

Article

Spatio-Temporal Hydrological Model Structure and Parametrization Analysis

Mostafa Farrag ^{1,*} , Gerald Corzo Perez ²  and Dimitri Solomatine ^{2,3,4,*}

¹ GFZ German Research Centre for Geosciences, Hydrology Section, Telegrafenberg, 14473 Potsdam, Germany

² Chair Group of Hydroinformatics, IHE Delft Institute for Water Education, P.O. Box 3015, 2601 DA Delft, The Netherlands; g.corzo@un-ihe.org

³ Water Resources Section, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands

⁴ Water Problems Institute of RAS, Gubkina 3, 119333 Moscow, Russia

* Correspondence: mostafa.farrag@gfz-potsdam.de (M.F.); d.p.solomatine@tudelft.nl (D.S.)

Abstract: Many grid-based spatial hydrological models suffer from the complexity of setting up a coherent spatial structure to calibrate such a complex, highly parameterized system. There are essential aspects of model-building to be taken into account: spatial resolution, the routing equation limitations, and calibration of spatial parameters, and their influence on modeling results, all are decisions that are often made without adequate analysis. In this research, an experimental analysis of grid discretization level, an analysis of processes integration, and the routing concepts are analyzed. The HBV-96 model is set up for each cell, and later on, cells are integrated into an interlinked modeling system (Hapi). The Jiboa River Basin in El Salvador is used as a case study. The first concept tested is the model structure temporal responses, which are highly linked to the runoff dynamics. By changing the runoff generation model description, we explore the responses to events. Two routing models are considered: Muskingum, which routes the runoff from each cell following the river network, and Maxbas, which routes the runoff directly to the outlet. The second concept is the spatial representation, where the model is built and tested for different spatial resolutions (500 m, 1 km, 2 km, and 4 km). The results show that the spatial sensitivity of the resolution is highly linked to the routing method, and it was found that routing sensitivity influenced the model performance more than the spatial discretization, and allowing for coarser discretization makes the model simpler and computationally faster. Slight performance improvement is gained by using different parameters' values for each cell. It was found that the 2 km cell size corresponds to the least model error values. The proposed hydrological modeling codes have been published as open-source.

Keywords: conceptual-distributed model; raster-based; model structure analysis; Muskingum routing; OAT sensitivity analysis; HBV



Citation: Farrag, M.; Perez, G.C.; Solomatine, D. Spatio-Temporal Hydrological Model Structure and Parametrization Analysis. *J. Mar. Sci. Eng.* **2021**, *9*, 467. <https://doi.org/10.3390/jmse9050467>

Academic Editor: Chung-yen Kuo

Received: 20 March 2021

Accepted: 16 April 2021

Published: 26 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Among different types of hydrological models, an increasing interest has been shown in spatially distributed models for their ability to derive spatially distributed and, in particular, to support the analysis of land use impact on landscape hydrology and water quality management [1–4]. For these applications, water balance components and fluxes throughout the landscapes have to be identified, and thus land-surface heterogeneity representation should be considered [5]. Extensive studies have been carried out to develop physical models, but their extensive requirements for data limit their use considerably [1]. On the other hand, distributed conceptual hydrological models are mainly based on mass conservation, usually calibrated with simple rainfall-runoff information, and still preserve certain physical basis in their structure [2,6–8].

Nowadays, there are several physically based distributed hydrological models used for practical purposes, such as the Institute of Hydrology Distributed Model (IHDM) [2], System Hydrological European (SHE) [3], and many others. However, these models are

rarely used for operational flood forecasting purposes [3,4]. Most of the hydrological processes are expressed explicitly in these models as it follows a “Bottom-up” approach where physical equations are applied in the form of a deterministic conservation equation for mass, momentum, and energy on the catchment spatial discretization then aggregated upward to the whole catchment [1]. The use of this continuum approach rarely gives results with quality that justifies the extensive demand of input information and computational load required at the applied spatial discretization [1]. As an alternative, conceptual models follow a “top-down” philosophy where each spatial discretization (cell) is treated as a single lumped catchment. Hydrological processes in conceptual models are expressed implicitly using state variables and parameters that might correspond directly to catchment physical properties. Thus, these parameters need to be obtained through calibration [1,5]. The top-down approach leads to a parsimonious interpretation of the system properties, e.g., TOPMODEL [9] and mHM [10]. Despite the previous classification of models, it is not easy to assign a model to any of these approaches, as not all the hydrological processes can be expressed purely with one philosophy rather than a combination of approaches [6]. The choice of a particular model type depends on the modeling objectives, data availability, and the desired modeling approach (large-scale deterministic approach [11], basin-scale approach [7], or site-specific approach [12]). In this study, a conceptual distributed model was chosen as a compromise between a lumped and fully distributed model. It consists of multiple HBV96 [13] lumped conceptual models in each cell, linked by routing models.

The process of building a distributed hydrological model is considered to be a series of decisions. First is the model type, which in most cases, is chosen based on the objective of the model, and in this study, it was chosen to be a conceptual distributed model. Second is the model structure and spatial discretization. Third is the parameter specification approach to improve the parsimony of the model (by limiting the number of parameters to be calibrated or measured) [2,14].

The spatial discretization required for catchment modeling is an uncertain area of the modeling process, many argue that it depends on dominant hydrological processes, modeling objectives, and available data [5]. Small landscape features such as agricultural fields can be represented explicitly in small catchments, and on the other hand, in large catchments, such fine representation is not feasible. Therefore, simplification or aggregation of parameters for specific discretization is necessary [15,16].

Dehotin in 2008 [17] stated that the choice of process conceptualization and resolution of spatial discretization are two complementary elements, as the latter allows the separation between processes and determine whether the processes can be represented explicitly or parameterized (simplified) and that the conceptual/physical representation of a process that worked for one resolution may not work for another resolution.

Due to the landscape heterogeneity and the hierarchical nature of the hydrological network, several types of discretization have been proposed, such as subcatchments [13], regular network [18] and HRUs (hydrological response units) [7]. Wood et al. (1988) [19] applied a concept of representative elementary area (REA) that defines subcatchments of areas that allow for average representation of the hydrological process similar to representative elementary volume (REV) in porous media; however, [15] disputed the robustness of the method, especially given that it does not consider the hierarchical structure of the river network, which was taken into account by [1,14,16,20] when they proposed the representative elementary watersheds (REWs).

The HBV model structure has been subjected to many changes since Bergström [21] proposed four different model structures, differing in the number of runoff components, with one response reservoir, two linear response reservoirs (surface and base flow), with the last two models having two reservoirs with two linear components (surface flow and interflow) in the upper reservoir beside the baseflow reservoir. Lindström [13] later replaced the two linear responses of the upper zone with one nonlinear response.

Many researchers have changed the structure of the conceptual model and coupled it with a routing model to be adapted in a conceptual distributed model. Ref. [22] used the concept of “Flex-Topo” [23] to build a distributed model of spatial elements, “Fields”, resulting from the intersection of subcatchment layer with the HRUs layer; each HRU has the same parameters and HBV model structure, and calculations are performed and updated for each field (different state variables), and flow from each subcatchment is routed to the outlet using a lag function used by [13]. Mayr et al. [24] used HBV-ETH [21,24,25] to build a distributed model that improves the representation of snowmelt in highly glacierized basins. Furthermore, they adjusted the model structure to calculate the accumulation in each cell using the cell curvature. The runoff from each cell is then summed without routing to obtain the total catchment flow. Boia [4] combined saturation excess runoff and infiltration excess to form the runoff generation model and coupled it with a 2D kinematic wave model for overland flow to solve the problem of catchment sinks storing water, and a 1D kinematic wave model for channel routing.

In this research, we aim at exploring a distributed modeling system to improve the understanding of the difference in performance between different grid discretization levels. This is done for a tropical river basin that is exposed periodically to extreme events. We also aim to provide a clear identification of the changes in discretization, and to understand if the change in performance with discretization is due to the lack of a well-defined routing scheme. We explore two methods and benchmarked the errors obtained in a validation process. The Jiboa River Basin in El Salvador is used as a case study.

The paper is organized as follows: the following section describes the study catchment; Section 3 presents the methodology and experimental setup. Section 4 includes the results and discussion. The last section presents conclusions.

2. Case Study

The study catchment is in the Jiboa River in the central part of El Salvador. The catchment (Figure 1) spread over 432 km² with the discharge gauge Puente Viejo locates at the outlet of the catchment at (13.52° N, 88.98° W). The catchment consisted of four topographical areas: San Vicente Volcano with an elevation of 2171 m a.s.l (above sea level) and very steep sides, Ilopango Lake (covers an area of 70 km²), Balsamo mountain range area with an elevation of 1000 m a.s.l, and the coastal plain area with elevation ranges from 0–100 m a.s.l.

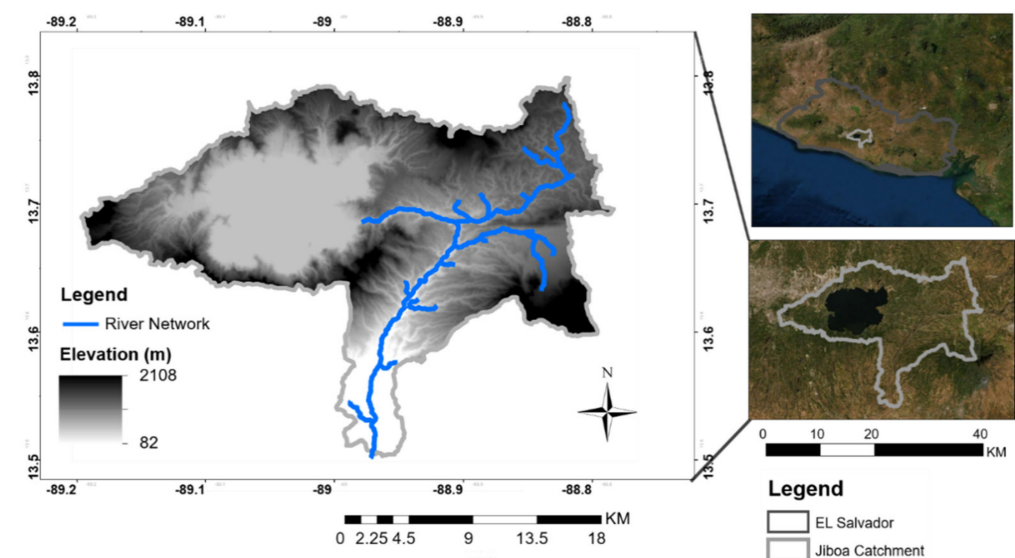


Figure 1. Location map for the Jiboa Catchment.

3. Methods

3.1. Hydrological Model Setup

This section describes the general methodology followed in this research for developing and testing the distributed models. First, we describe the approach for discretizing the landscape and connecting multiple model elements. Second the hydrological processes conceptualization in the model structure and the iterative process to improve model representation of the fast runoff response. Later we define the model parameters and parameterization approaches.

