

Article

# Adapting hydrological regionalization techniques to reconstruct rainfall fields in Haiti

Clertine Guerrier<sup>1</sup>, Alfonso Gutierrez-Lopez<sup>2,\*</sup>

<sup>1</sup> Water Research Center, International Flood Initiative Representative in Haiti (IFI-Haiti), Intergovernmental Hydrological Programme (IHP), Autonomous University of Queretaro, Queretaro 76010, Mexico

<sup>2</sup> Water Research Center, International Flood Initiative, Latin-American and the Caribbean Region (IFI-LAC), Intergovernmental Hydrological Programme (IHP), Autonomous University of Queretaro, Queretaro 76010, Mexico

\* **Corresponding author:** Alfonso Gutierrez-Lopez, [alfonso.gutierrez@uaq.mx](mailto:alfonso.gutierrez@uaq.mx)

## CITATION

Guerrier C, Gutierrez-Lopez A.  
Adapting hydrological regionalization techniques to reconstruct rainfall fields in Haiti. *Journal of Geography and Cartography*. 2025; 8(1): 9752. <https://doi.org/10.24294/jgc9752>

## ARTICLE INFO

Received: 21 October 2024

Accepted: 25 December 2024

Available online: 3 January 2025

## COPYRIGHT



Copyright © 2025 by author(s).

*Journal of Geography and Cartography* is published by EnPress Publisher, LLC. This work is licensed under the Creative Commons Attribution (CC BY) license. <https://creativecommons.org/licenses/by/4.0/>

**Abstract:** The hydroclimatological monitoring network in Haiti was inadequate before 2010 due to a lack of meteorological stations and inconsistent data recording. In the aftermath of the January 2010 earthquake, the monitoring network was reconstructed. In light of the prevailing circumstances and the mounting necessity for hydroclimatological data for water resource management at the national level, it is of paramount importance to leverage and optimize the limited available data to the greatest extent possible. The objective of this research is to develop regional equations that facilitate the transfer of climatic data from climatological stations to locations with limited or absent data. Physiographic and climatological characteristics are used to construct the hydrologic information transfer equations for sites with limited or no data. The validity of the regionalization techniques was assessed using cross-validation. The results enable estimation of hydrological events through the specific patterns of behavior of each region of the country, identified in cartography of homogeneous zones.

**Keywords:** hydrological regionalization; homogeneous regions; information transfer; hydrological regime; Haiti

## 1. Introduction

A report by the World Health Organization (WHO) and the United Nations Children's Fund [1] revealed that approximately 2.2 billion individuals globally lack access to safe drinking water. This figure represents one in three individuals worldwide who are unable to access the vital liquid. In Haiti, only 12% of households have access to safe drinking water at their place of residence. This problem is particularly impactful for those residing in rural areas.

According to the World Bank, in 2020, only 43% of Haiti's rural population had access to an essential supply of drinking water, a statistic that represented a decline from 48% in 2015 and 50% in 1990. As evidenced by various studies, these river basins are distributed across approximately 30 river basins, which can be divided into six major river basins: the North-West, North, North-Central, South-Central, South-East and South-West basins. These regions encompass an area of 27,750 km<sup>2</sup> [2].

The country receives an average annual precipitation of 1500–2000 mm, of which only 10% infiltrates into the ground to replenish the aquifers. Consequently, a significant proportion of rainwater is transformed into runoff, which carries sediments and harmful products [3]. One of the most significant challenges in water

management in Latin America and the Caribbean is the dearth of a comprehensive network of climatological and hydrometric stations [4].

In Haiti, as is the case in the most countries in Latin America and the Caribbean, there is a scarcity of climatological monitoring networks, which are subject to numerous limitations. It is also noteworthy that Haiti is vulnerable to natural disasters, underscoring the need for a robust and reliable monitoring network to facilitate accurate forecasting. The country has been grappling with a plethora of challenges pertaining to water management for over three decades. These include deficiencies in water planning, floods, droughts, and the absence of a robust water supply system, urban drainage issues, and inadequate hydraulic engineering designs. Despite the passage of two decades, the fundamental problems remain unresolved. In some instances, residents are compelled to consume river water that has been contaminated. Furthermore, several urban centres in the country are vulnerable to flooding, with incidents occurring even when minimal precipitation is recorded.

