



Article Impact of the Mean Areal Rainfall Calculation on a Modular Rainfall-Runoff Model

Jose Valles ^(D), Gerald Corzo * ^(D) and Dimitri Solomatine

Department of Hydroinformatics and Socio-Technical Innovation, IHE Delft, Institute for Water Education, 2611 AX Delft, The Netherlands; jose.valles.leon@gmail.com (J.V.); d.solomatine@un-ihe.org (D.S.)

* Correspondence: g.corzo@un-ihe.org

Received: 16 October 2020; Accepted: 30 November 2020; Published: 2 December 2020



Abstract: Hydrological models are based on the relationship between rainfall and discharge, which means that a poor representation of rainfall produces a poor streamflow result. Typically, a poor representation of rainfall input is produced by a gauge network that is not able to capture the rainfall event. The main objective of this study is to evaluate the impact of the mean areal rainfall on a modular rainfall-runoff model. These types of models are based on the divide-and-conquer approach and two specialized hydrological models for high and low regimes were built and then combined to form a committee of model that takes the strengths of both specialized models. The results show that the committee of models produces a reasonable reproduction of the observed flow for high and low flow regimes. Furthermore, a sensitivity analysis reveals that Ilopango and Jerusalem rainfall gauges are the most beneficial for discharge calculation since they appear in most of the rainfall subset that produces low Root Mean Square Error (RMSE) values. Conversely, the Puente Viejo and Panchimalco rainfall gauges are the least beneficial for the rainfall-runoff model since these gauges appear in most of the rainfall subset that produces high RMSE value.

Keywords: modular modelling; high and low flow; sensitivity analysis; mean areal rainfall

1. Introduction

Floods problems in El-Salvador have been present since a long time ago. According to Gierloff-Emden [1], a hurricane in 1934 produced urban and river floods due to 500 mm of rainfall in three days. Additionally, Hurricane Fifi in 1974 produced flood damages in the coastal area of El-Salvador and Hurricane Mitch caused USD 388.1 million worth of economic losses, which accounts for 3.26% of the Gross Domestic Product (GDP) of the country [2]. Furthermore, Hydrological drought has also affected El-Salvador in recent times. In 2015, important rivers and reservoir levels in the country reached a critical state due to the El Niño phenomenon. This type of extreme event develops slowly and affects larger areas than flood conditions.

Different rainfall-runoff models have been implemented in El-Salvador—most of them were built for flood forecasting purposes. For instance, Häggström and Lindström [3] calibrate the conceptual Hydrologiska Byråns Vattenbalansavdelning (HBV) model to six rivers in Central America. In 2004, the Hydrological National Service of El-Salvador [4] reported a hydrological model for flood forecasting purposes in the Lempa Catchment. Additionally, Riverside Technology [5] implemented a semi-distributed Sacramento rainfall-runoff model in Lake Ilopango and Jiboa Catchment, which was implemented in a modular forecast system (Delft-FEWS) standalone version. However, these models exhibit limitations in reproducing low flow conditions accurately. Therefore, a different model structure could be built such that it can consider not only high flows but also low flow conditions.

In literature, many rainfall-runoff modelling approaches have been proposed to reproduce both high and low flows in a single deterministic model. For instance, Solomatine [6] and Corzo et al. [7,8]

proposed a modular modelling approach in which different local models are responsible for a particular region of the input space and then these local or specialized models are combined to form a committee of models which improves modelling results. Furthermore, Corzo et al. [9] incorporated various flow regimes in a modular Artificial Neural Network (ANN) model architecture. Fenicia et al. [10] and Kayastha et al. [11] improved hydrological model prediction using a fuzzy logic combination scheme of specialized models' outputs obtained from conceptual models. Overall, the modular modelling technique has been proved to improve model performance since it overcomes the structural limitations that a single rainfall-runoff model may display. Nevertheless, these models could still be affected by precipitation inaccuracy. Therefore, it is also important to evaluate the representation of the rainfall and quantify the variability of the model performance produced by poor rainfall estimates.

The aim of this study is to evaluate the impact of the mean areal rainfall on a modular model with a soft combination of local models in a catchment in El-Salvador that can reproduce high and low flow regimes. The modular modelling approach has been evaluated in different case studies [10,11] leading to a higher model performance compared to a single conceptual model. However, the modular rainfall-runoff model has mainly been applied to aggregated basins and not semi-distributed systems that include lake dynamics, which implies an important challenge in the forecasting process. Furthermore, the sensitivity analysis with respect to rainfall can help to find out the most influential rainfall gauges as well as the least beneficial, taking into consideration their usefulness for runoff calculation.

2. Case of Study

2.1. Description

The Jiboa River is one of the five important rivers in El-Salvador and is located at the central part of the country with a total drainage area of 604.87 km², a lake surface area of 70.64 km² and a channel length of 61.48 km. The mean elevation of the catchment is 238 meters with a maximum elevation of 2106 meters above mean sea level. In the lower part of the basin, large coastal areas are commonly catalogued as flood prone areas in the basin.

