Predicting spatiotemporal variation in runoff in a data-sparse region: analyses of a whole-country model for Panama

Shriram Varadarajan[1,1], José Fabrega[2,2], and Brian Leung[1,1]

[1]McGill University
[2]Universidad Tecnológica de Panamá

September 12, 2021

Abstract

Panama faces seasonal floods and droughts as well as rising freshwater demands ranging from domestic consumption to hydropower and the operation of the Panama Canal. A process-based hydrological model of the country is a desirable scenario planning tool to complement the existing national water security plan. In Panama as in much of the Global South, sufficient observed data do not exist for all watersheds to calibrate complex hydrological models. Understanding and improving the performance of uncalibrated hydrological models could greatly expand their utility in such regions. In this study, we build and validate an uncalibrated Soil and Water Assessment Tool (SWAT) model for Panama. We extend the default precipitation submodel and demonstrate the importance of sufficiently accounting for for spatial autocorrelation patterns in precipitation inputs: we found large improvements over the default model, not only for monthly means (NSE = 0.88, from NSE= 0.69 for default SWAT), but especially for standard deviations (NSE = 0.59, from 0.27) and maxima (NSE = 0.51, from 0.21) of discharge across locations and months. We found a strong seasonal trend and regional differences in the spatial autocorrelation of rainfall, suggesting that this phenomenon should not be modeled statically. The resulting precipitation and hydrology models provide important baseline information for Panama, especially on variability and extremes, and could serve as a template for other regions with limited data.

Predicting spatiotemporal variation in runoff in a data-sparse region: analyses of a whole-country model for Panama

Authors: Shriram Varadarajan[1], José Fabrega[4], Brian Leung[1,2,3]

Affiliations:

1. Department of Biology, McGill University, Montreal, Quebec, Canada H3A 1B1
2. Bieler School of Environment, McGill University, Montreal, Quebec, Canada, H3A 2A7
3. Smithsonian Tropical Research Institute, PO Box 0843-03092, Panama City, Panama
4. Universidad Tecnológica de Panamá, Campus Victor Levi Sasso, Ancón, Vía Centenario, Panamá City, Panamá

Corresponding author: Shriram Varadarajan, shriram.varadarajan@mail.mcgill.ca

Other contact emails:

jose.fabrega@utp.ac.pa

brian.leung2@mcgill.ca

Funding:

i NSERC (National Science and Engineering Research Council of Canada) Discovery Grant to Brian Leung

ii MSSI (McGill Sustainability Systems Initiative) grant to Brian Leung

Keywords

Hydrological modeling, Soil and Water Assessment Tool, Neotropical hydrology, Sustainability

Abstract

Study region

Panama faces seasonal floods and droughts as well as rising freshwater demands ranging from domestic consumption to hydro-power and the operation of the Panama Canal. A process-based hydrological model of the country is a desirable scenario planning tool to complement the existing national water security plan.

Study focus

In Panama as in much of the Global South, sufficient observed data do not exist for all watersheds to calibrate complex hydrological models. Understanding and improving the performance of uncalibrated hydrological models could greatly expand their utility in such regions. In this study, we build and validate an uncalibrated Soil and Water Assessment Tool (SWAT) model for Panama. We extend the default precipitation submodel and demonstrate the importance of sufficiently accounting for for spatial autocorrelation patterns in precipitation inputs: we found large improvements over the default model, not only for monthly means (NSE = 0.88, from NSE= 0.69 for default SWAT), but especially for standard deviations (NSE = 0.59, from 0.27) and maxima (NSE = 0.51, from 0.21) of discharge across locations and months.

New hydrological insights for region

We found a strong seasonal trend and regional differences in the spatial autocorrelation of rainfall, suggesting that this phenomenon should not be modeled statically. The resulting precipitation and hydrology models provide important baseline information for Panama, especially on variability and extremes, and could serve as a template for other regions with limited data.

1. Introduction

The potential impacts of changes in the hydrological cycle brought about by climate change and human activity range from shortages of potable water (FAO 2018, p. 31) to increasing frequency and impact of floods and droughts (e.g., Hirabayashi et al. 2013). Changing patterns of precipitation and hydrology also strongly influence local vegetation and may negatively affect biodiversity (e.g., distance to water and precipitation were found to be two of the best predictors of species distributions, Bradie & Leung 2017). The capacity to predict the nature of changes in hydrological patterns will be crucial for effective risk mitigation and building systemic resilience. Given spatial heterogeneity in these dynamics, identifying critical regions at high risk for reductions in water availability and increases in extreme events must be an integral part of planning and management strategy. The need for this type of analysis is even more pronounced in the Global South, where the threat of climate change is compounded by economic and infrastructural inequality (Roberts 2001, Chapagain et al. 2020).

In Panama, our study area, a national framework for water resource management already exists in the form of the Plan Nacional de Seguridad Hídrica or National Water Security Plan (Comité de Alto Nivel de Seguridad Hídrica, 2016), which predicts a rise in water insecurity as human consumption reaches 50% of freshwater availability in the country by 2050. Freshwater is also a key resource for the Panama Canal system, which requires 52 million gallons per ship transit. The Canal Authority came close to having to impose draft restrictions due to lack of water during the wet season in 2015, an El Niño year (Autoridad del Canal de Panamá, 2015). The canal is uniquely important not only to Panama's economy with $2.6 billion in revenue (OECD 2017), but also as a key node in the global shipping trade. Additionally, about 45% of the country's electrical capacity is accounted for by hydropower (Autoridad de Servicios Públicos 2021), making variability in flow patterns critical to predict. Along with droughts like the one in 2015, floods and associated landslides are also problems faced by the country (e.g., Wohl and Ogden, 2013). These changes are projected to have further downstream effects ranging from agricultural yield changes to the persistence of Chagas disease (Fábrega et al. 2013). Beyond human impacts, Panama is also at the center of one of the world's most biodiverse regions (Myers et al. 2000), and the rich tropical forests and aquatic ecosystems that support this diversity are heavily reliant on the health of its waterways.