3.1.1. Hydrological Grid-Based Model “Hapi”

This work is based on an extended version of the HBV-96 model structure that provides grid-based conceptual integration of processing river basin. The structure is based on the mass conservation principle keeping the hydrological process as the main variables (Equation (1))

$$P - E - Q = \frac{d}{dt}(SP + SM + UZ + LZ + L) \quad (1)$$

where P is the precipitation, E is the actual evapotranspiration, Q is the discharge, SP , SM , UZ , LZ are the snowpack, soil moisture, upper zone, lower zone, and lake state variables, respectively, all in mm/time step, Equation (1) in its discrete form is represented in the Hapi model like in the following form (Equation (2)):

$$P_{i,t+1} - E_{i,t+1} - Q_{i,t+1} = f\left(\frac{E}{Et}(SP_{i,t} + SM_{i,t} + UZ_{i,t} + LZ_{i,t} + L_{i,t})\right) \quad (2)$$

where i , $t + 1$ refer to the cell at time step $t + 1$, and f is a routing method (Maxbas or Muskingum).

3.1.2. Spatial Discretization and Spatial Connectivity

Two types of models with different spatial discretizations were considered in this study. In the first type, the lake is represented at the outlet of one subcatchment, and the rest of the catchment is represented as another subcatchment. Semi-distributed model structure is used. Each subcatchment is assigned mean areal forcing data and different parameters; state variables evolve independently and do not interact with the state variables of the other subcatchments. The calculated runoff by each subcatchment is routed directly to the outlet using a Maxbas triangular function [13] (Figure 2). (This semi-distributed model of type 1 is used as a benchmark model with which the two distributed model variants of the second type are compared).

The second type of model uses spatial discretization on a regular grid (raster-based), where the Hapi Hydrologic framework [26] is used to build two variants of conceptual distributed models. Hapi uses multiple geo-registered layers of raster data sets to describe the drainage basin in order to perform the hydrologic computations and pixel-by-pixel routing in the correct order following the drainage network to the outlet. Pixels are assumed to be laterally disconnected from each other, so state variables are uncoupled and change independently.

Figure 2 presents the two considered variants (frameworks) of raster-based models of the second type. The first one, FW-1 (framework 1) (Figure 2A), routes the runoff from each cell directly to the outlet using Maxbas function, each cell has a Maxbas value depends on its flow path length; discharge at the outlet is obtained by summing up the routed runoff from all cells. In the second raster-based model, FW-2 (Figure 2B), runoff is routed from one cell to another following the flow direction input (see Section 3.1.4) using the Muskingum method (see Section 3.1.3). For both FW-1 and FW-2, only superficial discharge (surface runoff and interflow generated from upper zone UZ reservoir in the HBV model structure) is routed, while flow generated from the lower zone LZ is assumed to be lumped for the whole catchment (see Section 3.1.3).

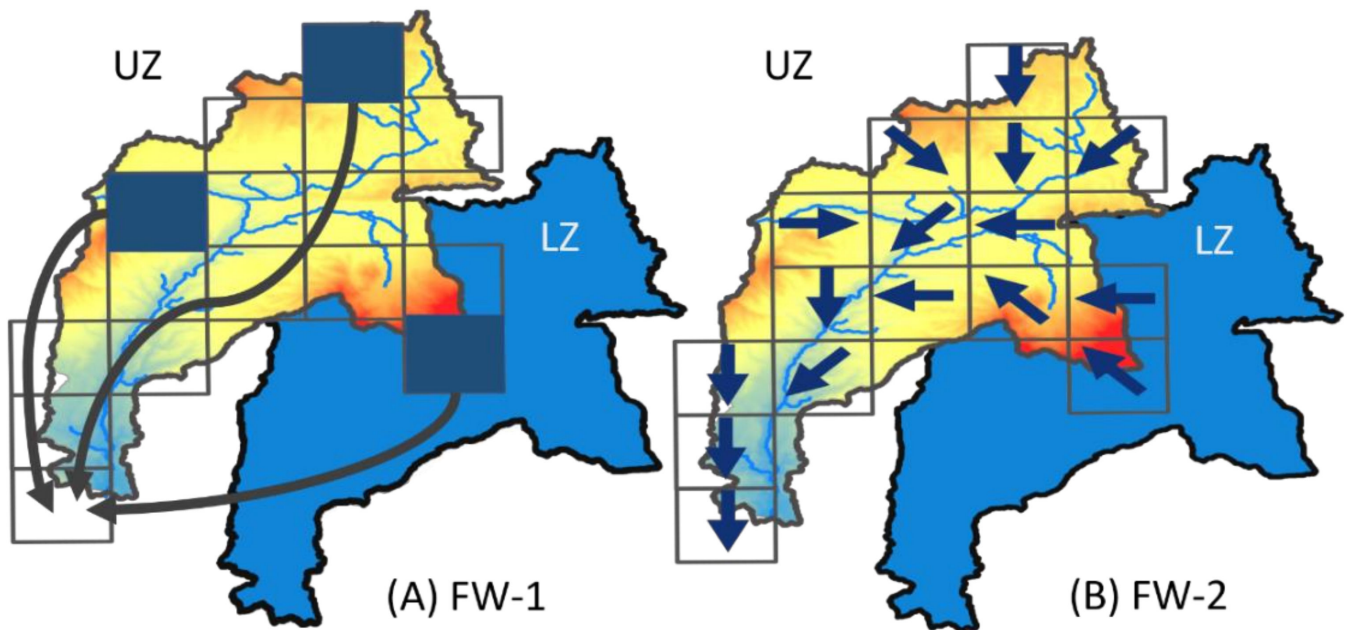


Figure 2. Two variants of raster-based conceptual distributed models (of type 2): (A) FW-1 routes the discharge directly to the outlet using Maxbas function, and (B) FW-2 routes the discharge from one cell to another following the river network using the Muskingum method.

3.1.3. Model Structure

Hydrological processes within the adopted spatial discretization (subcatchment and cell) are represented using the same lumped HBV model structure [13]. HBV consists of three tanks (soil moisture, upper response, and lower response) without considering the snow subroutine process (as followed in this study). For each spatial discretization element (cell), infiltration, soil moisture, percolation, and runoff generation processes are calculated. The runoff process is represented using a nonlinear function of actual soil moisture and precipitation in the upper reservoir; the nonlinear function simulates the surface runoff and interflow. The baseflow is simulated using a linear function of the lower reservoir tank. Both reservoirs interact in series by constant capillary rise and percolation. The primary parameters of the HBV are shown in Table 1, with the range of values used in the calibration for each parameter. More details on the HBV model can be found in [13].

Several trials are made to adapt the model structure following an iterative calibration process to understand better which hydrologic process needs to be improved in the model structure. In the iterative process, we either modify the conceptualization of the processes or change the parameter range of values used in the calibration process using OAT sensitivity analysis (see Section 3.2).

Following the iterative process, it is noticed that the model significantly underestimates the whole flood season, which indicates the poor representation of the fast runoff response of the model. The interaction between soil moisture and upper zone tanks has to be regulated differently in order to increase the sensitivity of the model structure to the catchment fast response. Both soil moisture and upper zone tanks are connected through recharge (Equation (3) and Figure 3) and capillary rise; by checking the sensitivity of the runoff to parameters of both processes, it was found that model performance (RMSE error) is susceptible to a slight change in the value of beta.

$$\frac{R}{IN} = \left(\frac{SM}{FC} \right)^\beta \tag{3}$$

where R is the recharge from soil moisture tank to upper zone tank, IN is the infiltration to the soil moisture tank, SM is the soil moisture storage in mm, FC is the max soil moisture storage capacity, β is a nonlinear runoff parameter that conceptualizes the degree of soil saturation and its effect on the recharge response.

Table 1. HBV parameters and the ranges used in the calibration process.

Parameter Name	Description	Units	Lower Bound	Upper Bound
RFCF	Precipitation correction factor	-	0.93	1.3
FC	Maximum soil moisture storage	mm	150	500
Beta	Nonlinear runoff parameter	-	0.01	5
ETF	Evapotranspiration correction factor	-	0.0	1.25
LP	Limit for potential evaporation	%	0.1	0.55
CFLUX	Maximum capillary rate	mm/h	0.05	0.55
K	Upper storage coefficient	1/h	0.00055	0.008
K1	Lower storage coefficient	1/h	0.0035	0.012
A	Nonlinear response parameter	-	0	0.3
Perc	Percolation rate	1/h	0.15	0.7
Clake	Lake correction factor	%	0.85	1.15
Maxbas	Transfer function length	h	1	7
K	Travel time	h	$\Delta t/2(1 - X) < K < \Delta T/2X$	
X	Weighting coefficient	-	0	0.5

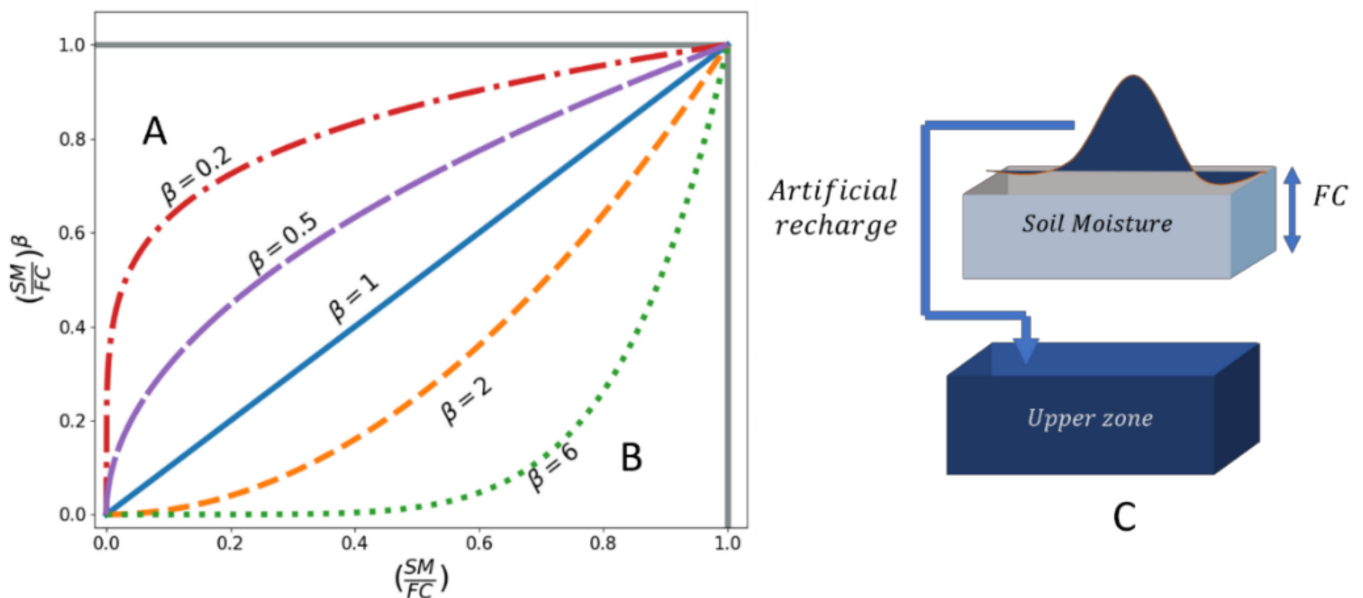


Figure 3. Elements of the HBV Structure: (A) Convex soil storage behavior (beta less than 1—model structure-2), (B) Concave soil storage behavior (beta more than 1—model structure-1), (C) Artificial recharge (model structure-2).