The transfer of hydrological data to locations without a meteorological station can be accomplished through the utilization of regional equations. Regional equations are constructed on the basis of the assumption of by region homogeneity. The objective of this study is to present the steps of hydrological regionalization, adapted for Haiti, where climatological records are severely limited or non-existent.

### **1.1. Up-to-date bibliography**

The current literature includes a number of publications on hydrologic regionalization techniques. A substantial number of contributions are published on a regular basis in almost all engineering-related topics, with a particular focus on those related to water resources. A typical example is the utilization of regional models to estimate surface runoff in ungauged basins, which has recently been combined with machine learning [5,6] and satellite observations [7]. To be specific, regionalization techniques are employed in the majority of Latin American and Caribbean countries to facilitate an understanding of the behaviour of homogeneous regions. For instance, Brazil utilises hydrological regionalization for the management of water resources, given the vast extent of its territory [8]. Furthermore, the estimation of extreme events in ungauged basins is extremely reliable when conducted within the context of an effective hydrological regionalization [9].

## **2. Materials and methods**

The quality of the data is dependent on a number of factors, including the process of information collection, the size of the records and the representation of the information on the site. This presents two significant challenges for the hydrologist. The first issue represents the lack of suitable tools to assess the suitability of the data to be used. This requires the definition of precise concepts and their subsequent monitoring. The second issue attends the need to design projects in areas where climatological data or hydrometric records are limited or non-existent. There is consensus among studies on the use of hydrological regionalization as a tool for hydrological transfer [5,6,10–12]. Likewise, Gutierrez-Lopez and Aparicio [4] put

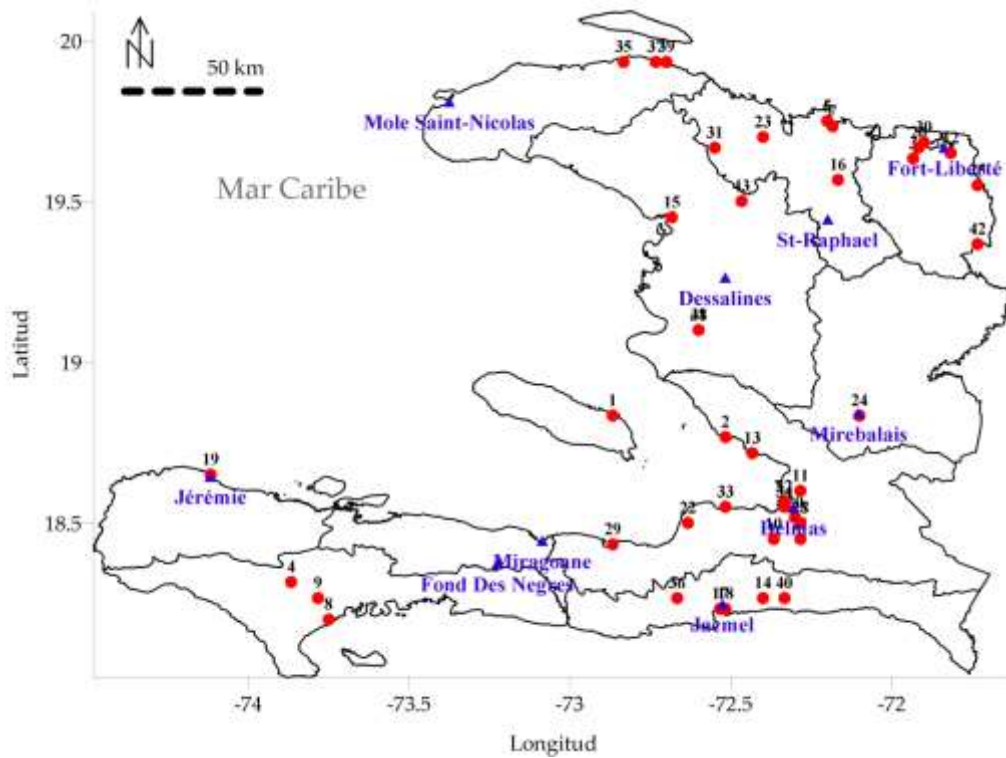
forth a series of steps that must be undertaken for a proper hydrological regionalization.