In this study, we use the Soil and Water Assessment Tool (SWAT, Arnold et al. 1998) to build a countrywide hydrological model of Panama. SWAT is a process-based model that incorporates information about meteorology, physical geography, and human land use to simulate the entire hydrological cycle of the study area. Spatial variation is made explicit in SWAT by splitting each watershed into 'subbasins' of non-branching stream segments and their drainage areas, and further splitting each subbasin into a set of Hydrological Response Units (HRUs) which represent a particular combination of slope, land use, land management, and soil type. SWAT can thus provide sophisticated and holistic hydrological projections for given patterns of changes in its inputs.

The utility of SWAT in simulating single watersheds with model parameters calibrated to local conditions, at least at the monthly timestep, is well established (e.g., Perez-Valdivia et al., 2017). However, the calibration procedures are complex and require comprehensive and high-quality hydrological data (Abbaspour et al. 2015), which are not available for many watersheds in most regions of the Global South. Yet, models such as SWAT use physics-based equations, and thus in principle could have predictive power even without calibration. While uncalibrated SWAT models have been shown to perform well in some contexts (Srinivasan et al. 2010), its performance needs to be tested in different regions, and for different phenomena of interest (e.g., mean, variation and extremes in flow). Improvements to model structure and parameter calibration can be considered separate avenues of improving predictive power (Butts et al. 2004) and an improved soil submodel for SWAT has been shown to improve streamflow and nitrate load predictions even in the absence of parameter calibration (Qi et al. 2020). In particular, the ability to adequately describe environmental inputs will likely be a key determinant in the ability to predict hydrological patterns, and will also pose a greater challenge in the Global South compared to regions with more complete data. That said, precipitation gauges are generally more numerous and available than hydrological measurement stations, and given that precipitation patterns are the most direct driver of hydrological phenomena, precipitation input would be a logical focus of attention.

In the present study we thus propose modifications to the precipitation submodel of SWAT, and demonstrate that it is crucial to capture regional and spatio-temporal autocorrelation patterns in precipitation. For both the default SWAT model and one with our modified precipitation algorithm, we examine predictive power for water availability (mean monthly discharge) as well as variability (standard deviation and maxima of discharge) across space and time. We also examine the performance of the models within each watershed, and identify characteristics of watersheds that explain the variation in this performance.

2. Methods

2.1 SWAT Model Setup

The SWAT model requires a set of spatially explicit inputs for the study area: a digital elevation model (DEM), a soil map, a land use map, and a set of weather station locations. The weather stations further must be provided with precipitation, solar radiation, relative humidity, and wind data in the form of either (i) records for each day of the simulation or (ii) parameters for a rainfall distribution that the model samples from on each simulated day. The data sources used for each of the above in the current study are summarized in Table 1. Data on river discharge from the ETESA (Empresa de Transmisión Eléctrica, S.A; https://www.etesa.com.pa/) hydrological monitoring network from the period 2005 – 2015 was used for model validation, while ETESA precipitation data from the periods 1990 – 2000 and 2005 – 2015 were used for fitting the precipitation submodel parameters and running the validation simulation respectively. All of the above rain gauges and hydrological monitoring stations are mapped in Figure 1.

Watershed delineation was carried out in ArcSWAT. A threshold of 5000 cells was chosen as the minimum inflow into an outlet for which a subbasin would be defined, which amounts to a drainage area of about 40.5 km^2 given the DEM cell size at the equator. Areas smaller than this which drain directly into the sea or either neighboring country were not part of the model, resulting in a model delineation covering roughly 65,000 km^2 or 86% of the total land area of Panama. SWAT further assigns each non-branching segment of stream its own subbasin unit and calculates Hydrological Response Units (HRUs) within each subbasin based on existing combinations of soils, land use, and slope. SWAT generates daily mean discharge output (m^{3}s^{-1}) at the outlet of each subbasin, so additional outlets were manually defined at the location of each hydrological measurement station (52 in total) to provide direct comparison points. This resulted in a delineation of 980 subbasins in total.

2.2 Precipitation interpolation

Daily precipitation data from ETESA was downloaded for 249 rain gauge locations across Panama, of which 120 were active during the simulation period of 2005 - 2015, though many had substantial temporal gaps in their records. SWAT requires daily precipitation values for each subbasin (980 in total in the current model) and thus some method of interpolation is required to fill both spatial and temporal gaps in the data coverage. For spatial gaps, the method used by default is a nearest neighbor (or Thiessen polygon) interpolation from each subbasin centroid to the nearest rain gauge. For temporal gaps, the default method is sampling from an empirically determined rainfall distribution at rain gauge locations using a skew-normal distribution (see Neitsch et al., 2011). Means, standard deviations, skew, and wet-dry transition probability values were calculated at all gauge locations using the observations spanning the period of 1990 - 2000. Henceforth this method will be referred to as the default model.