Schumann [27] mentioned the different interpretations and responses that model structure exhibit just by adopting a different range of values for β . In [27], it has also been stated that during a rainfall event, soil moisture tank would be filled up rapidly; recharge rate will be high, and that could be represented using Equation (3) with a convex behavior of soil storage capacity (beta is less than 1 (Figure 3A)). In contrast, using the same equation with a concave behavior (beta is greater than 1 (Figure 3B)) will indicate that the part of the basin with full soil storage (saturated) during rainfall event is small, and consequently, that the recharge is small.

In the calibration iterative process, while using the convex behavior of soil moisture storage (beta less than 1, which will be referred to as “Model Structure-1”) with the distributed model, the beta value resulting from the calibration is minimal (less than 0.1). Very small beta values (power coefficient almost zero) mean that the soil moisture subroutine’s effect is very minimal (all water that comes to the soil tank goes directly to the upper zone tank). This behavior of passing all the water to the upper zone indicates that the model is trying to generate more runoff to overcome the runoff underestimation, and it has already reached its highest capability (beta is almost zero), and, still, model performance is not good.

The concave soil moisture behavior is adjusted and used in the distributed model along with another rule that passes water to the upper zone once soil moisture exceeds the field capacity FC (Figure 3C) as an artificial recharge which will be referred to as “Model Structure-2”. This rule in the model structure forces the calibration algorithm to give smaller FC values to increase the artificial recharge, which in return increases the subsurface runoff and makes the simulated hydrograph closer to the observed during the rainy season.

Based on the iterative process of adopting a hypothesis about the catchment behavior (convex or concave behavior) then check the adequacy of the hypothesis in the model performance and visually inspect whether the expected shape of the hydrograph is obtained, the semi-distributed model will be implemented using both Model Structure 1 and 2, and it will be the base case, while Model Structure 2 will be used for both FW-1 and FW-2.

The lake is simulated explicitly with Model Structure-1 using the storage–discharge curve relationship described by [13,28]. The outflow from basins upstream of the lake is summed and used as an inflow to the lake (represented as a separate tank with the lake storage as a state variable). The new lake storage is then updated with the recent inflow, direct lake precipitation, and evaporation, the outflow from the lake can then be obtained using the storage–discharge curve.

The superficial discharge generated from the upper zone tank in Model Structure-1 and 2 is assumed to be distributed, and computations are performed on a cell scale. The runoff is routed either directly to the outlet using Maxbas Triangular function as in FW-1, or routed from one cell to another following the flow direction input (see Section 3.1.4) as in FW-2. In the FW-1 model, only one Maxbas parameter value is used as an input to the model regardless of the number of cells, and this value is assigned to the cell with the longest flow path length, and Maxbas values for all other cells are calculated proportionally based on the flow path length of each cell (Figure 4).

The Muskingum [29] routing method is used for routing the flow from cell to cell in FW-2, as this method has the ability to alter the coefficient for each cell at every computational time step (and, as observed in [30], it also requires limited data for the computations, but with modern computers, this becomes not very relevant). The Muskingum method employs the continuity equation to predict the magnitude, volume, and temporal patterns of flow as it translates downstream of a channel (Equations (4) and (5))

$$I - Q = \frac{\Delta S}{\Delta t} \tag{4}$$

$$\Delta S = K * [X * I^m + (1 - x) * Q^m] \tag{5}$$

where I , Q are Inflow and Outflow respectively, Δt is the time step, and ΔS is the storage. K is the travel time, X is a weighting coefficient to determine the linearity of the water surface, x values are between 0 and 0.5, and m is an exponential constant varying from 0.6 for the rectangle channel to 1 [31]. In this study, we use the linear Muskingum version, which uses m equal to 1.

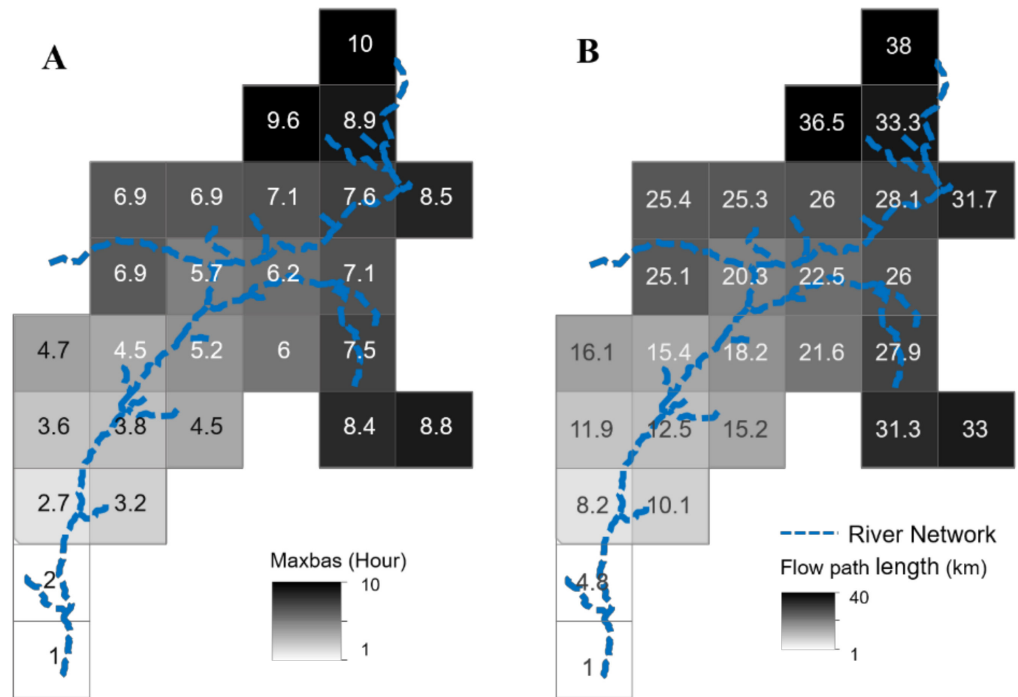


Figure 4. (A) Flow path length input to the FW-1 model, and (B) the model Maxbas values calculated for each cell.

The Muskingum equation can be written as (Equations (6) and (7)):

$$Q_{j+1} = C_0 * I_{j+1} + C_1 * I_j + C_2 * Q_j \tag{6}$$

$$C_0 = \frac{\Delta t - 2KX}{2K(1 - X) + \Delta t}, C_1 = \frac{\Delta t + 2KX}{2K(1 - X) + \Delta t}, C_2 = \frac{2K(1 - X) - \Delta t}{2K(1 - X) + \Delta t} \tag{7}$$

where $J + 1$ and J denote the present time step and the previous time step, respectively, C_0 , C_1 , and C_2 are three coefficients that can be calculated using K and X (Equation (7)). The main difference between the Muskingum–Cunge method and the Muskingum method is that the estimation of the parameters K and X [29] derives an expression to calculate each coefficient; however, in this study, both parameters K and X are freely calibrated parameters.

Based on the range of values for parameter X and K , and the temporal resolution Δt (hourly time step), the value of the coefficient C_0 might be negative, Mishra 2001 [32] suggested that negative values of C_0 should be avoided in Equation (8) to accomplish numerical stability. The time step Δt should be small enough to achieve the assumption of linearity of flow rate over the timestep, especially if the response time of the catchment is short [33]. To keep the numerical method theoretically stable, another constraint was suggested by Mishra and Singh, (2001) [32], as described in Equation (9)

$$\frac{\Delta T}{K} > 2X \tag{8}$$

$$\Delta t < 2K(1 - X) \tag{9}$$

Both K and X are obtained through optimization by defining a possible range of values for the optimization algorithm to search for the optimum values, by defining Equations (8) and (9) as boundaries in the calibration process, the feasible search space for parameter K and X will be as shown in Figure 5.

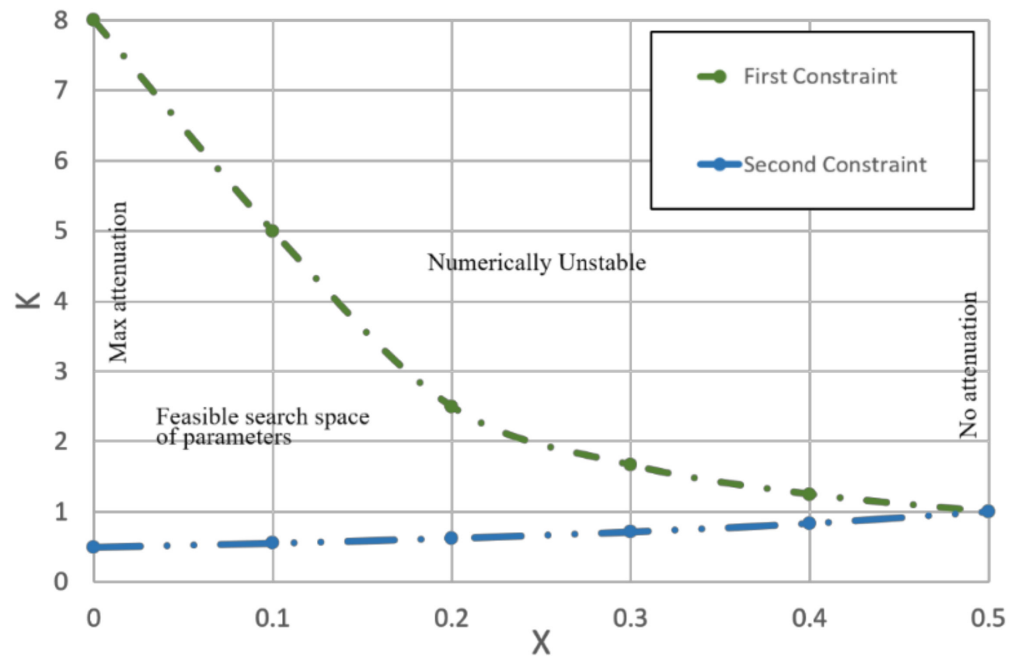


Figure 5. Feasible domain for Muskingum parameters K and X parameters based on the stability conditions.

In this study, the Hapi model [26] was employed (the source code is available in GitHub); it is a framework for building raster-based conceptual distributed hydrological models where computations are performed on a pixel-by-pixel basis. It also enables multi-scale modeling and thus suitable for simulations from coarse to relatively fine resolutions. Meteorological data (precipitation, evapotranspiration, and temperature) are supplied in the form of gauge inputs or raster data. Hapi uses HBV-96 as a conceptual model to represent the hydrological processes within a cell, and the Muskingum method as a routing model between cells, with the potential of using any other conceptual hydrological model in cells instead of HBV-96.

3.1.4. Model Inputs and Outputs

In the models used, the model computations are performed on a pixel-by-pixel basis. Meteorological data (precipitation, evapotranspiration, and temperature) data can be processed in a raster form. A simple statistical formula like inverse squared distance weighting (ISDW) (Equation (10)) can be used to distribute the values of gauge data over different cell resolutions as was done in this study (15 gauges).

$$Z_{newpoint} = \frac{\sum_{i=1}^n \left(\frac{z_i}{d_i^2} \right)}{\sum_{i=1}^n \left(\frac{1}{d_i^2} \right)} \quad (10)$$

where Z is the gauge’s value, i is the number of gauges, and d is the distance between each gauge and the cell center.