The climate in the study area is divided into two seasons: the wet period from May to October, which is characterized by convective storms and some extreme rainfall events, and the second season is the dry period from November to April which is characterized by trade winds and temperature ranges from 20 to 40 $^{\circ}$ C. More details can be seen in Table 1.

Table 1. Summary of climate parameters in the Jiboa Catchment. Reproduced from Master Plan for development and use of water resources in El-Salvador. Ministry of Agriculture and Livestock of El Salvador (1981) [12].

Climate Period	Wet season	Dry season
Period	May-October	November-April
Rainfall	Convective storm	Some convective storm
Sunlight	7.6 hours per day	9.4 hours per day
Wind	Wind speed of 5.6 km/h	Trade winds of 6.5 km/h
Special phenomenon	Extreme rainfall and drought	North wind component
Temperature ranges	20 to 35 °C	19 to 36 °C
Relative humidity	Monthly average of 80%	Monthly average of 66%

For this research, the part of the catchment that drains to the Puente Viejo station with an area of 431.93 km² is considered in the modelling setup. Figure 1 presents the catchment, the rainfall stations considered, and the semi-distributed model structure proposed in the study.



Figure 1. Jiboa and Lake Ilopango catchment area that drains to Puente Viejo water level gauge and the surrounding rainfall gauges.

2.2. Data Available

The Ministry of Environment and Natural Resources of El-Salvador (MARN for its acronym in Spanish) has provided rainfall, water level and discharge data for the 1970 to 2015 period. These data are necessary for the calibration and verification of the rainfall-runoff model since they provide a mean to mimic the observed discharge at Puente Viejo Station.

The rainfall dataset is divided into conventional and telemetry stations. The former records rainfall data in a daily time scale for the 1970 to 2015 period. The latter records rainfall estimates in a minute time step for the 2002 to 2015 period (see Figure 2). This rainfall gauge network includes Puente Viejo Station since it has a rainfall sensor in its setup.



Figure 2. Period of record for telemetry rainfall stations surrounding the Jiboa Catchment.

Even though most of the data correspond to daily measurements, telemetry stations (minute time scale) have been selected to calibrate and validate the modular rainfall-runoff model for three reasons:

- The time of concentration of the catchment is of the order of hours.
- Automatic rainfall gauge represents the most recent data in the study area.

• For operational forecast, telemetry rainfall gauges are more suitable than daily gauges in this case study.

Water level data at Puente Viejo station were provided by MARN for the 2012 to 2016 hydrological year period. These data were converted into discharge units using a rating curve equation shown in Equation (1), where a, b and H_o are the rating curve coefficient and h is the water level. The rating curve coefficient for different hydrological years and valid periods is presented in Table 2.

$$Q = b \cdot (h - H_0)^a \tag{1}$$

 Table 2. Rating curve coefficients for different hydrological years at Puente Viejo Station.

Gauge/Station	Hydrological Year	Valid Period	Ho	а	b
Puente Viejo	2012–2013 May 2012 to April 2013		0.15	1.89	13.20
	2013–2014	May 2013 to November 2013	0.16	1.90	13.27
		November 2013 to April 2014	0.02	3.00	15.45
	2014–2015	May 2014 to April 2015	0.15	1.25	13.46
	2015–2016	May 2015 to April 2016	0.18	1.73	20.04

A rainfall-runoff model requires potential evapotranspiration (Ep) as an input. For this study, the Ep was estimated using the potential evapotranspiration–elevation relationships shown in Equation (2), where x is the elevation above mean sea level (m), y is the potential evapotranspiration (Ep) in millimetres and A, B and C are equation parameters. These parameters are shown in Table 3 for each month and were determined using the Hargreaves method and linear regression equations [13].

$$y = Ax^2 + Bx + C \tag{2}$$

Table 3. Potential evapotranspiration–elevation equation parameters (A, B and C). Reproduced from Integrated and Dynamic Water Balance in El-Salvador. National Service for Territorial Studies of El-Salvador (2005) [13].

Month	Α	В	С
January	-7.0×10^{-6}	-0.0170	152.95
February	$-8.0 imes10^{-6}$	-0.0164	156.04
March	-1.0×10^{-5}	-0.0170	190.42
April	-9.0×10^{-6}	-0.0202	194.52
May	-7.0×10^{-6}	-0.0237	186.89
June	-7.0×10^{-6}	-0.0181	165.79
July	-9.0×10^{-6}	-0.0127	175.28
August	$-8.0 imes10^{-6}$	-0.0122	170.12
September	-8.0×10^{-6}	-0.0126	150.42
October	-1.0×10^{-5}	-0.0094	149.56
November	-9.0×10^{-6}	-0.0119	142.18
December	-7.0×10^{-6}	-0.0168	144.91

3. Materials and Methods

3.1. Modular Modelling with Soft Combination

Solomatine [6] described modular modelling as a set of specialized or local models each of which is responsible for a region of the input space. In recent years, these modular models have been subject to

many experiments using data-driven and conceptual models such as local models. These experiments have demonstrated that the accuracy of the output is improved since it overcomes the structural limitations that a single deterministic model may exhibit.

The modular principle divides the input vector into different specialized or local models (hard partitioning). Thereafter, these local models are combined with a soft combination scheme at the boundary between them to produce a composite solution [6].