In our modified method, interpolation of precipitation gauge data at each subbasin centroid was carried out for each day, and separately for each of six climatic regions that the country was split into (Figure 2). We first tested a single-step inverse-distance weighting model, with a single distance-decay parameter (α, see Eq 1) fit by region and month (this model was named ‘**RDW1**’, for ‘single-step regional distance weighting’). Then, we tested a two-step method (‘**RDW2**’) at each location which incorporated an explicit prediction of rainfall occurrence: (1) first we predicted the probability of a wet or dry day (occurrence) using a logistic regression on the inverse distance-weighted mean of observations in the region (using a threshold observation <

0.5 mm, dry gauges were given a value of 0 and wet gauges a value of 1); (2) then, given the probability of a wet day, we performed a binomial trial; and if a wet day was generated, we interpolated the quantity of rain, again using an inverse distance weighted mean of all observed quantities of precipitation within the region for that day. Dry days were assigned a quantity of 0 mm. The two steps involved a single parameter each, α_1 and α_2, which controlled the decay of the relative weighting with distance.

$$p_{ikt} = \frac{\sum p_{jkt} e^{-\alpha_{kt} d_{ijkt}}}{\sum e^{-\alpha_{kt} d_{ijkt}}} \qquad \text{Eq 1}$$

Where p_{ikt} was the precipitation value of interest (i.e., either wet/dry or quantity of rain), p_{jkt} denoted all measured precipitation values, d_{ijkt} was the distance between points i and j, and α_{kt} was a shape parameter.

These α_{kt} values were fit separately to each monthly time interval (t), and each of 6 regions (k) in Panama, to capture expected differences in spatial and seasonal autocorrelation patterns. In the RDW2 model, the predicted total quantity of rain across a region on a given day (q1) was reallocated to those locations that were predicted to be wet on that day (with a total quantity q2, where q2 ≤ q1), i.e. the predicted quantity in each of these locations was scaled by the ratio q1/q2, to prevent underestimation caused by the independent prediction of occurrence using binomial trials.

Each fitted parameter (α in Eq 1) represented the strength of the distance-weighting, with high α severely penalizing information from gauges further away from the target point and prioritizing immediate neighbors. The fitting process, minimizing the sum of squared deviations, was run using the gauge data from 1990 – 2000, and the fitted algorithm was then used to interpolate daily values at the 980 subbasin centroids using gauge data from 2005 - 2015 for the validation run.

2.3 Model evaluation

As the objective was to gain insight into patterns of water distribution, the hydrological model results were compared against river discharge data from ETESA, which was not used to calibrate or parameterize any part of the models. R^2, Nash-Sutcliffe Efficiency (NSE, Nash and Sutcliffe 1970), and percentage bias ('Pbias') of mean model predictions against mean observed daily discharge values were calculated for each month and station across the entire simulation period and these were used as metrics of the ability to predict average flow.

We also used NSE, R^2, and Pbias to examine the ability of the hydrological model to estimate variation in runoff at a given location, using as metrics (i) standard deviation of discharge and (ii) the magnitudes of the 3 highest daily discharge events across the simulation period in each location-month combination. We chose the latter to represent of the extreme highs of the discharge distribution for that combination and as a coarse indicator of flood risk.

Finally, we used NSE and R^2 to examine the model's ability to simulate the observed monthly time series of mean discharge within each basin from 2005 – 2015. Instead of spatial variation across locations, this procedure tested the ability of the model to capture temporal variation within each watershed (using the lowermost hydrological station in each of 35 watersheds). We posited several variables across that could explain variation in model predictiveness, namely: (i) elevation of the observation (as we did not account for orographic effects explicitly), (ii) existence of a precipitation gauge within the same subbasin as the observation and (iii) number of precipitation gauges in the region (both as measures of the relevance and quantity of precipitation information), (v) number of subbasins in the watershed (as larger watersheds could have more complex behaviour), (vi) number of subbasins downstream from the observation (as interior reaches could behave differently regardless of elevation), and (vi) simulated standard deviation of mean monthly discharge at the location (as the ability to predict flow might depend on variability). We used a stepwise forward selection algorithm in R to arrive at the linear combination of these variables and their pairwise combinations with the lowest AIC value.

3. Results

3.1 Precipitation interpolation

In the regional distance weighted (RDW1 and RDW2) models, there was a strong seasonal signal in the autocorrelation patterns as represented by the fitted distance-decay parameters (α, Figure 3). Maximum values of the α parameters correspond to the greatest weighting of the closest stations and, correspondingly, the fastest decay in weighting with distance. These maxima consistently occurred for both precipitation occurrence and quantity in April and October, and these months represent the two transitions between the wet (generally May to November) and dry seasons. Over the remainder of each season, parameter values decline and then rise gradually (for both the single parameter in RDW1 and the occurrence parameter in RDW2) or remain generally low (for the quantity parameter in RDW2) until the next seasonal transition. Variation across region was comparatively higher for both RDW2 parameters than for the single RDW1 parameter.

3.2 Model evaluation

The standard SWAT model using default precipitation interpolation performed reasonably well for mean discharge across locations with an NSE of 0.69. Improving the precipitation interpolation approach yielded even better predictions; NSE = 0.88 using the RDW2 model and

NSE = 0.89 using the RDW1 model. While the default model worked well for mean discharge (NSE = 0.69), it was less able to capture variability, with NSE = 0.26 for standard deviation of locations-month combinations. In contrast, for standard deviation, the RDW2 model achieved NSE = 0.59 and RDW1 NSE = 0.53, which compared very favorably with the default model. For predicting maxima of daily discharge at each location-month combination, we found again that the default model performed poorly, with NSE = 0.22, while the RDW2 model achieved a higher NSE of 0.53, and RDW1 again performed similarly to RDW2 with NSE = 0.51. Notably, however, Pbias values were significantly larger in magnitude for the RDW1 model than for RDW2 across all analyses. These results are summarized in Table 2 and Figure 4.