Hapi uses multiple geo-registered raster data to describe the drainage basin. All the raster data required were derived from DEM, where it is used to delineate the catchment and subdivide it into two subcatchments (in the case of the semi-distributed model). A series of GIS modules were used to perform operations to generate these input data, such as pit removal to fill depressions in the DEM, flow direction, flow accumulation, river network, and flow path length (Figure 6).

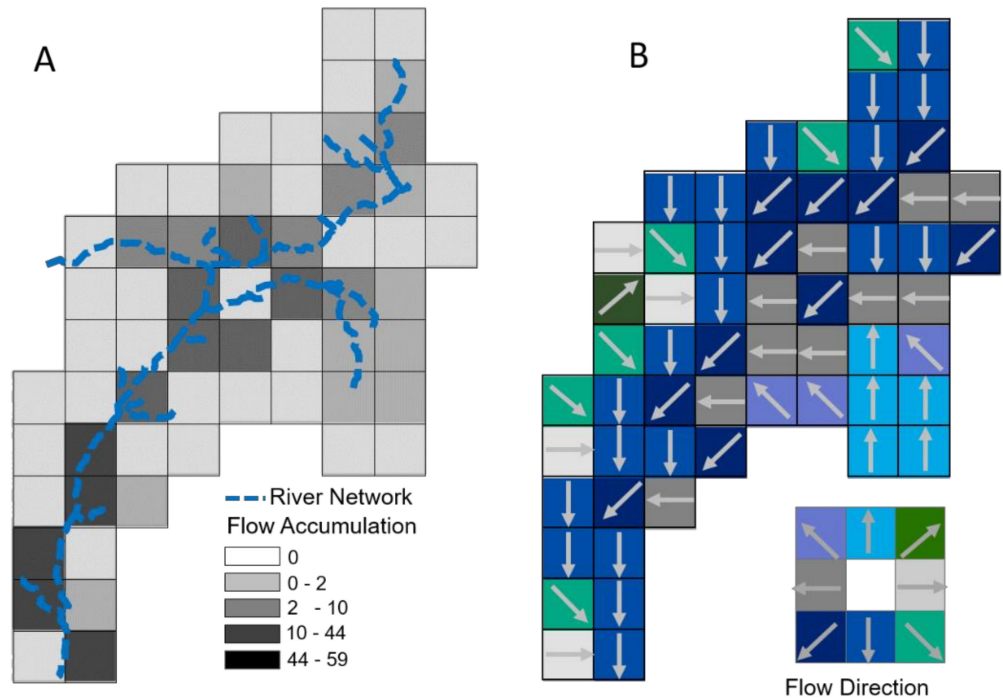


Figure 6. Raster input data for Hapi: (A) Flow Accumulation; (B) Flow Direction.

DEM is resampled many times in steps (not directly from 30 m to 4 km) to prepare the raster input data for different resolutions. To keep the topographical features of the river and flow direction in coarser resolution as close as possible to the original shape of the river network, DEM depressions are filled after each resampling.

3.1.5. Model Parameters

Two model structures are used in this study (Figure 3), being different only in the surface runoff generation process. Both model structures use ten parameters (Table 1) to describe the whole range of hydrological processes in the catchment (snow subroutine is excluded). Furthermore, additional parameters were considered for the lake and the routing methods (Maxbas and Muskingum K and X parameters).

3.1.6. Parameterization Approaches

Two simple parameterization approaches are followed in the distributed models in this study; the first approach assumes that all cells use the same set of parameters. The second approach assumes different parameter values for each cell, considering that in a large catchment, most of the hydrological processes change significantly with the change of land use or elevation. Lower zone response is considered uniform in all cells for both parameterization approaches.

3.2. Model Calibration

Meteorological input data are split into two sets with similar statistical properties for calibration and verification, respectively. The available data with good quality is two-and-a-half years long (June 2012 to November 2014), including two dry seasons and three wet seasons. As the model is built mainly for flood forecasting purposes, data on the two wet seasons is included in the training dataset (till December 2013), including the highest peak in the hydrograph.

Conceptual model parameters are not measurable, and thus calibration procedure has to be used to determine model parameters. The calibration process consisted of three steps: (1) Pre-calibration of the model using a genetic algorithm (GA); (2) Sensitivity analysis of the model to model parameters; and (3) Calibration of the model using harmony search (HS) algorithm (Figure 7).

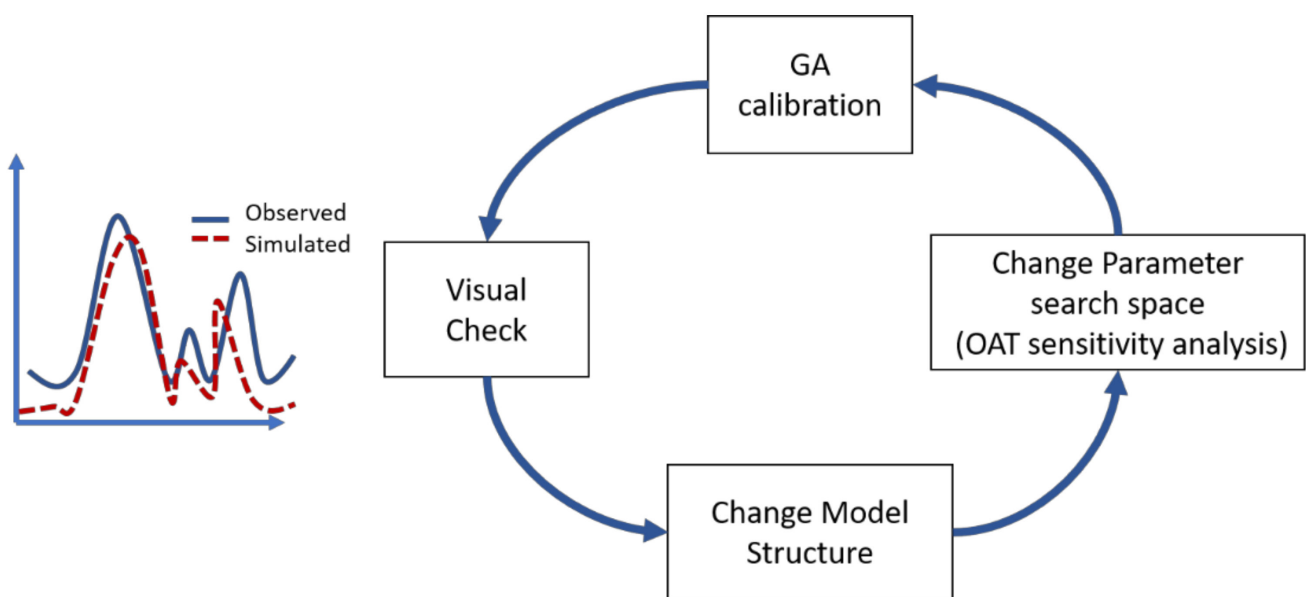


Figure 7. Iterative calibration process.

Starting from initial boundaries for each parameter to define the search space, GA runs for 30 generations with 200 populations to minimize overall RMSE as an objective function. Afterward, One-at-a-time (OAT) sensitivity analysis is performed to re-explore all parameters' search space and to redefine the boundaries for the most influential parameters to the model performance criteria. The interaction between parameters is checked using Sabol's concept by plotting relative values of all parameters in one graph [34] (Figure 8). The previous calibration steps are performed many times as an iterative process to define the model structure and the parameter range (see Section 3.1.5).

Choice of parameter boundaries is crucial for the calibration results, and initial choice may not be the best. After the GA runs for the entire 30 generations or converges in the middle, the evolution from one generation to another focuses the search space around good results. The GA may ignore parts of the multidimensional search space after a certain number of generations, trying to converge faster. Using the OAT sensitivity analysis re-explores the entire range of each parameter after reaching the optimum set of parameter values. Redefining the boundaries of each parameter to either narrow the range to make the algorithm feel the influence of a minimal change in one parameter value or to expand one of the boundaries of a parameter as the best value locates very close to the boundary (parameter K1-j). With the new search region, the randomized search may find a better result. (For example, model performance improves significantly with a very slight change in parameter Perc-j see Figure 8).

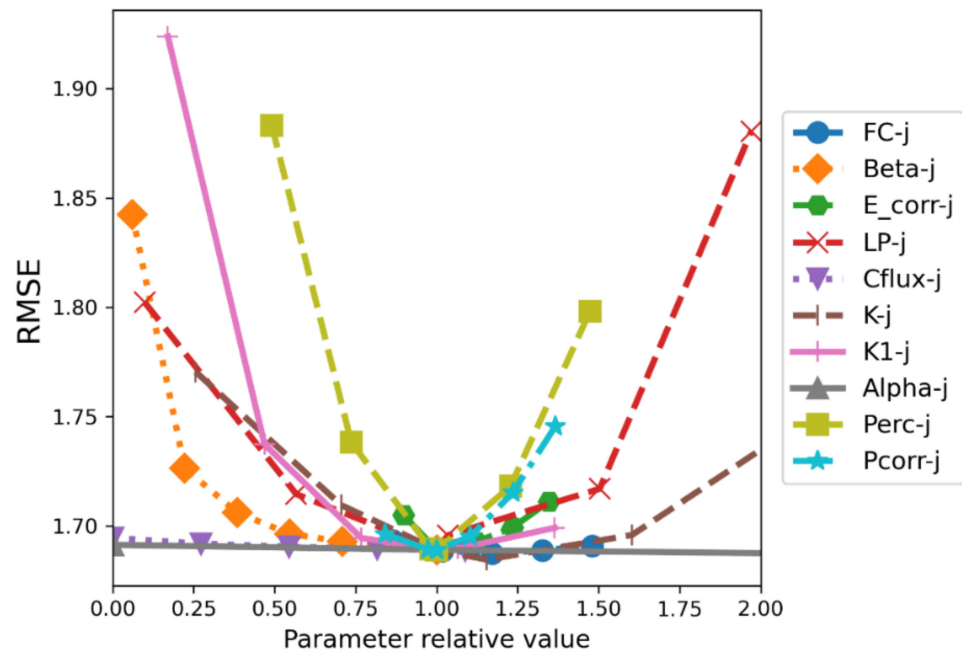


Figure 8. OAT sensitivity analysis with RMSE as evaluation criteria; the relative values of parameters are plotted in one graph.

This final step of the calibration is carried out using a different type of randomized search (HS) as the calibration algorithm with the new boundaries. HS generates a new vector of parameters from all the existing vectors in the harmony memory, while GA generates offsprings (new parameter sets) from only two parents [35]. Our experiments showed that such hybridization leads to better results than just using GA at both stages.