In this research, the hard partitioning with a soft combination of models using a fuzzy logic scheme will be used since problems of compatibility among the models could occur due to the hard partitioning. The smooth combination of local or specialized models can be reached with a fuzzy logic scheme as proposed by Solomatine [6] and implemented in conceptual models recently. For instance, Fenicia et al. [10] improved hydrological model prediction involving a combination of specialized models outputs obtained from conceptual models in the Alzette River basin in Luxembourg. These two local models were combined using a fuzzy logical trapezoidal membership function with discharge threshold parameters of δ and γ .

In addition to the previous experiment, Kayastha et al. [11] extended earlier work from Fenicia et al. [10] using a fuzzy logic scheme to combine specialized models. The authors used four different types of weighting schemes to evaluate the performance of specialized models and two membership functions to combine the two flow conditions.

In this study, the "Hydrologiska Byråns Vattenbalansavdelning" (HBV) conceptual model has been used to transform the rainfall into runoff. For this purpose, the HBV code in MATLAB programming language provided by IHE-Delft, has been modified to explicitly include lake dynamics. The updated HBV Lake model requires lake physical characteristics such as storage–discharge curve to compute the runoff simulation.

The spatial disaggregation of the inputs and model parameters is usually divided into three different approaches: lumped, semi-distributed and distributed modelling. Lumped modelling analyses the catchment as a single unit, semi-distributed modelling divides the catchment into a smaller subcatchment, and distributed modelling performs the model computation in a gridded manner. Figure 1 shows the semi-distributed model structure selected for this study whose HBV outflows were summed and then routed using the triangular weighting function (MAXBAS). The reason for using a semi-distributed model instead of a lumped model is because the two sub-basins exhibits different physical processes. For instance, the Jiboa sub-basin response is faster than the Lake Ilopango sub-basin which is controlled by the water volume in the water body.

After this process, the input data were split into calibration and verification sets. The aim of this process is to calibrate the model parameters with a lot of data points that includes different flow regimes and to evaluate the model performance in an independent dataset. The statistical summary of the calibration and verification dataset is presented in Table 4. The calibration dataset cover both dry and wet periods and the verification period only covers the wet period. It was difficult to split the data into statistically similar sets since the procedure was constrained by data available and the wish to keep data in a contiguous block.

Table 4. Statistical summary of the runoff data at Puente Viejo Station for the Jiboa River basin.

Statistical Properties	Complete Dataset	Calibration Dataset	Verification Dataset	
Period (month/year)	June/2012–December/2013 July/2015–December/2015	June/2012-December/2013	July/2015–December/2015	
Number of data points	17,779	13,433	4346	
Average (m ³ /s)	3.68	3.45	4.38	
Minimum (m ³ /s)	0.79	1.24	0.79	
Maximum (m ³ /s)	35.74	31.22	35.74	
Standard deviation (m ³ /s)	3.18	2.67	4.34	

The aim of using specialized or local models instead of only one single model is to reproduce high and low flow regimes accurately. These specialized or local models were found using a multi-objective optimization which minimized two weighted objective functions—RMSE_{LF} and RMSE_{HF}—shown in Equations (3) and (4), where n is total number of time steps, $Q_{sim,i}$ is the simulated flow for time step i and $Q_{obs,i}$ is the simulated flow for time step i.

$$RMSE_{LF} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Q_{s,i} - Q_{o,i})^2 \cdot W_{LF,i}}$$
(3)

$$RMSE_{HF} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Q_{s,i} - Q_{o,i})^2 \cdot W_{HF,i}}$$
(4)

The two weighting functions W_{LF} and W_{HF} , for the objective function formulas, were calculated using the weighting function (Figure 3) and Equations (3) and (4), where the $Q_{o,max}$ is the maximum observed flow.

$$W_{\rm HF} = \left(\frac{Q_{\rm o,i}}{Q_{\rm o,max}}\right)^2 \tag{5}$$

$$W_{\rm LF} = \left(\frac{Q_{\rm o,max} - Q_{\rm o,i}}{Q_{\rm o,max}}\right)^2 \tag{6}$$



Figure 3. Weight functions for high and low flows. Reproduced with permission from Kayastha et al., Hydrology and Earth System Sciences; published by Copernicus Publications on behalf of the European Geoscience Union, 2012 [11].

The NSGA-II Genetic multi-objective optimization algorithm was used in the study to find the specialized models HBV_{LF} and HBV_{HF} . The outcome of this framework analysis is the Pareto-optimal set of solutions. All the trade-off solutions are equally important, which means that it is difficult to prefer one solution over the other without any further information about the problem. The two specialized models— HBV_{HF} and HBV_{LF} —represent the best models that minimize the two objective functions RMSE_{HF} and RMSE_{LF}, respectively. However, different optimal combinations could be used since every

solution has its strength and limitation in describing the observed discharge [10]. Moreover, two single objective optimizations, separately, could also have also been applied as alternative procedures to obtain the two specialized HBV models. For the combining process of the two specialized models, the two-membership function shown in Figure 4 and indicated by Kayastha et al. [11] was used. The linear function or MFtype-A represents the degree of membership associated with the flow condition. For instance, the degree of membership is 0 when the simulated flow is below the discharge threshold γ , it increases linearly between the γ and δ parameters, and it is 1 when the simulated flow is above the discharge threshold δ . On the contrary, the degree of membership is 1 when the simulated flow is below the Υ , it decreases linearly between the transitional parameters, and it is 0 when the simulated flow is above the δ parameter [10]. Similarly, the quadratic function or MFtype-B follows the same logic using a quadratic equation instead of a linear equation in the region where the two specialized models work together.