We also examined the ability to predict temporal patterns within each watershed for the default and RDW models across the ten-year period (Figure 5). We found that the default model performed generally poorly, being a worse predictor in 24 out of 35 sites than simply using the observed mean (i.e. NSE < 0). The RDW2 model performed significantly better: while 8 sites still performed poorly (NSE < 0), 15 (43%) sites had satisfactory performance at NSE > 0.5, and the median NSE was 0.4 across locations. The RDW1 model performed intermediately, with no sites achieving NSE > 0.5 and a median NSE of 0.21. RDW2 was chosen as the overall best model due to this as well as the lower magnitude of bias mentioned above.

While there were areas of failure, the performance of RDW2 was largely predictable. 71% of the variation in NSE across locations was explained by six variables and two interaction terms. The variables were (i) elevation, (ii) total number of rain gauges in region, (iii) number of downstream subbasins, (iv) size of watershed (i.e., number of subbasins), (v) simulated standard deviation of discharge, and (vi) presence of (of which (ii), (iv), and (vi) were significant), and the significant interaction terms were between watershed size and number of downstream subbasins, and between watershed size and simulated standard deviation (Table 3).

4. Discussion

Panama faces a variety of issues related to potential changes in the water cycle ranging from shortages of drinking water and hydropower to increased impact of floods and droughts. While sophisticated hydrological modeling tools such as SWAT exist, the data available with which to build and calibrate such models is generally more limited in Panama and other countries of the Global South. Thus, while Srinivasan et al. (2010) demonstrated that an uncalibrated SWAT model predicted streamflow similarly to calibrated ones in the Upper Mississippi basin of the USA, testing the performance of such a model in Panama is necessary, given differences in environmental conditions and limitations of the available input data. Indeed, we found that while the default SWAT model performed well for predicting monthly mean flow across watersheds, it fared poorly for predicting monthly standard deviations and maxima. Substantial improvements were obtained across all three metrics by using a 2-stage interpolation algorithm for precipitation (i.e., our RDW2 model).

These findings highlight the importance of validating model performance in different regions, but also the potential promise of uncalibrated models even in locations where hydrological data are limited. Standard deviations and maxima, which the default model predict poorly, represent information about streamflow distributions that are crucially important in the predictive modeling of flood risk (e.g. van der Wiel et al. 2019). Furthermore, variability in water availability is a critical indicator of potential water scarcity, and has significant impacts on human water use, despite often being overlooked in favour of annual means (Damkjaer &

Taylor, 2017). Modeling studies also show that hydropower output is sensitive to variability in hydroclimatic inputs (Arriagada et al., 2019, Chowdhury et al. 2020). While future hydroclimate projections for Panama have been made for monthly mean discharge (Fabréga et al. 2013), the present study lays the groundwork for improved projections based on a more sophisticated hydrological model with higher spatial resolution and finer prediction of variability and extremes.

The relative predictive failure of the default model was due to the precipitation model which by default matches each subbasin to its nearest precipitation gauge, and fills temporal gaps in the daily records of each gauge by sampling from an empirical distribution that models the behavior of that gauge (Neitsch et al. 2011). This sampling is done independently of any other gauge value on that day. As these temporal gaps occurred frequently (the mean precipitation gauge was only active on 60% of days from 2005 to 2015), the result of this independent sampling would tend to average out variability across subbasins, and contribute to the observed pattern of underestimated variances and extremes of streamflow in the default model. Such spatial and especially temporal gaps in rainfall data have been shown to have significant negative impacts on SWAT performance (Tan and Yang 2020). In contrast, due to the use of distance-weighted interpolation fitted on seasonal and regional spatial patterns of precipitation, estimated values across gauges within a region on a given day more faithfully replicated real precipitation patterns in the RDW1 and RDW2 models.

While both distance-weighted models were clear improvements on the default model, RDW2 also outperformed RDW1 markedly in terms of percentage bias of all metrics, with RDW1 generally underpredicting streamflow volume and variability. Most often, distance-weighted interpolation for precipitation has been modeled as a single step calculating quantity of rain (e.g.

Chen and Liu, 2012; Cheng et al., 2017; Tuo et al., 2016; Xue et al., 2018), as in the RDW1 model. Yet, two-step interpolation treating (i) the occurrence of precipitation and (ii) the quantity separately has been found to regenerate more realistic patterns of spatial variability for daily precipitation (Hwang et al. 2012), as these two factors need not be linearly related and cannot be captured with a single distance function. RDW2 is thus, overall, the best model of the three tested.

Better accounting for spatial relatedness of precipitation gauges was also crucial for estimating fluctuations in streamflow across time within each watershed (as opposed to variation across watersheds discussed above). The uncalibrated default SWAT model generally performed no better than simply using the mean flow in most watersheds (median NSE < 0), again highlighting the importance of testing in different regions and for different metrics. With our RDW2 model, the median watershed NSE was 0.4 and 43% of locations achieved NSE > 0.5, defined to be 'satisfactory' performance by Moriasi et al. (2007). Furthermore, we could largely identify where failures in the RDW2 model occurred, explaining 71% of the variation in basin NSEs, and showed that the total number of gauges in the region and the presence of a gauge in the subbasin itself were both predictors of higher NSE at a given location.

Additionally, our findings suggest that spatial patterning of rainfall varied over time, with peaks in parameter values of the distance-weighted interpolation kernel in the months of April and October. High parameter values indicate that nearby gauges are much more predictive of precipitation at a point than ones further away. Low parameter values on the other hand indicate a broader averaging, and more regional forcing. The two months of highest parameter values, April and October, coincide with the periods of change in patterns of observed variation in rainfall (Fabrega et al., 2013) as well as periods of strongest increase and decrease in average rainfall respectively (Kusunoki et al., 2019). Further elucidation of the processes that lead to the variation in spatial autocorrelation captured by the current model may aid in the development of a dynamical procedure that can account for nonstationarity. While seasonal and regional patterns of spatial autocorrelation in precipitation and their effects on extreme events are being studied in some arid and semi-arid areas in regions such as China (Xu et al., 2021) and Iran (Darand et al., 2017, Rousta et al. 2017), they remain understudied in the tropics where they may also be of importance.