Model Performance Metrics

Model performance is assessed using traditional statistical goodness-of-fit metrics, such as root mean square error RMSE (Equation (11)) and Nash-Sutcliffe efficiency (NSE) (Equation (12)) [36]. RMSE is also used as an objective function for the calibration of all models. RMSE and NSE focus on the high values (peaks) in the time series and downplay lower values, while NSE with logarithmic values (NSEln) (Equation (13)) puts greater emphasis on low flow values. Furthermore, to estimate how much the model manages to reproduce the streamflow volume correctly, the mean cumulative error ‘WB’ is used (Equation (14)). Finally, the Kling-Gupta efficiency index (KGE) [37] is used as an aggregate metrics (Equation (15)).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_{obs_i} - Q_{sim_i})^2} \tag{11}$$

$$NSE = 1 - \frac{\sum_{i=1}^n (Q_{obs_i} - Q_{sim_i})^2}{\sum_{i=1}^n (Q_{obs_i} - Q_{avg})^2} \tag{12}$$

$$NSEln = 1 - \frac{\sum_{i=1}^n (\ln(Q_{obs_i}) - \ln(Q_{sim_i}))^2}{\sum_{i=1}^n (\ln(Q_{obs_i}) - \ln(Q_{avg}))^2} \tag{13}$$

$$WB = 100 * \left(1 - \left| 1 - \frac{\sum_{i=1}^n Q_{sim_i}}{\sum_{i=1}^n Q_{obs_i}} \right| \right) \tag{14}$$

$$KGE = 1 - \sqrt{(c - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \tag{15}$$

where Q_{obs} and Q_{sim} are the observed and simulated discharge at time t , respectively, Q_{avg} is the mean observed discharge, and n is the number of time steps. c is the cross-correlation between Q_{obs} and Q_{sim} , α measures the variability in the data $\alpha = \sigma_{Q_{sim}} / \sigma_{Q_{obs}}$, $\sigma_{Q_{sim}}$, $\sigma_{Q_{obs}}$ are the standard deviation in the simulated and observed discharge, respectively, $\beta = Q_{simAvg} / Q_{avg}$.

3.3. Data

The basin is instrumented by a telemetry system of 15-min time step rainfall gauges (Figure 9A), which provides data for the period between 2002 and 2015. Only one water level gauge at the catchment outlet is available and provides data between 2012 and 2016. For the calibration, the rainfall and water level measurements were aggregated to hourly time step, while the water level data was further converted into discharge using a rating curve equation) Equation (16) (and the equation coefficient are listed in Table 2. Due to gaps in the rainfall data, only the period between June 2012 to November 2014 is considered in the study.

$$Q = a(H - H_0)^b \tag{16}$$

where H is the water level in meters, H_0 , a , and b are coefficient changes based on the date of the water level value, and Q is the discharge in m^3/s .

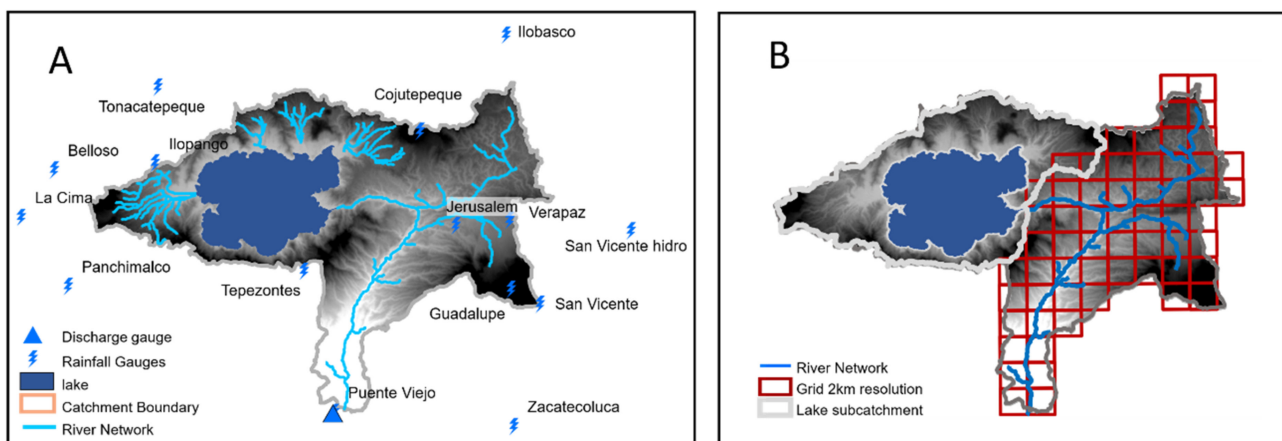


Figure 9. Jiboa Catchment. (A) Boundaries and rainfall gauge, (B) the spatial model setup, where the lake and the upstream area ‘Lake subcatchment’ are represented separately using HBV lumped model and the representation of the rest of the catchment changes from lumped to distributed with different spatial resolutions.

Table 2. Discharge rating curve coefficients for Puente Viejo station.

Period	H_0	b	A
05/2012–04/2013	0.15	1.89	13.20
05/2013–11/2013	0.16	1.90	13.27
11/2013–04/2014	0.02	3.00	15.45
05/2014–04/2015	0.15	1.25	13.46
05/2015–04/2016	0.18	1.73	20.04

Monthly potential evapotranspiration data were obtained for each gauge using a potential evapotranspiration–elevation relation derived by the Hargreaves method and linear regression [38]. Hourly temperature estimates were calculated from one available daily gauge (Cojutepeque) using sine function approximation, and were used for the whole catchment.

The lake is simulated using a storage–outflow relation, where storage changes from inflow, direct rainfall, and evaporation are updated at each time step, and the lake-outflow is calculated using the newly calculated storage. Inverse square distance weighting method

was used to calculate distributed data for different spatial resolution for all data required for the hydrological simulation (precipitation, temperature, potential evapotranspiration). All the data available provided by the Ministry of Environment and Natural Resources of El Salvador (MARN).

3.4. Model Setup

The lake’s existence in the catchment made it impossible to represent the whole catchment as one lumped model; therefore, the catchment is divided into two parts, the lake and the Jiboa subcatchments (Figure 9B). The lake and all the upstream areas to the lake are represented as one lumped model using Model Structure-1, referred to as a lake subcatchment. The rest of the catchment, which will be referred to as the “Jiboa subcatchment”, is simulated using different spatial discretizations (lumped or cells) as follows:

- Lumped model: two models are built using the lumped spatial discretization SemiDist-1 and SemiDist-2, the former is built using Model Structure-1, and the latter is built with Model Structure-2.
- Conceptual distributed model built with FW-1 considering lumped catchment parameters FW-1-L using different spatial resolutions (4 km, 2 km, 1 km, and 500 m).
- Conceptual distributed model built with FW-1 considering distributed catchment parameters (different parameter for each cell) FW-1-D using different spatial resolutions (4 km, 2 km, 1 km, and 500 m).
- Conceptual distributed model built with FW-2 considering distributed catchment parameters (different parameter for each cell) using different spatial resolutions (4 km, 2 km, 1 km, and 500 m)

The experimental model setup followed in this research, starting from the model type, spatial discretization, model structure, routing method, and parameterization approach, is shown in Figure 10. For model descriptions, see Table 3.

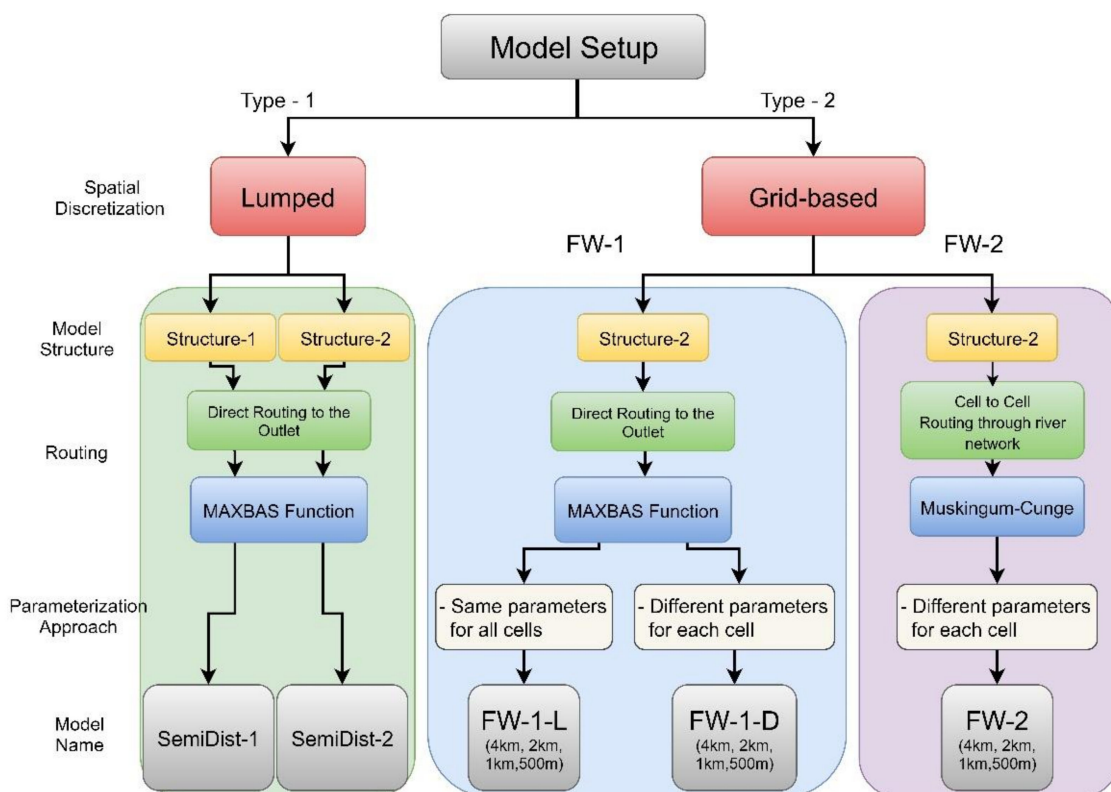


Figure 10. Model setup.

Table 3. Experimental setup.

Model Name	Model Description	Resolution
SemiDist	Two models are built using the lumped spatial discretization SemiDist-1 and SemiDist-2, the former is built using Model Structure-1, and the latter is built with Model Structure-2.	Semi-distributed
FW-1-L	Conceptual distributed model built with FW-1 and lumped catchment parameters FW-1-L.	4 km, 2 km, 1 km, and 500 m
FW-1-D	Conceptual distributed model built with FW-2 and distributed catchment parameters.	4 km
FW-2	Conceptual distributed model built with FW-2 and distributed catchment parameters.	4 km, 2 km, 1 km, and 500 m

4. Results and Discussion

Runoff dynamics in the catchment are significantly influenced by the lake, which is the reason subcatchments have been defined based on the outlet of the lake. Therefore, the semi-distributed model is used as the simplest and the base case in this study. However, the elevation differs a lot between the four topographical areas in the catchment, Temperature from only one gauge is available and used in the study, which might affect the performance of the models; however, the effect will be the same on all models, therefore the comparison between models would still be valid.

It is worthy of mentioning that we explored CMORPH rainfall satellite data with an hourly temporal resolution and $8 \times 8 \text{ km}^2$ spatial resolution; however, it shows very poor quality compared to ground-based gauges. The kriging method wasn't used due to the non-Gaussian nature and the skewness of the rainfall data, being discontinuous and having many zeros; the kriging method requires rainfall data to be transformed to the normal distribution, and for this reason, the BoxCox and normal score transform method were investigated but did not fully succeed at transforming the rainfall to normally distributed values. On the basis of our knowledge of the localized orographic pattern of rainfall at the east part of the catchment where there is a volcano, in addition to the convective rainfall pattern in the catchment, the ISDW method is preferred over the inverse distance weighting method (IDW) as it gives smaller weights ($1/d^2$ instead of $1/d$) with greater distances, which suits the rainfall nature in the area.