$$M_{LF} = \begin{cases} 1, h < \gamma \\ \left(\frac{h - \gamma}{\delta - \gamma}\right)^{N}, \gamma \le h < \delta \\ 0, h \ge \delta \end{cases}$$
(7)

$$M_{\rm HF} = \begin{cases} 0, \ h < \gamma \\ 1 - \left(\frac{h - \gamma}{\delta - \gamma}\right)^{\rm N}, \ \gamma \le h < \delta \\ 1, \ h \ge \delta \end{cases}$$
(8)



Figure 4. Fuzzy membership functions for combining local models. (a) Trapezoidal membership function (MFtype-A), (b) a quadratic membership function (MFtype-B) for high and low flow models. Reproduced with permission from Kayastha et al., Hydrology and Earth System Sciences; published by Copernicus Publications on behalf of the European Geoscience Union, 2012 [11].

The membership functions M_{LF} and M_{HF} were computed using Equations (7) and (8) where γ and δ are the discharge threshold parameters, N is the power value referring to linear (N = 1) or quadratic (N = 2) functions and the variable h is the division between the observed flow at any time step and the maximum flow in the dataset. The outputs of the two specialized or local models were combined according to Equation (9) where the M is the membership function for low or high flows (M_{LF} and M_{HF}) calculated in Equations (7) and (8) and HBV is the specialized model output for high and flow regimes (HBV_{LF} and HBV_{HF}).

$$Q_{c,i} = \frac{(M_{LF} \cdot HBV_{LF,i} + M_{HF} \cdot HBV_{HF,i})}{M_{LF} + M_{HF}}$$
(9)

The threshold parameters (γ and δ) were manually calibrated using the Root Mean Square Error (RMSE) as objective function (See Equation (10)).

RMSE =
$$\sqrt{\frac{1}{n} \sum_{i=1}^{n} (Q_{s,i} - Q_{o,i})^2}$$
 (10)

For comparison purposes, a single HBV model was calibrated using the Root Mean Square Error (RMSE) as an objective function. This model was built with the same calibration and verification dataset shown in Table 4.

3.2. Sensitivity Analysis of the Mean Areal Rainfall Input

Rainfall data from gauges form an important input in rainfall-runoff models since they are the primary drivers of runoff. These input data are often considered independently of the modelling procedure. In hydrology, many methods such as the arithmetic average, Thiessen polygons, contour area weighted average (i.e., Isohyetal) and algorithm hypsometric method have been proposed to estimate the average areal rainfall that feeds the hydrologic model. However, these methods could produce a bias in runoff simulations since there is a sensitivity of the rainfall-runoff model output to the characteristics of the rain gauge network used as the model input.

The improvement in computational power and calibration process opened a window of opportunities to researchers to investigate the sensitivity of hydrological models to rainfall estimates errors. Yet, before these computational improvements, many authors such as Crawford and Linsley [14], Sutcliffe [15], Hamlin [16] and Morrissey et al. [17] had already discussed the problem from a theoretical perspective.

The impact of the variability of rainfall estimates on runoff simulations can lead to flow forecast inaccuracy [18]. Furthermore, the temporal and spatial nature of rainfall volume may or may not be captured by rainfall gages, and thus a poor runoff calculation is produced by the model. Therefore, it is important to analyze the sensitivity of the model with an imperfect estimation of rainfall input.

Anctil et al. [19], Andréassian et al. [20], Arnaud et al. [21], Dawdy and Bergmann [22] and Melching [23] indicated that the variability of rainfall estimates could be compensated by adjusting model parameters. This variability is attenuated by the low-pass filter behavior that the rainfall-runoff models can exhibit [24,25]. Hence, it should consider the characteristics of this filter in order to determine the quality and quantity of rainfall estimates required to achieve a degree of accuracy in flow prediction.

Andréassian et al. [20] examined the quality of rainfall time distribution and the total depth based on the analysis of how areal rainfall is transformed into runoff in three lumped conceptual models. Furthermore, the authors proposed the Goodness of Rainfall Estimate (GORE) and BALANCE index to analyze the link between the performance of the model and the representativeness of the mean areal rainfall estimate used to run the model.