Conclusions

Our findings indicate that uncalibrated hydrological models such as SWAT can be predictive, and that a key limitation in Panama had been the default precipitation sub-model. Improving description of precipitation input by incorporating information about regional and seasonal differences in spatial autocorrelation patterns dramatically improved predictions across a number of metrics, including means, standard deviations, and maxima of monthly streamflow across watersheds, as well as the time series of monthly flow in each watershed. As precipitation gauges tend to be common and relatively simple to set up, the application of hydrological models across large and heterogeneous spatial contexts becomes much more feasible, even in regions where data limitations make hydrological calibration difficult.

Data availability

All data of simulation results and code used in the present analysis are available at https://doi.org/10.5281/zenodo.6111112 .

References

Abbaspour, K. C. *et al.* A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model. *Journal of Hydrology* **524**, 733–752 (2015).

Arnold *et al.* SWAT: Model Use, Calibration, and Validation. *Transactions of the ASABE* **55**, 1491–1508 (2012).

Arriagada, P., Dieppois, B., Sidibe, M. & Link, O. Impacts of Climate Change and Climate Variability on Hydropower Potential in Data-Scarce Regions Subjected to Multi-Decadal Variability. *Energies* **12**, 2747 (2019).

Autoridad del Canal de Panamá (ACP),. Canal de Panamá suspende restricción al calado de buques – Canal de Panamá. https://micanaldepanama.com/canal-de-panama-suspende-restriccion-al-calado-de-buques/ (2015). Accessed 15 Feb 2022.

Autoridad de Servicios Públicos (ASEP). 2021. Available from: https://www.asep.gob.pa/wp-content/uploads/electricidad/estadisticas/2020/segundo_semestre/oferta.pdf. Accessed 15 Feb 2022.

Bradie, J. & Leung, B. A quantitative synthesis of the importance of variables used in MaxEnt species distribution models. *Journal of Biogeography* **44**, 1344–1361 (2017).

Butts, M. B., Payne, J. T., Kristensen, M. & Madsen, H. An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. *Journal of Hydrology* **298**, 242–266 (2004).

Chapagain, D., Baarsch, F., Schaeffer, M. & D'haen, S. Climate change adaptation costs in developing countries: insights from existing estimates. *Climate and Development* **12**, 934–942 (2020).

Chen, F.-W. & Liu, C.-W. Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan. *Paddy Water Environ* **10**, 209–222 (2012).

Cheng, M. *et al.* Performance Assessment of Spatial Interpolation of Precipitation for Hydrological Process Simulation in the Three Gorges Basin. *Water* **9**, 838 (2017).

Chowdhury, A. F. M. K., Dang, T. D., Bagchi, A. & Galelli, S. Expected Benefits of Laos' Hydropower Development Curbed by Hydroclimatic Variability and Limited Transmission Capacity: Opportunities to Reform. *Journal of Water Resources Planning and Management* **146**, 05020019 (2020).

Comité de Alto Nivel de Seguridad Hídrica. Plan Nacional de Seguridad Hídrica 2015-2050: Agua para Todos. Panamá, República de Panamá. (2016)

Damkjaer, S. & Taylor, R. The measurement of water scarcity: Defining a meaningful indicator. *Ambio* **46**, 513–531 (2017).

Darand, M., Dostkamyan, M. & Rehmani, M. I. A. Spatial autocorrelation analysis of extreme precipitation in Iran. *Russ. Meteorol. Hydrol.* **42**, 415–424 (2017).

F.A.O. The future of food and agriculture – Alternative pathways to 2050. (2018).

Fábrega, J. *et al.* Hydroclimate projections for Panama in the late 21st Century. *Hydrological Research Letters* **7**, 23–29 (2013).

Hirabayashi, Y. *et al.* Global flood risk under climate change. *Nature Climate Change* **3**, 816–821 (2013).

Hwang, Y., Clark, M. R., Rajagopalan, B. & Leavesley, G. H. Spatial interpolation schemes of daily precipitation for hydrologic modeling. *Stochastic Environmental Research and Risk Assessment* vol. 26 295320 (2012).

Kusunoki, S., Nakaegawa, T., Pinzón, R., Sanchez-Galan, J. E. & Fábrega, J. R. Future precipitation changes over Panama projected with the atmospheric global model MRI-AGCM3.2. *Clim Dyn* **53**, 5019–5034 (2019).

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

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B. & Kent, J. Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858 (2000).

Nash, J. E. & Sutcliffe, J. V. River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology* **10**, 282–290 (1970).

Neitsch, S.L., Williams, J.R., Arnold, J.G. and Kiniry, J.R. (2011) Soil and Water Assessment Tool Theoretical Documentation Version 2009. Texas Water Resources Institute, College Station.

Perez-Valdivia, C., Cade-Menun, B. & McMartin, D. W. Hydrological modeling of the pipestone creek watershed using the Soil Water Assessment Tool (SWAT): Assessing impacts of wetland drainage on hydrology. *Journal of Hydrology: Regional Studies* **14**, 109–129 (2017).

Qi, J. *et al.* SWAT ungauged: Water quality modeling in the Upper Mississippi River Basin. *Journal of Hydrology* **584**, 124601 (2020).

Roberts, J. T. Global inequality and climate change. *Society and Natural Resources* **14**, 501–509 (2001).

Rousta, I., Doostkamian, M., Haghighi, E., Ghafarian Malamiri, H. R. & Yarahmadi, P. Analysis of spatial autocorrelation patterns of heavy and super-heavy rainfall in Iran. *Adv. Atmos. Sci.* **34**, 1069–1081 (2017).