Table 4 presents the performances of all models. The semiDist-1 model shows better results, with an RMSE value of 1.33, than the SemiDist-2, with an RMSE value of 1.65. FW-2 4 km^2 model performance is slightly better than the performance of SemiDist-2, both having the same model structure, while it shows lower performance than SemiDist-1, the reason is clearly related to the conceptualization of direct runoff process. Another possible explanation, as stated by Das et al. (2008) [39], could be the overparameterization of the distributed models with high and low resolutions, and it could also be that averaging in space compensates for stochastic errors. A possible source of error that affects the performance of the models with different cell resolutions is that precipitation and evapotranspiration were obtained using the ISDW interpolation method (Equation (10)), and interpolation methods tend to have higher estimation error with higher resolution. Moreover, the generated values do not reflect the spatial variation of the precipitation or evapotranspiration; rather, they reflect the statistical method's inherent effect on gauges' values at different distances.

Table 4. Model performance statistics for all models during calibration and validation periods.

Cell Size (Km ²)	RMSE		Low Flow NSE		High Flow NSE		WB		KGE	
	Cal.	Val.	Cal.	Val.	Cal.	Val.	Cal.	Val.	Cal.	Val.
FW-1-L (4)	1.69	1.9	0.57	0.8	0.6	0.62	97.8	73.7	0.66	0.57
FW-1-L (2)	1.65	1.89	0.58	0.82	0.61	0.62	96.7	74.9	0.69	0.56
FW-1-L (1)	1.66	1.91	0.59	0.81	0.61	0.62	96.2	74	0.71	0.56
FW-1-L (500 m ²)	1.68	1.89	0.57	0.81	0.6	0.62	96.1	74.1	0.68	0.57
FW-1-D (4)	1.67	1.84	0.54	0.82	0.61	0.65	94.31	74.74	0.71	0.59
FW-1-D (2)	1.63	1.9	0.59	0.81	0.63	0.62	97.08	73.03	0.7	0.55
FW-1-D (1)	1.66	1.94	0.58	0.82	0.61	0.61	97.84	72.5	0.68	0.54
FW-1-D (500 m ²)	1.69	1.88	0.53	0.82	0.6	0.63	96.24	73.49	0.69	0.56
FW-2 (4)	1.6	1.95	0.62	0.81	0.64	0.6	99.48	72.18	0.69	0.53
FW-2 (2)	1.67	1.98	0.61	0.8	0.61	0.58	98.79	71.54	0.67	0.53
FW-2 (1)	1.76	1.98	0.58	0.81	0.56	0.59	97.8	72.57	0.65	0.54
FW-2 (500 m ²)	1.84	1.97	0.53	0.82	0.53	0.59	95	73.53	0.63	0.54
SemiDist-I	1.33	1.19	0.64	0.87	0.75	0.85	0.83	0.92	99.31	97.45
SemiDist-II	1.65	1.72	0.59	0.82	0.62	0.69	0.71	0.65	99.49	85.1

The performance statistics of all models for the calibration period (June 2012 to December 2013) are presented in a normalized form in Figure 11. FW-2 4 km² distributed model has the best performance among all distributed models during the calibration periods, with an RMSE value of 1.6, Nash-Sutcliffe value of 0.64, and a value of 0.62 for logarithmic Nash-Sutcliffe. RMSE for all models of FW-1 ranges between 1.63 and 1.69, whereas for FW-1-L models, RMSE ranges between 1.65 and 1.69, and for FW-1-D RMSE ranges between 1.63 and 1.69, as shown in Table 4. The simulated hydrograph with FW-2 4 km² is benchmarked to the gauge hydrograph in Figure 12.

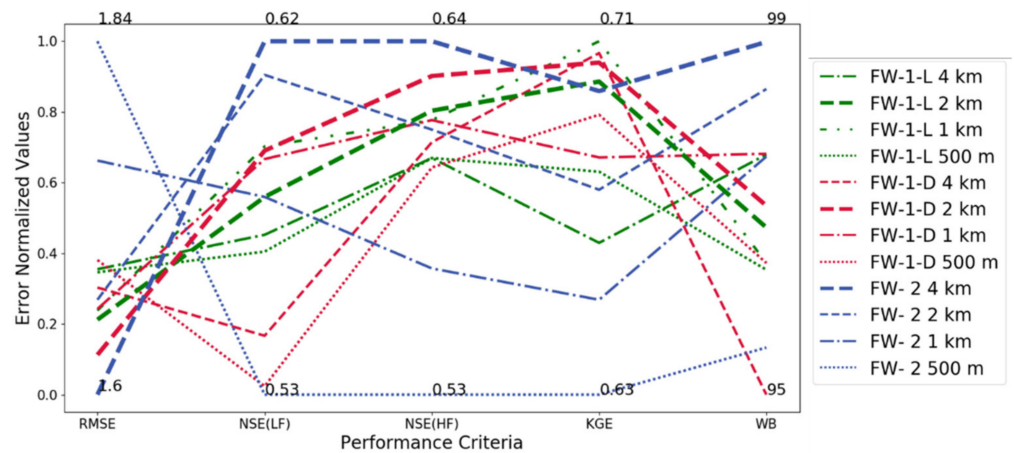


Figure 11. Normalized performance criteria for the calibration period for all distributed models. The minimum and maximum error values are presented at the top and bottom of the figure. γ -axis represents the percent of model error between the maximum and minimum values.

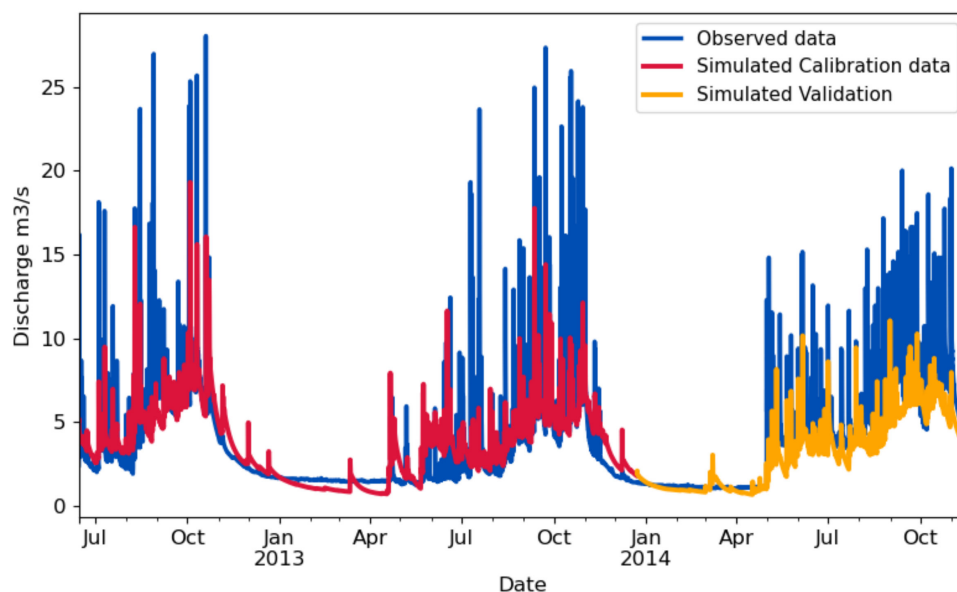


Figure 12. Flow hydrograph result from FW-2 4 km².

For the optimal set of parameters obtained by calibration for both SemiDist-1 and SemiDist-2 model, see Table 5.

Table 5. Optimal set of parameters for Jiboa subcatchment for the SemiDist-1 and SemiDist-2 models.

Model	RFCF	FC	B	E_corr	LP (%)	CFLUX (mm/h)	K (1/h)	K1 (1/h)	A	Perc (mm/h)	MAX-BAS
SemiDist-1	1	547.7	0.0443	0.589	0.071	1.421	0.0053	0.0063	0.096	0.304	3
SemiDist-2	1.07	359.3	4.19	0.942	0.045	0.333	0.0043	0.0062	0.020	0.438	5

Figure 12 shows the simulated and observed flow hydrograph using FW-2 4 km².

On average, the medium and low flow are well estimated, considering that low flow is dominated by the outflow from the lake during the dry season, and it is slightly affected by the spatial representation of the Jiboa subcatchment. Peak flows are underestimated, as shown in Figure 13. FW-1-L 2 km² and FW-1-L 1 km² are the closest to the observed hydrograph in terms of the magnitude and shape of the peak (rising and falling limb), FW-1-D models with resolutions of 4 km² and 2 km² comes after in capturing the highest peak, whereas FW-2 (4 km²) comes next. FW-2 4 km² has the highest NSE and NSE(Lf), which represents the model performance to reproduce high and low flow values, respectively, followed by FW-1-D 2 km² model then FW-1-L 1 km² (Figure 11 and Table 4).

The results regarding the effect of increasing the number of parameters on model performance from lumped to distributed parameters are as expected (Figure 14): FW-1-L has lower performance statistics (higher error) than FW-1-D; however, the difference in performance is minimal and using lumped parameters leads to better results with 500 m² resolution than that with distributed parameters (Figure 14). Even though the calibration process consists of three independent steps and two different calibration algorithms are used to obtain better parameters and not to end in a local minimum, with more model parameters, there is a need for more iterations during the calibration process. Besides, an increase in the number of parameters with finer resolutions, and the fact that overestimation of flow in one part of the catchment can be compensated by underestimation of the flow in another part; this is because only one flow gauge at the outlet is used for the calibration. These points could be the reason for the slight difference in model performance with an increase in the model parameters number.

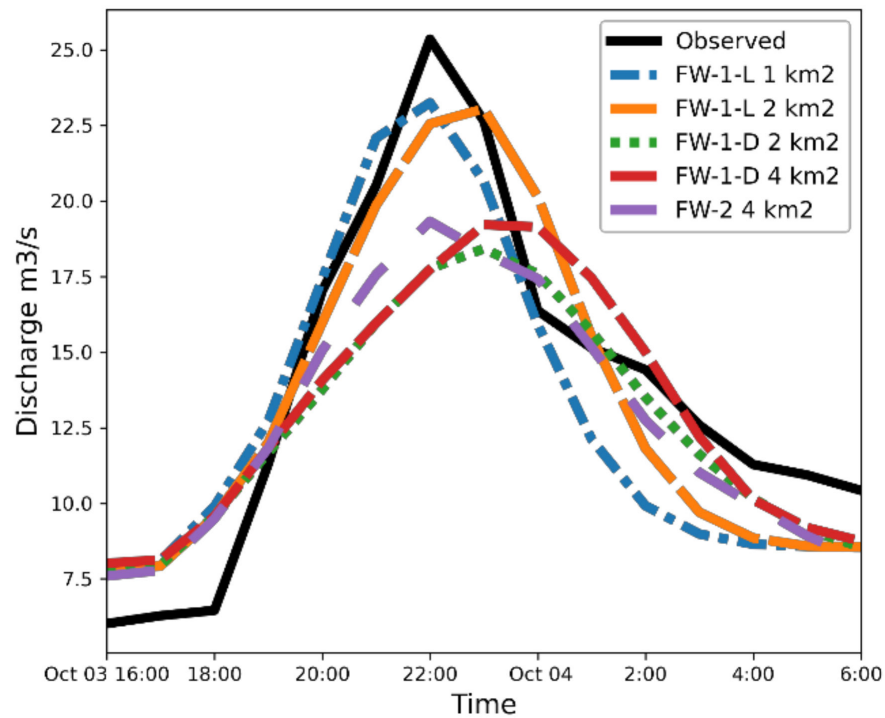


Figure 13. Models’ performance in reproducing the highest peak during the calibration period.