The Goodness of Rainfall Estimation (GORE) index assesses the quality of rainfall time distribution based on the best estimate or reference areal rainfall (P) since the true precipitation is impossible to measure. The GORE index varies from $-\infty$ to 1. It takes a value of 1 when the estimated rainfall (P^E) equals the reference areal rainfall and decreases as the estimates become poorer. The index is a transposition of the Nash–Sutcliffe criterion in the rainfall domain that uses square roots of the rainfall estimates to reduce the weights of extreme events [19,20].

$$GORE = 1 - \frac{\sum_{i=1}^{n} \left(\sqrt{P_{i}^{E}} - \sqrt{P_{i}}\right)^{2}}{\sum_{i=1}^{n} \left(\sqrt{P_{i}} - \overline{\sqrt{P}}\right)^{2}}$$
(11)

The BALANCE index quantifies the overestimation (BALANCE > 1) or underestimation (BALANCE < 1) of the reference rainfall by a given rain gauge subset. This index compares the sum of the subset of rainfall gauge (P) to the sum of reference rainfall (P^E) over the available period [19,20]. Equation (10) defines the BALANCE index.

$$BALANCE = \frac{\sum_{i=1}^{n} P_i^E}{\sum_{i=1}^{n} P_i}$$
(12)

The approach consists of comparing a reference rainfall estimate with the mean areal rainfall obtained with a subset of gauges. These subsets were randomly obtained using a different number of rainfall gauge combinations. For the reference rainfall estimate, the input rainfall of the modular rainfall-runoff model was used since it includes all the available rain gauges and provides the "best" rainfall estimate.

For the generated mean areal subset, an arithmetic mean scheme was used to calculate the mean areal rainfall. The selected scheme is simple to implement and its results could be used to evaluate the ability of the selected rainfall subset to represent the reference rainfall time series which was calculated using a weighting scheme.

Figure 5 describes the methodology followed in this sensitivity analysis with respect to rainfall estimates, where the input rainfall gauges were randomly generated in order to obtain 411 rainfall subsets. These subsets represent a different number of possible rainfall gauge combinations and were randomly generated using k different rain gauges combinations (for k = 2, ..., 9). The number of combinations of rainfall gauges are shown in Table 5 where, for instance, 120 different combinations or subsets are possible using five rainfall gauges. The number of possible combinations was constrained by the wish to keep at least one rainfall gauge in one sub-basin. Otherwise, the modular model will not have a rainfall input in one sub-basin.



Figure 5. Approach used for the sensitivity analysis of the model input in the study area.

Number of Rain Gauges	Number of Possible Combinations			
2	18			
3	63			
4	111			
5	120			
6	83			
7	36			
8	9			
9	1			
Total	441			

Table 5. Number of possible combinations or subsets using different number of rainfall gauges.

After this process, each subset was used as rainfall input in the modular rainfall-runoff model without changing its calibrated model parameters. Finally, the Root Mean Square Error (RMSE), the Goodness of Rainfall Estimate (GORE) and the BALANCE indexes were calculated for each subset.

4. Results and Discussion

4.1. Modular Rainfall-Runoff Modelling

The aim of using specialized or local models instead of only one single model is to reproduce accurate high and low flow regimes. These specialized or local models (HBV_{LF} and HBV_{HF}) were found using a multi-objective optimization which minimized two weighted objective functions—RMSE_{LF} and RMSE_{HF}, shown in Equations (3) and (4). Furthermore, the two-weighting functions, W_{LF} and W_{HF} , for the objective function formulas, were calculated using the WStype-I function (Figure 3). The outcome of this framework analysis is the Pareto-optimal set of solutions which is shown in Figure 6. All the trade-off solutions are equally important, which means that it is difficult to prefer one solution over the other without any further information about the problem.



Figure 6. Pareto-optimal solutions in the objective space (blue) and specialized models (red).

The two specialized models (HBV_{HF} and HBV_{LF}) represent the best models that minimize the two objective functions $RMSE_{HF}$ and $RMSE_{LF}$, respectively. However, different optimal combinations could be used since every solution has its strength and limitation in describing the observed discharge [10].

The calibrated parameters of the two specialized models are shown in Table 6. For one specialized model, there are 17 calibrated model parameters in total where eight parameters correspond to the Lake Ilopango sub-basin, eight parameters correspond to the Jiboa sub-basin and one parameter is the transformation function MAXBAS.

For comparison purposes, a single conceptual HBV model was built and calibrated using the RMSE formula presented in Equation (10). The aim of the comparison was to evaluate how much the modular rainfall-runoff model improves the model efficiency using different discharge threshold parameters γ and δ .

Model Parameter	Units	Lower Bound	Upper Bound	Low Flow Model (HBV _{LF})		High Flow Model (HBV _{HF})	
	Cinto			Lake Ilopango	Jiboa	Lake Ilopango	Jiboa
FC	mm	50	500	320.26	462.86	319.77	463.16
LP	-	0.001	1	0.06	0.24	0.09	0.12
K0	mm/h	0.001	1	0.31	0.02	0.37	0.01
K1	mm/h	0.001	1	0.77	0.02	0.75	0.84
α	-	0.001	1	0.95	0.08	0.98	0.02
β	-	0.001	1.5	0.16	0.97	0.11	1.05
CFLUX	mm/h	0.001	4	2.81	3.95	2.76	3.93
PERC	mm/h	0.001	0.6	0.50	0.33	0.53	0.15
MAXBAS	h	2	15	6.77	7	5.65	5

Table 6. Optimal parameters for the two sub-basins and local models HBV_{HF} and HBV_{LF} .