Srinivasan, R., Zhang, X. & Arnold, J.. SWAT Ungauged: Hydrological Budget and Crop Yield Predictions in the Upper Mississippi River Basin. *Transactions of the ASABE* **53**, 1533–1546 (2010).

Tan, M. L. & Yang, X. Effect of rainfall station density, distribution and missing values on SWAT outputs in tropical region. *Journal of Hydrology* **584**, 124660 (2020).

Tuo, Y., Duan, Z., Disse, M. & Chiogna, G. Evaluation of precipitation input for SWAT modeling in Alpine catchment: A case study in the Adige river basin (Italy). *Science of The Total Environment* **573**, 66–82 (2016).

van der Wiel, K., Wanders, N., Selten, F. M. & Bierkens, M. F. P. Added Value of Large Ensemble Simulations for Assessing Extreme River Discharge in a 2 °C Warmer World. *Geophysical Research Letters* **46**, 2093–2102 (2019).

Wohl, E. & Ogden, F. L. Organic carbon export in the form of wood during an extreme tropical storm, Upper Rio Chagres, Panama. *Earth Surface Processes and Landforms* **38**, 1407–1416 (2013).

Xu, L., Zheng, C. & Ma, Y. Variations in precipitation extremes in the arid and semi-arid regions of China. *International Journal of Climatology* **41**, 1542–1554 (2021).

Xue, F., Shi, P., Qu, S., Wang, J. & Zhou, Y. Evaluating the impact of spatial variability of precipitation on streamflow simulation using a SWAT model. *Water Policy* **21**, 178–196 (2018).

Tables and figures

Table 1. Data sources for building the SWAT model.

Data layer	Source
DEM (Digital Elevation Model)	USGS Earth Resources Observation And Science (EROS) Center. (2017). Shuttle Radar Topography Mission (SRTM) 1 Arc-Second Global [Data set]. U.S. Geological Survey. https://doi.org/10.5066/F7PR7TFT
Soil map	FAO-UNESCO Soil Map of the World, accessible at https://data.apps.fao.org/map/catalog/srv/eng/catalog.search#/metadata/446ed430-8383-11db-b9b2-000d939bc5d8
Land use map	iii For fitting the precipitation submodel, simulation period 1990 – 2000; iv For the validation simulation, 2005 – 2015; "Panama 2012 Forest Cover and Land Use", STRI GIS Data Portal, accessible at https://stridata-si.opendata.arcgis.com/maps/SI::panama-2012-forest-cover-and-land-use-tile-layer/about
Precipitation & discharge	ETESA hydrological and meteorological stations, STRI meteorological stations
Other climate variables (Solar radiation, wind, relative humidity, temperature)	National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR) data, available at https://globalweather.tamu.edu/

Table 2. Summary of simulation results; all statistics calculated by location and calendar month across the whole country.

Model	Summary statistics (monthly discharge, m^3/s)								
	Mean			Standard deviation			Maxima		
	R^2	NSE	pbias	R^2	NSE	pbias	R^2	NSE	pbias
Default	0.70	0.69	-11.4	0.34	0.27	0.3	0.32	0.21	1.0
1-step regional distance-weighted (RDW1)	0.90	0.89	-15.3	0.60	0.53	-30.1	0.52	0.49	-23.5
2-step regional distance-weighted (RDW2)	0.88	0.88	-9.5	0.61	0.59	-0.9	0.53	0.51	-4.1

Table 3. Linear regression model summary for predictors of within-basin NSE value. AIC = 5.74, adjusted R^2 of prediction = 0.71.

Predictor	Coefficient	Std. Error	Significance ($p < 0.05$)
Intercept	0.55	0.07	*
1. Elevation	-0.17	0.09	
2. Total number of gauges in region	0.12	0.05	*
3. Number of downstream subbasins	0.24	0.13	
4. Size of watershed (# subbasins)	-0.28	0.08	*
5. Simulated standard deviation of discharge	-0.05	0.06	
6. Presence of rain gauge within subbasin	0.12	0.05	*
Interaction term 3*4	-0.38	0.09	*
Interaction term 4*5	-0.39	0.11	*
Interaction term 1*3	-0.26	0.17	
Interaction term 2*5	0.11	0.09	

Figure 1: Map of meteorological and hydrological stations used in simulation. Basin area upstream of gauge locations used for validation highlighted in blue (30 basins), all other regions were ungauged for the period of 2005 – 2015. Not all meteorological stations are necessarily active at any given point.