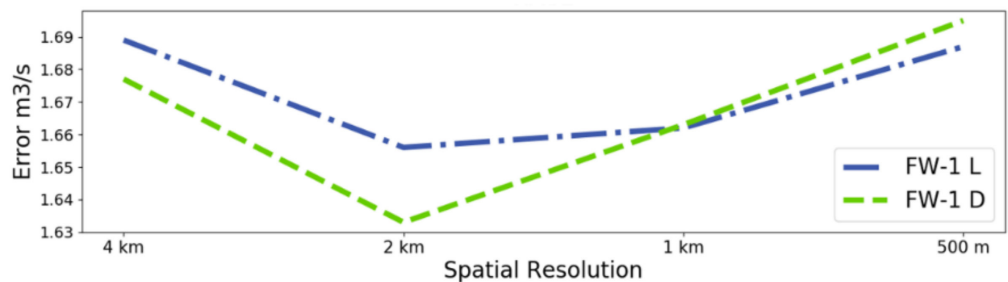


Figure 14. The effect of considering distributed parameters with FW-1 on model performance (RMSE).

The decrease in performance with the increase of the spatial resolution confirms the hypothesis by [40–43] that complex process interaction at small scales can be represented by simple approaches at a larger scale due to the self-organizing capacity of large systems. Therefore, the more we use finer resolution with the conceptual model, the more we lose the self-organizing capacity, and at a certain scale, we should switch to more complex models. Both FW-1-L and FW-1-D models show similar behavior: error decreases with the increase in model resolution to reach the lowest error value (RMSE 1.63) at a resolution of 2 km², and then it starts increasing again (Figure 14).

It was expected that increasing the complexity of the model would lead to overfitting. However, considering that both calibration and validation datasets are statistically different, and the validation dataset has an even higher error with the semi-distributed model (Table 4), a slight increase in error corresponding to an increased spatial resolution does not indicate overfitting, especially given that it decreases at a resolution of 500 m² (Figure 15).

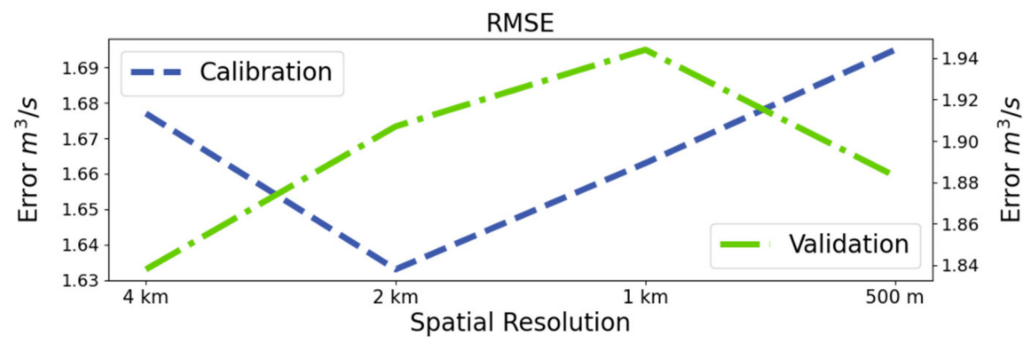


Figure 15. RMSE values of FW-1-D during calibration and validation periods.

Based on the nonlinear constraints that govern the values of the routing parameters K and X (Figure 5), the point in the feasible parameter space that gives minimum attenuation (X equals 0.5) corresponds to a 1-time step of translation (K equals 1). The point that gives the minimum translation (K equals 0.5) corresponds to the maximum attenuation (X equals 0). For fine spatial resolutions, the flood wave is assumed to travel between cells with minimal attenuation and small translation, which conflicts with the feasible space of parameter values created by the constraints. Therefore, for 500 m cell resolution, flood wave is supposed to be as close as possible to remain unattenuated (X equals 0.5), and traveling time should be around 0.13 h ($500\text{ m} / (1 \times 60 \times 60)$), assuming flood speed of 1 m/s). However, these values are outside the feasible search space of K and X to maintain the stability of the method with the given time step (hourly time step), so the only solution for this drawback is to use coarser cell resolution considering that using a finer time step is not always possible. Figure 16 shows the effect of Muskingum parameters on the highest flood wave simulated using different cell sizes. Using small spatial resolutions led to more attenuating and translation of the hydrograph, which in return results in a decline in the model performance with smaller cell sizes. The previous trend between cell size and model performance shows the strong influence of the routing method and cell size on the accuracy of the model predictions.

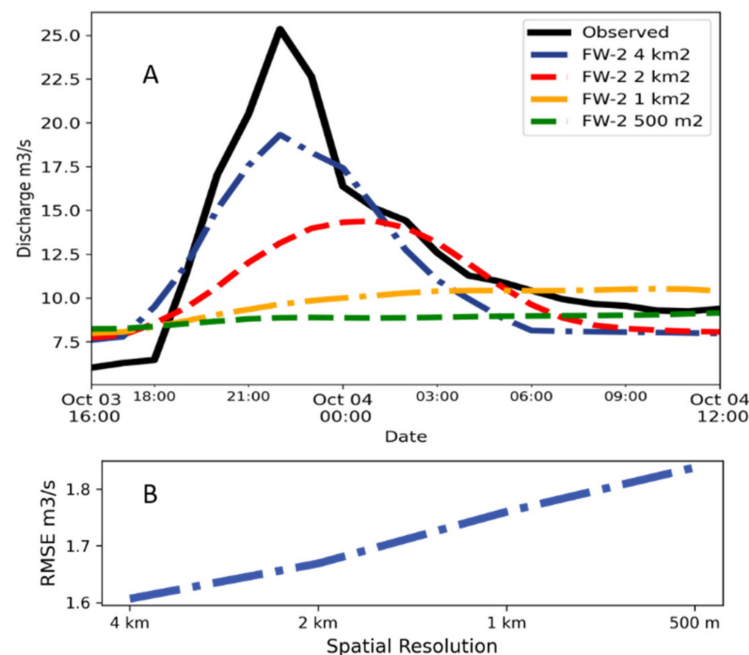


Figure 16. Effect of the Muskingum routing parameters (K and X) for different cell resolution on (A) the highest flood wave, and (B) model performance.

It is important to stress the following: the lumped representation of the whole catchment, including the lake (implicitly represented in the lower reservoir and not at the outlet of the catchment), did not capture the catchment baseflow response. Thus, it is very important to define a separate state variable to reflect the lake characteristics (lake storage, water level, or lake surface area). The lake state variable is updated based on the interaction between the lake and other hydrological responses (inflow, direct rainfall, evaporation, and outflow). Modeling the lake implicitly without this state variable would lead to a misinterpretation of both lower response from the lower reservoir and outflow from the lake. Therefore, explicit representation of the lake, which requires dividing the catchment into sub-basins with the lake being at the outlet of one of the sub-basins, is the simplest way to simulate the catchment. Consequently, the semi-distributed model structure is used as a base case for this study.

The lake is a very dominating hydrological feature, where base flow in dry seasons comes mainly from the lake's outflow (Figure 17). Using a rating curve (storage–outflow relation) in lake outflow calculation overwrites all the upstream hydrological processes and gives all the importance to the initial lake volume at the beginning of the simulation; the initial lake volume cannot be corrected by considering a warm-up period at the beginning of the simulation.

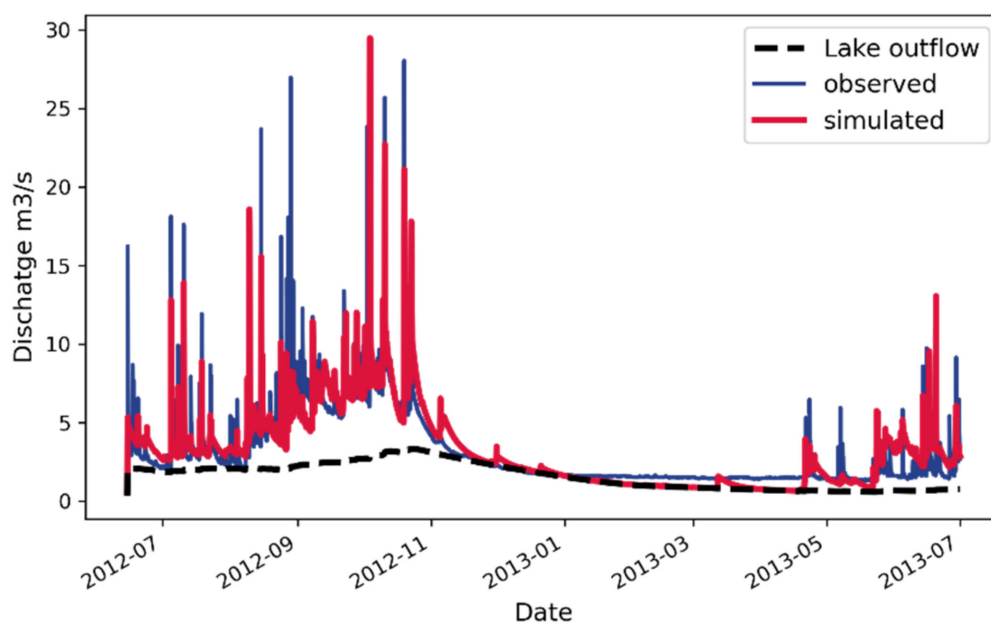


Figure 17. Contribution of the lake outflow in the whole hydrograph at the outlet.

5. Conclusions

In this study, we studied the performance of several hydrological (semi-distributed) model structures on an example of modeling the Jiboa River Basin in El Salvador. Two routing models are considered. The second concept is the spatial representation, where the model is built and tested for different spatial resolutions (500 m, 1 km, 2 km, and 4 km). The models were calibrated using a combination of two randomized search algorithms, complemented by the sensitivity analysis.

The results allow us to conclude that the spatial sensitivity of the resolution is highly linked to the routing method; the routing sensitivity influenced the model performance more than the spatial discretization, allowing for coarser discretization makes the model simpler and computationally faster. Slight performance improvement is gained by considering spatial catchment characteristics. It was found that the 2 km cell size corresponds to the least model error values.

The results show the same trend of error regarding increasing the resolution of spatial discretization from 4 km² to 500 m² for FW-1 with both lumped parameters and distributed parameters, while it shows a decrease in all performance criteria with increasing the resolution with FW-2 due to the high attenuation of peaks that Muskingum method performs with fine resolutions.