The performance results for calibration and validation can be seen in Table 7 where all the modular rainfall-runoff models show better model performance than a single HBV model. Additionally, it shows a special case of the combining scheme using the same value of 0.5 for the threshold parameters γ and δ . This condition means no soft combination or smooth transition between the boundaries. This combined model also provides better performance than a single conceptual model, although is not better than the others modular models.

Table 7. Performance of different modular models compared to the single optimal HBV conceptual rainfall-runoff model.

Dataset	Model	MFtype	Υ	δ	RMSE _{LF}	RMSE _{HF}	RMSE
	Single HBV	n.a.	n.a.	n.a.	1.309	0.608	1.709
	Ū		0.1	0.5	1.028	0.487	1.349
			0.2	0.4	1.046	0.491	1.373
		А	0.2	0.5	1.053	0.492	1.382
			0.2	0.6	1.062	0.499	1.399
Calibration	Modular		0.5	0.5	1.112	0.533	1.349
	Model		0.1	0.5	1.046	0.491	1.372
			0.2	0.4	1.044	0.493	1.373
		В	0.2	0.5	1.042	0.490	1.368
			0.2	0.6	1.046	0.490	1.372
			0.5	0.5	1.112	0.533	1.349
Verification	Single HBV	n.a.	n.a.	n.a.	1.914	1.229	2.867
	-		0.1	0.5	1.488	1.131	2.351
			0.2	0.4	1.516	1.135	2.380
		А	0.2	0.5	1.496	1.130	2.354
			0.2	0.6	1.492	1.127	2.346
	Modular		0.5	0.5	1.516	1.138	2.383
	Model		0.1	0.5	1.524	1.139	2.396
			0.2	0.4	1.544	1.141	2.415
		В	0.2	0.5	1.517	1.135	2.381
			0.2	0.6	1.503	1.131	2.362
			0.5	0.5	1.516	1.138	2.383

Furthermore, the best committee model or "MC-Model" is provided by MFtype-A using a discharge threshold of 0.1 and 0.5 for low (γ) and high (δ) flow, respectively. From this solution, it appears that the high flow model works in a wider space than the low flow model; this could be related to the fact that the classical RMSE formula is oriented to stress large deviation in high flows and, therefore, it requires a wider space for the HBV_{HF} to adjust the model.

Kayastha et al. [11] concluded that a very narrow region where the two specialized models work together may lead to force the modular model to produce large changes in output with minor changes in average flow. For this procedure, the best modular model ensures a smooth transition between the boundaries or transitional parameters where the two specialized models work together. Nevertheless, these parameters were manually calibrated, and they could display other values with a narrower space using an optimization algorithm. Further developments in improving these parameters could be carried out in future works.

Figure 7 shows the modular rainfall-runoff models in the objective functions space. It can be seen that all the combined models provide a better model performance than the specialized or local models and a single HBV model. This result was expected since Fenicia et al. [10] and Kayastha et al. [11] proved that modular modelling is able to overcome the limitation of an HBV conceptual model in reproducing different flow regimes in the catchment.



Figure 7. Modular rainfall-runoff models compared to a single HBV model and local models.

The scatter plot between the observed and simulated discharge in the calibration period is shown in Figure 8. From this figure, it can be seen that some points exhibit a bias between the observed and simulated discharge (i.e., reference dash line). The model results exhibit underestimation of the peaks which may be produced by poor mean areal rainfall input and/or inaccuracy of observed flow since high values are normally not measured.

Figure 9 displays the observed and simulated flows for the calibration period. From this graph, it can be observed that there is a good agreement between the MC-Model and the observed flow in terms of the overall behavior of volume, time of peaks and water balance (e.g., accumulated volume).



Figure 8. Scatter plot and coefficient of determination (\mathbb{R}^2) between the observed and simulated discharge in the calibration dataset.



Figure 9. Comparison between the observed and simulated (**a**) and accumulated volume (**b**) with the Coefficient of Determination (\mathbb{R}^2), Mean Absolute Error (MAE) and Relative Volume Error (RVE) performance metrics in the calibration dataset.

4.2. Sensitivity Analysis with Respect to Rainfall Input

The rainfall-based indexes (i.e., GORE and BALANCE) and the Root Mean Square Error (RMSE) model performance were compared. Figure 10a presents the results of the model performance of the modular rainfall-runoff model and the ability of a rainfall subset to represent the reference rainfall (GORE index). The scatterplot reveals an interesting point of discussion:

• The model error decrease with a better description of the rainfall estimates. This result was expected since there is a link between model errors and areal rainfall estimates.

