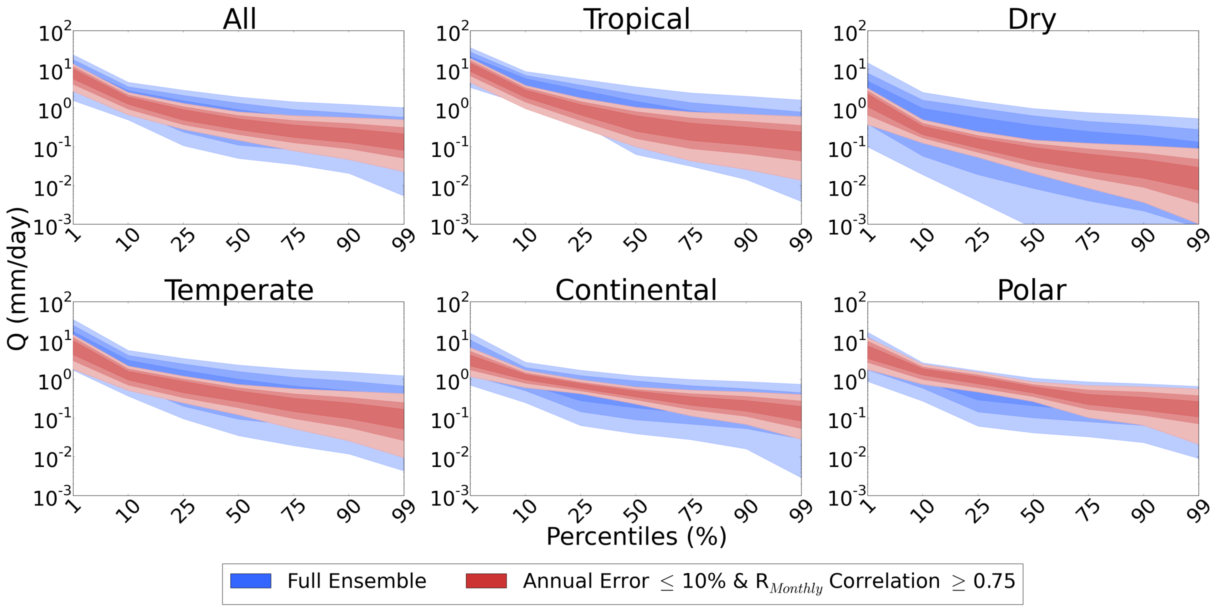

Figure 6. Climate averaged ensemble spread in the daily flow duration curve. The spread in flow duration curve is calculated for all 10,000 ensemble members. The blue shading shows the spread of the entire ensemble while the red shading shows the spread for parameter sets that have an annual mean runoff within 10% of the observed runoff and normalized monthly runoff correlation above or equal to 0.75.

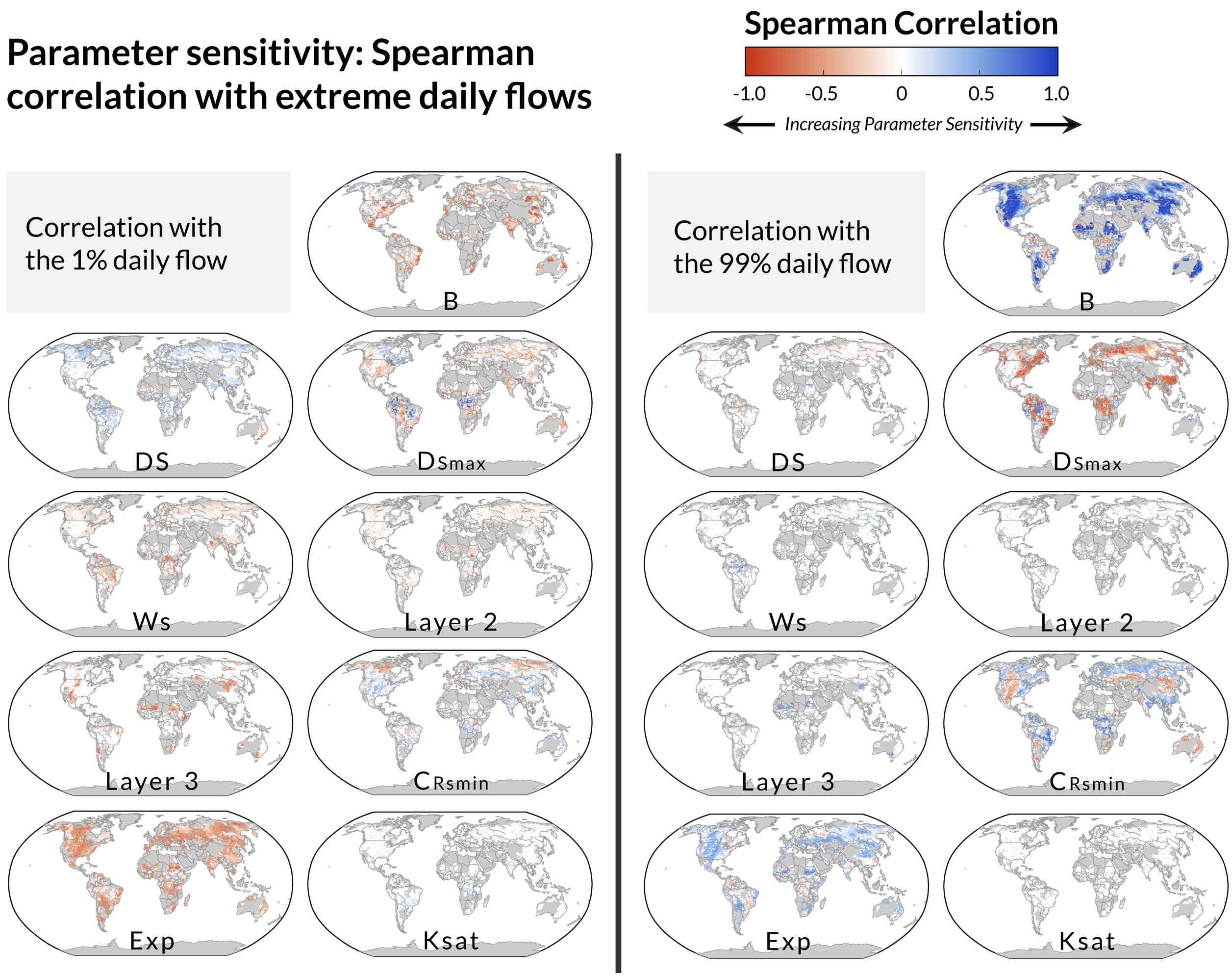

Figure 7. Global maps of the spearman correlation between the simulated extreme daily flows (1^{st} and 99^{th} percentile) and the corresponding VIC parameter. The correlations are calculated using the ensemble members that fulfill the strongest error criteria (relative error below 10% and monthly correlation above 0.75).