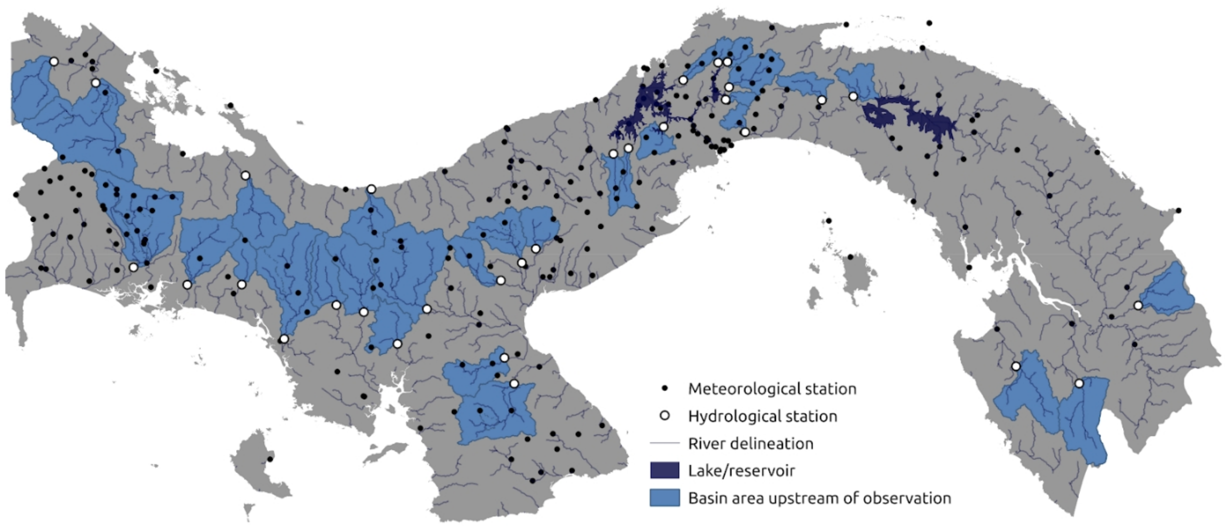

Figure 2: Climatic regions as delineated on the area covered by the SWAT model. 1 – Caribbean side of the Tabasará mountains, 2 – Pacific side of the Tabasará mountains, 3 – Azuero peninsula, 4 – Central Panama, 5 – East-Central Panama, 6 – Darien region

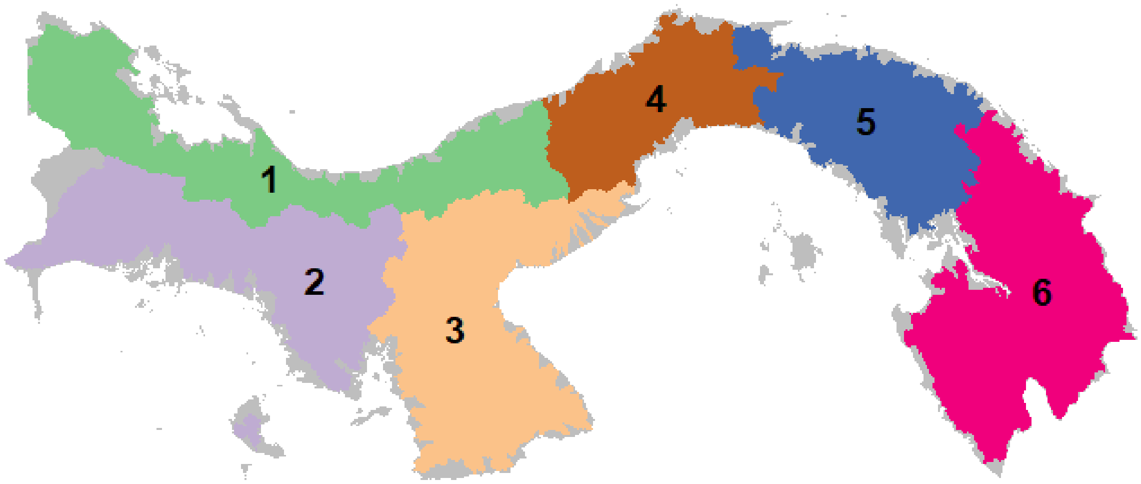

Figure 3: Distance weighting parameter values for RDW1 (α_0) and RDW2 (α_1 for occurrence, α_2 for quantity) model. A higher value of the parameter indicates that closer neighbours are weighted much higher than ones further away, and a lower value indicates a slower distance-decay function and thus a more even distribution of weights across the region.