- It is possible to obtain a low RMSE value with a low network quality (i.e., GORE index). To illustrate this, the scatter plot (see Figure 10a) displays a GORE value of 0.82 and RMSE of 1.64 m³/s. This result was obtained using Ilopango and Jerusalem rainfall gauges and an arithmetic mean for areal rainfall. In fact, this outcome is surprising since it is well-known that more rainfall gauges provide better rainfall description. However, it appears that these rainfall gauges play an important role in the runoff calculation in the hydrological model.
- The range of the model error is 15.48 m³/s which means the model output exhibits a high variability in the model performance produced by poor rainfall estimates. This situation could be related to rainfall gauges (e.g., Panchimalco and Puente Viejo gauges) that are not useful for discharge calculations in the study area. This outcome does not imply that these rainfall gauges should be removed from the rainfall network at Jiboa Catchment since they could be used for other purposes such as rainfall monitoring for early warning systems.
- Some arithmetic rainfall subsets can represent the true reference rainfall which was obtained with a weighted mean areal rainfall (Thiessen polygon) since the Goodness of Rainfall Estimate (GORE) efficiency is around 0.8. This outcome could mean that it is possible to represent the mean areal rainfall using the arithmetic formula and some rainfall gauges that contribute to streamflow calculation in the study area.



Figure 10. Impact of the quality of rainfall estimate on the modular model in calibration period. (a) Goodness of Rainfall Estimate (GORE) index value and (b) BALANCE index values.

Additionally, Figure 10b shows other important patterns:

- The lowest model errors are produced by underestimation (BALANCE less than one) of the total amount of reference rainfall. Conversely, the highest model error is produced by an overestimation (BALANCE greater than one) of the reference rainfall.
- There are some rainfall subsets which are similar to the total amount of reference rainfall (BALANCE equal to one) but the model error is higher. This outcome means that it is not sufficient to match the total reference rainfall to achieve a low model error since the temporal characteristics of the rainfall are important as well.

Further analysis was also made regarding the number of rainfall gauges and their impact on runoff calculation. The box-and-whisker diagram in Figure 11 displays the variability of model error using a different combination of rainfall gauges. This diagram shows interesting results:

• The interquartile range (i.e., the difference between the 75th and 25th percentiles) and the median model error diminish when more rainfall gauges are used in the rainfall calculation.

- Using two rainfall gauges provides the lowest RMSE value. This result indicates that there are two rainfall gauges which are more beneficial for discharge calculation than the others. Nevertheless, this number of rainfall gauges also shows a wide model error variability.
- The minimum set value increases with more rainfall gauges. This situation could be related to the fact that some rainfall gauge combinations affect the model performances.



Figure 11. Box plot of model error (Root Mean Square Error (RMSE)) and the number of rainfall gauges.

5. Conclusions

This research presents the implementation of a modular rainfall-runoff model to reproduce high and low regimes in a small catchment in El-Salvador which is affected by flood events as well as streamflow drought. This modular model has been used in many case studies leading to a better representation of the catchment behaviour, but few experiments have been carried out in catchments that includes a lake dynamic and semi-distributed modelling system. Additionally, this study includes an analysis of the model output with respect to mean areal rainfall estimates to evaluate the representation of the rainfall and its effect on the model performance. The conclusions drawn from the research are:

- The best modular rainfall-runoff model, the MC-Model, was obtained using the trapezoidal membership function (MFtype-A) and a discharge threshold parameter of 0.1 for low flow and 0.5 for high flow for this case study. This solution provided a wider space for the high flow model than the low flow model since the classical RMSE formula is oriented to stress large deviation in high flows and, therefore, it requires a wider space for HBV_{HF} to adjust the model.
- The MC-Model exhibited a reasonable agreement with respect to the observed discharge for most of the data. In fact, it showed a better model reproduction of the high and low flow conditions than a single conceptual HBV model in the calibration and verification period. However, the model results exhibit underestimation of the peaks which may be produced by poor mean areal rainfall input and/or inaccuracy of observed flow since high values are normally not measured.
- The sensitivity analysis with respect to rainfall input revealed that Ilopango and Jerusalem rainfall gauges are the most beneficial for runoff calculation since they appear in most of the rainfall subset that produces low RMSE values. Conversely, Panchimalco and Puente Viejo rainfall gauges are the least beneficial for discharge calculation since these gauges appear in most of the high RMSE values. This outcome does not imply that the latest rainfall gauges should be removed from the telemetry rainfall network since they are used for other purposes such as monitoring rainfall for early warning systems and not only for hydrological modelling.

• There are some rainfall subsets that produce a BALANCE index close to one, but the model error is higher than the reference model. This situation could mean that the temporal distribution of the rainfall over the whole period is also necessary to achieve a low model performance.

Author Contributions: Conceptualization, J.V., G.C. and D.S.; data analysis, J.V.; formal analysis, J.V.; supervision, G.C., and D.S.; writing—original draft preparation, J.V.; writing—review and editing, G.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors wish to express their gratitude to the Ministry of Environment and Natural Resources of El-Salvador (MARN) for providing the rainfall, water level and discharge data of Jiboa Catchment.