There are also limitations to this study. Only a single catchment is considered, so one should be careful with generalizations; only one model structure (HBV) was evaluated for model components. No comparison with other model structures of the same class was carried out. Propagation of uncertainty from inputs and parameters to the modeling results was not studied (albeit model sensitivity to parameters has been explored). The main forcing data (precipitation, evapotranspiration, and temperature) are interpolated from gauge data using the ISDW method for all resolutions. Due to the complexity of the distributed model parameter interaction and dependency, the calibration process is a difficult task and is highly affected by the initial limits of the parameters. Therefore, the genetic algorithm GA is used in an iterative approach to modify the model structure and to define the possible range for each parameter.

Future work will focus on uncertainty analysis of the built models and extending the Hapi model with more model formulations used in the model cells.

The Hapi modeling framework for building raster-based conceptual distributed hydrological models is open-source and is available on GitHub [26].

Author Contributions: D.S. and G.C.P. conceptualized the study, M.F. developed the Hapi algorithm, calibrated the model, analyzed the results, and drafted the initial manuscript. All authors contributed to the final manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Restrictions apply to the availability of these data, which were used under license for this study.

Acknowledgments: We thank the Ministry of Environment and Natural resources of El Salvador (MSRN) for providing data used in this study. We also thank Rotary Foundation (TRF) for the Fund provided within the Rotary fellowship.

Conflicts of Interest: The authors declare no conflict of interest.

Code Availability: The development of the Hapi source code is available in the public GitHub repository (<https://github.com/MAfarrag/Hapi>). Package releases are distributed using the Python package index (<https://pypi.org/project/HAPI-Nile>), and using conda package manager through the conda-forge channel (<https://anaconda.org/conda-forge/hapi>). For the release associated with this paper is 1.0.4 and assigned a DOI (Farrag & Corzo, 2021 [26]) through Zenodo. The documentation of the package is available on Read the Docs (<https://hapi-hm.readthedocs.io/>). Hapi is implemented using Python 3.6 and depends on Gdal and NumPy. Other dependencies are specified in the package setup. Hapi is available under the license GPL-3.0.

References

1. Reggiani, P.; Rientjes, T.H.M. Flux parameterization in the representative elementary watershed approach: Application to a natural basin. *Water Resour. Res.* **2005**, *41*, 1–18. [CrossRef]
2. Beven, K.; Calver, A.; Morris, E.M. The Institute of Hydrology distributed model Technical Report 98. *Inst. Hydrol. Wallingford* **1987**, 595–626.
3. Abbott, M.B.; Bathurst, J.C.; Cunge, J.A.; O’Connell, P.E.; Rasmussen, J. An introduction to the European Hydrological System—Systeme Hydrologique Europeen, “SHE”, 1: History and philosophy of a physically-based, distributed modelling system. *J. Hydrol.* **1986**, *87*, 45–59. [CrossRef]
4. Bao, H.; Wang, L.; Zhang, K.; Li, Z. Application of a developed distributed hydrological model based on the mixed runoff generation model and 2D kinematic wave flow routing model for better flood forecasting. *Atmos. Sci. Lett.* **2017**, *18*, 284–293. [CrossRef]
5. Abbott, M.B.; Refsgaard, J.C. *The Role of Distributed Hydrological Modelling in Water Resources Management*; Springer: Dordrecht, The Netherlands, 1996; Volume 22, ISBN 978-94-010-6599-3.

6. 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. *J. Hydrol.* **2004**, *298*, 242–266. [[CrossRef](#)]
7. Leavesley, G.H.; Lichty, R.W.; Troutman, B.M.; Saindon, L.G. Precipitation-runoff modeling system; user's manual. *Water Resour. Investig. Rep.* **1983**, *83*, 207. [[CrossRef](#)]
8. Iizuka, T.; Arita, K.; Yamamoto, I.; Yamamichi, S.; Yamaguchi, H.; Matsuki, T.; Sone, S.; Yabuta, H.; Miyasaka, Y.; Kato, Y. Low Temperature Recovery of Ru/(Ba, Sr)TiO₃/Ru Capacitors Degraded by Forming Gas Annealing. *Jpn. J. Appl. Phys.* **2000**, *39*, 2063–2067. [[CrossRef](#)]

9. Beven, K.J.; Kirkby, M.J. A physically based, variable contributing area model of basin hydrology. *Hydrol. Sci. Bull.* **1979**, *24*, 43–69. [[CrossRef](#)]
10. Samaniego, L.; Kumar, R.; Attinger, S. Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resour. Res.* **2010**, *46*, 1–25. [[CrossRef](#)]
11. Liang, X.; Lettenmaier, D.P.; Wood, E.F.; Burges, S.J. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *J. Geophys. Res. Atmos.* **1994**, *99*, 14415–14428. [[CrossRef](#)]
12. Harbaugh, A.W. *MODFLOW-2005: The U.S. Geological Survey Modular Ground-Water Model—the Ground-Water Flow Process*; USGS: Reston, VA, USA, 2005.
13. Lindström, G.; Johansson, B.; Persson, M.; Gardelin, M.; Bergström, S. Development and test of the distributed HBV-96 hydrological model. *J. Hydrol.* **1997**, *201*, 272–288. [[CrossRef](#)]
14. Reggiani, P.; Hassanizadeh, S.M.; Sivapalan, M.; Gray, W.G. A unifying framework for watershed thermodynamics: Constitutive relationships. *Adv. Water Resour.* **1999**, *23*, 15–39. [[CrossRef](#)]
15. Fan, Y.; Bras, R.L. On the concept of a representative elementary area in catchment runoff. *Hydrol. Process.* **1995**, *9*, 821–832. [[CrossRef](#)]
16. Reggiani, P.; Sivapalan, M.; Majid Hassanizadeh, S. A unifying framework for watershed thermodynamics: Balance equations for mass, momentum, energy and entropy, and the second law of thermodynamics. *Adv. Water Resour.* **1998**, *22*, 367–398. [[CrossRef](#)]
17. Dehotin, J.; Braud, I. Which spatial discretization for distributed hydrological models? Proposition of a methodology and illustration for medium to large-scale catchments. *Hydrol. Earth Syst. Sci.* **2008**, *12*, 769–796. [[CrossRef](#)]
18. Refsgaard, J.C.; Storm, B.; Refsgaard, A. Recent developments of the Systeme Hydrologique Europeen (SHE) towards the MIKE SHE. *IAHS Publ.* **1995**, *231*, 427.
19. Wood, E.F.; Sivapalan, M.; Beven, K.; Band, L. Effects of spatial variability and scale with implications to hydrologic modeling. *J. Hydrol.* **1988**, *102*, 29–47. [[CrossRef](#)]
20. Tian, F.; Hu, H.; Lei, Z.; Sivapalan, M. Extension of the Representative Elementary Watershed approach by incorporating energy balance equations. *Hydrol. Earth Syst. Sci. Discuss.* **2006**, *3*, 427–498. [[CrossRef](#)]
21. Bergström, S. Development and Application of a Conceptual Runoff Model for Scandinavian Catchments. *Smhi* **1976**, *7*, 134. [[CrossRef](#)]
22. Fenicia, F.; Kavetski, D.; Savenije, H.H.G.; Pfister, L. From spatially variable streamflow to distributed hydrological models: Analysis of key modeling decisions. *Water Resour. Res.* **2016**, *52*, 954–989. [[CrossRef](#)]
23. Savenije, H.H.G. HESS opinions “topography driven conceptual modelling (FLEX-Topo)”. *Hydrol. Earth Syst. Sci.* **2010**, *14*, 2681–2692. [[CrossRef](#)]
24. Mayr, E.; Hagg, W.; Mayer, C.; Braun, L. Calibrating a spatially distributed conceptual hydrological model using runoff, Annual mass balance and winter mass balance. *J. Hydrol.* **2013**, *478*, 40–49. [[CrossRef](#)]
25. Braun, L.N.; Aellen, M. Modelling discharge of glacierized basins assisted by direct measurements of glacier mass balance. *IAHS Publ.* **1990**, 99–106.
26. Farrag, M.; Corzo, G. MAfarrag/HAPI: HAPI. *Zenodo* **2021**. [[CrossRef](#)]
27. Schumann, A.H. Development of conceptual semi-distributed hydrological models and estimation of their parameters with the aid of GIS. *Hydrol. Sci. J.* **1993**, *38*, 519–528. [[CrossRef](#)]
28. Bergström, S. The HBV model - its structure and applications. *Smhi Rh* **1992**, *4*, 35.
29. Cunge, J.A. On the subject of a flood propagation computation method (Muskingum method). *J. Hydraul. Res.* **1969**, *7*, 205–230. [[CrossRef](#)]
30. Johnson, D.; Miller, A. A spatially distributed hydrologic model utilizing raster data structures. *Comput. Geosci.* **1997**, *23*, 267–272. [[CrossRef](#)]
31. Chaudhry, M.H. *Open Channel Flow*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2007; ISBN 9780387301747.
32. Mishra, S.K.; Singh, V.P. Hysteresis-based flood-wave analysis using the concept of strain. *Hydrol. Process.* **2001**, *15*, 1635–1651. [[CrossRef](#)]
33. Tewolde, M.H.; Smithers, J.C. Flood routing in ungauged catchments using Muskingum methods. *Water SA* **2006**, *32*, 379–388. [[CrossRef](#)]
34. Rusli, S.R.; Yudianto, D.; Liu, J. Effects of temporal variability on HBV model calibration. *Water Sci. Eng.* **2015**, *8*, 291–300. [[CrossRef](#)]
35. Lee, K.S.; Geem, Z.W. A new meta-heuristic algorithm for continuous engineering optimization: Harmony search theory and practice. *Comput. Methods Appl. Mech. Eng.* **2005**, *194*, 3902–3933. [[CrossRef](#)]
36. Nash, E.; Sutcliffe, V. PART I-A DISCUSSION OF PRINCIPLES * The problem of determining river flows from rainfall, evaporation, and other factors, occupies a central place in the technology of applied hydrology. *J. Hydrol.* **1970**, *10*, 282–290. [[CrossRef](#)]
37. Gupta, H.V.; Kling, H.; Yilmaz, K.K.; Martinez, G.F. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.* **2009**, *377*, 80–91. [[CrossRef](#)]
38. SNET Balance hídrico integrado y dinámico en El Salvador. Available online: <https://www.snet.gob.sv/Documentos/balanceHidrico.pdf> (accessed on 25 April 2021).

39. Das, T.; Bárdossy, A.; Zehe, E.; He, Y. Comparison of conceptual model performance using different representations of spatial variability. *J. Hydrol.* **2008**, *356*, 106–118. [[CrossRef](#)]
40. Sivapalan, M. Process complexity at hillslope scale, process simplicity at the watershed scale: Is there a connection? *Hydrol. Process.* **2003**, *17*, 1037–1041. [[CrossRef](#)]
41. Dooge, J.C.I. Bringing It All Together. *Hydrol. Earth Syst. Sci.* **2005**, *9*, 3–14. [[CrossRef](#)]
42. Savenije, H.H.G. Equifinality, a blessing in disguise? *Hydrol. Process.* **2001**, *15*, 2835–2838. [[CrossRef](#)]
43. Gharari, S.; Hrachowitz, M.; Fenicia, F.; Savenije, H.H.G. Hydrological landscape classification: Investigating the performance of HAND based landscape classifications in a central European meso-scale catchment. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 3275–3291. [[CrossRef](#)]