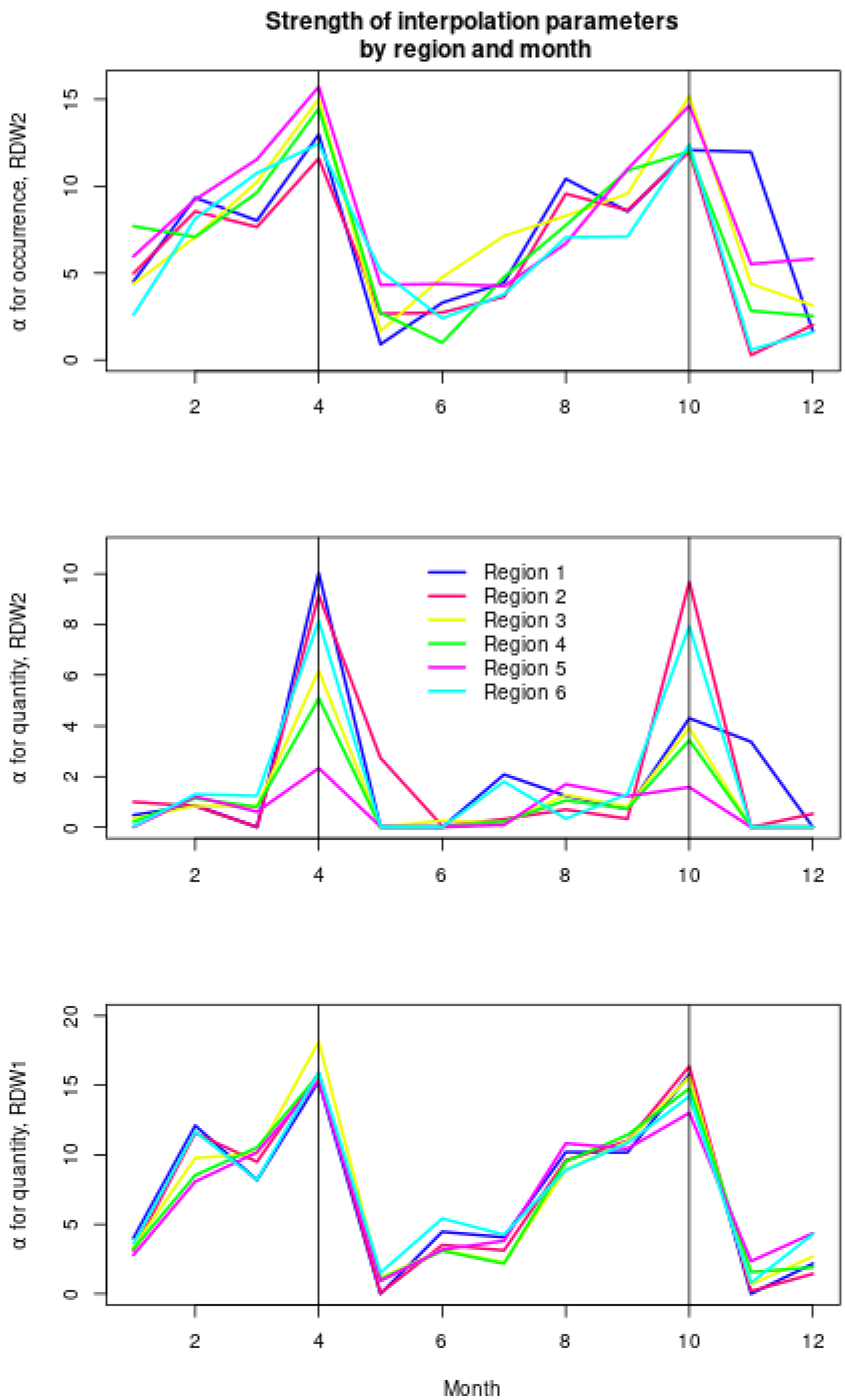

Figure 4: Scatterplots of model performance in predicting mean monthly discharge, standard deviations, and daily maxima of monthly discharge by location and calendar month

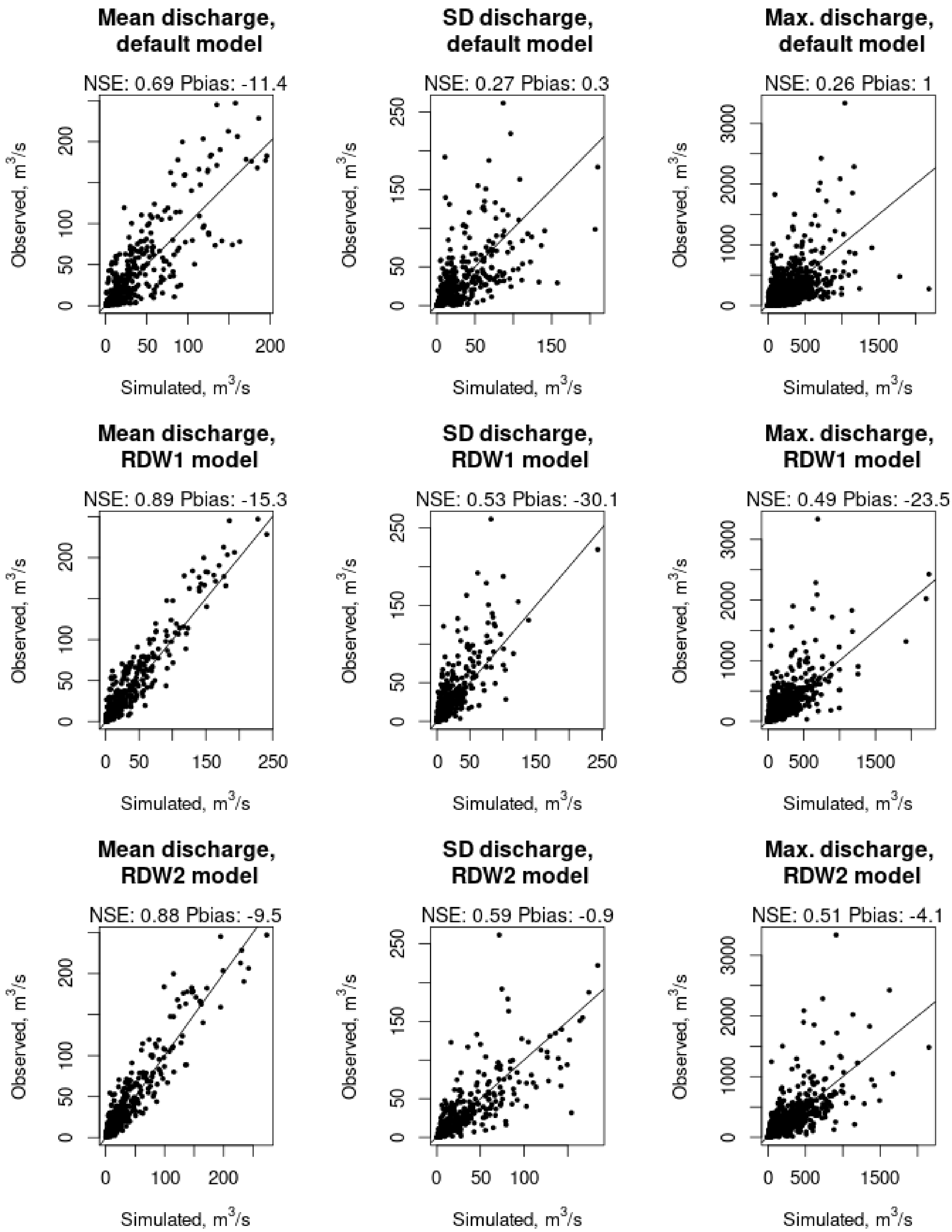

Figure 5: NSE for monthly mean prediction across 2005 – 2015 by basin for (a) default model, (b) RDW1 model, and (c) RDW2 model. Basins with NSE ≤ 0 are not labeled.