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

References

- 1. Gierloff-Emden, H.G. *La Costa de El-Salvador: Monografía Morfológica-Oceanográfica;* Ministerio de Educación, Dirección de Publicaciones: San Salvador, El-Salvador, 1976.
- 2. Comisión Económica para América Latina y el Caribe Evaluación de los daños ocasionados por el huracán Mitch. In *Sus Implicaciones Para el Desarrollo Económico, Social y el Medio Ambiente, Documento;* CEPAL1999-02-15; CEPAL: México, D.F., México, 1999; Volume 11026.
- 3. Häggström, M.; Lindström, G. *Application of the HBV Model for Flood Forecasting in Six Central American Rivers;* SMHI: Norrkoping, Sweden, 1990.
- 4. Servicio Hidrologico Nacional Sistema de alerta temprana por inundaciones. In *Experiencia en El-Salvador;* Servicio Nacional de Estudios Territoriales: San Salvador, El-Salvador, 2004.
- 5. *Riverside Technology Development and Implementation of a Regional Flash Flood Guidance and Early Warning System;* Riverside Technology: Fort Collins, CO, USA, 2012.
- Solomatine, D.P. Optimal modularization of learning models in forecasting environmental variables. In Proceedings of the 3rd International Congress on Environmental Modelling and Software, Burlington, VT, USA, 9–13 July 2006.
- Corzo, G.A.; Siek, M.; Solomatine, D.; Price, R. Modular data-driven hydrologic models with incorporated knowledge: Neural networks and model trees. In Proceedings of the International Congress IAHR, ASCE, Venice, Italy, 1–6 July 2007.
- 8. Corzo, G.; Solomatine, D. Knowledge-based modularization and global optimization of artificial neural network models in hydrological forecasting. *Neural Netw.* **2007**, *20*, 528–536. [CrossRef]
- 9. Corzo, G.; Solomatine, D. Baseflow separation techniques for modular artificial neural network modelling in flow forecasting. *Hydrol. Sci. J.* 2007, *52*, 491–507. [CrossRef]
- 10. Fenicia, F.; Solomatine, D.P.; Savenije, H.H.G.; Matgen, P. Soft combination of local models in a multi-objective framework. *Hydrol. Earth Syst. Sci.* 2007, *11*, 1797–1809. [CrossRef]
- 11. Kayastha, N.; Ye, J.; Fenicia, F.; Kuzmin, V.; Solomatine, D.P. Fuzzy committees of specialized rainfall-runoff models: Further enhancements and tests. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 4441–4451. [CrossRef]
- 12. MAG. PNUD Plan Maestro de Desarrollo y Aprovechamiento de los Recursos Hídricos; Ministerio de Agricultura y Ganadería: San Salvador, El-Salvador, 1981.
- 13. SNET. *Balance Hídrico Integrado y Dinámico en El-Salvador;* Servicio Nacional de Estudios Territoriales: San Salvador, El-Salvador, 2005; p. 118.
- 14. Crawford, N.H.; Linsley, R.K. Digital Simulation in Hydrology'Stanford Watershed Model 4. In *Stanford University Technical Report*; Stanford University: Stanford, CA, USA, 1966.
- 15. Sutcliffe, J.V. The Assessment of Random Errors in Areal Rainfall Estimation. *Int. Assoc. Sci. Hydrol. Bull.* **1966**, *11*, 35–42. [CrossRef]
- 16. Hamlin, M.J. The significance of rainfall in the study of hydrological processes at basin scale. *J. Hydrol.* **1983**, 65, 73–94. [CrossRef]
- 17. Morrissey, M.L.; Maliekal, J.A.; Greene, J.S.; Wang, J. The uncertainty of simple spatial averages using rain gauge networks. *Water Resour. Res.* **1995**, *31*, 2011–2017. [CrossRef]
- 18. Beven, J.K. Rainfall-Runoff Modelling the Primer; John Willey & Sons Ltd.: New York, NY, USA, 2000.

- 19. Anctil, F.; Lauzon, N.; Andréassian, V.; Oudin, L.; Perrin, C. Improvement of rainfall-runoff forecasts through mean areal rainfall optimization. *J. Hydrol.* **2006**, *328*, 717–725. [CrossRef]
- 20. Andréassian, V.; Perrin, C.; Michel, C.; Usart-Sanchez, I.; Lavabre, J. Impact of imperfect rainfall knowledge on the eficiency and the parameters of watershed models. *J. Hydrol.* **2001**, *250*, 206–223. [CrossRef]
- 21. Arnaud, P.; Lavabre, J.; Fouchier, C.; Diss, S.; Javelle, P. Sensitivity of hydrological models to uncertainty in rainfall input. *Hydrol. Sci. J.* 2011, *56*, 397–410. [CrossRef]
- 22. Dawdy, D.R.; Bergmann, J.M. Effect of rainfall variability on streamflow simulation. *Water Resour. Res.* **1969**, *5*, 958–966. [CrossRef]
- 23. Melching, C.S. Reliability estimation. In *Computer Models of Watershed Hydrology*; Water Resources Publications, LLC: Highlands Ranch, CO, USA, 1995.
- 24. Bras, R.L. Sampling of interrelated random fields: The rainfall-runoff case. *Water Resour. Res.* **1979**, *15*, 1767–1780. [CrossRef]
- 25. Bras, R.L.; Rodríguez-Iturbe, I. Network design for the estimation of areal mean of rainfall events. *Water Resour. Res.* **1976**, *12*, 1185–1195. [CrossRef]

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



© 2020 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 (http://creativecommons.org/licenses/by/4.0/).