## **2.1. Haiti's water resources**

The issue of water scarcity represents a significant challenge for a considerable proportion of the Haitian population, particularly those residing in remote areas remote from the capital city. This situation concerns particularly with regard to the provision of drinking water. The most recent statistics indicate that the national drinking water coverage rate is 64%, with 77% in urban areas and only 48% in rural areas. In terms of poverty, the distribution of drinking water is highly precarious throughout Haiti. A review of the data reveals that only 26 municipalities out of 133 (19.5%) receive access to piped water services that can be considered satisfactory. In particular, three departments stand out: The highest rates of access to drinking water are found in Artibonite, Centre and Grande Anse [13]. A considerable proportion of the Haitian population has severely restricted access to drinking water, despite the country's vast potential for renewable water resources, currently distributed in an uneven manner. As evidenced by empirical research, these entities are dispersed across approximately 30 watersheds, which are themselves divided into six principal river basins as previously noted [2]. The country receives an average annual precipitation of 40 billion cubic meters, of which only 10% infiltrates into the ground to recharge the aquifers. The rainy season is defined as lasting 9.8 months, commencing on 24 February and concluding on 19 December. The interval between consecutive periods of rainfall represents 31 days, during which precipitation of at least 13 millimeters is required to qualify as a valid instance. The month with the highest precipitation levels in Haiti is May, with an average rainfall of 54 millimeters. The period of the year during which precipitation is absent lasts for 2.2 months, from 19 December to 24 February. The month with the lowest precipitation levels in Haiti is January, with an annual mean of 11 millimeters of rainfall.

## **2.2. Variables selection**

The availability of hydro-climatological data in Haiti is limited. The country's meteorological network has been severely constrained for several decades, rendering the retrieval of information a challenging endeavour. As previously mentioned, the objective is to utilize the limited data that is available. The data presented in this study are derived from a number of meteorological stations distributed across the Haitian landscape. **Figure 1** illustrates the geographical distribution of meteorological stations that have been operational in Haiti. Stations marked with a red circle are characterized by a limited availability of data and are currently inactive. However, following the 2010 earthquake, some stations have continued in their operation, despite the numerous challenges that persist in Haiti. Consequently, only stations with reliable data were selected for inclusion in the study. **Table 1** presents a description of the physiographic and climatological characteristics of each station included in the study.



**Figure 1.** Location of weather stations in Haiti. Stations in operation (blue triangle). Stations out of operation (red circle).

**Table 1.** Characteristics used in the regionalization model for Haiti.

Variable	Symbology
Latitude	Lat
Longitude	Long
Altitude of the station (m)	Alt
Distance to ocean (m)	DistOc
Maximum annual precipitation height (mm)	HpMax
Minimum annual precipitation height (without taking zeros) (mm)	HpMin
Kurtosis of the maximum annual precipitation height, in mm	HpKurt
100-year precipitation height of the Gumbel Distribution (mm)	Gumb100
Scaling parameter of the Gumbel Distribution	Gumb-u
Scaling parameter of the Poisson/Exponential Distribution	Fuit-lamb

### 2.3. Independence of time series

The recommended methodology for conducting this analysis entails verifying the consistency of the variables to be utilized [14–16]. It is typically advised that, prior to undertaking any further analysis, tests of independence be conducted on historical data samples [17]. Moreover, a regression analysis or an analysis of variance is strongly advised to establish the relationship between variables (regional equations) [18]. It is recommended that an analysis of the covariance function be conducted, followed by a test of independence through the use of a correlogram. In the context of a time series comprising the maximum annual precipitation values ( $X_t$ ), the time lag ( $k$ ) is defined as the time lag in years of each record. In this context ( $N$ )

represents the total number of years of the data set, while  $(\bar{X})$  denotes the mean value of the records. In accordance with this concept, the correlograms  $r(k)$  of the AR( $p$ ) models serve as estimators of the variance ( $C_0$ ) and autocovariance ( $C_k$ ) (Equation (1)). This can be expressed in terms of the lag autocovariance function  $k(C_k)$  as follows. Values close to zero in the correlogram indicate independence, homogeneity and consistency of the series.

$$r(k) = \frac{C_k}{C_0} = \frac{\frac{1}{N-k} \sum_{t=1}^N (X_t - \bar{X})(X_{t+k} - \bar{X})}{\frac{1}{N} \sum_{t=0}^N (X_t - \bar{X})(X_t - \bar{X})} \quad (1)$$

## 2.4. Cartography of hydrologically homogeneous regions

The considerable heterogeneity of regions represents a significant challenge when attempting to regionalize watersheds. The delineation of hydrologically homogeneous regions represents the most challenging aspect of the hydrological regionalization process [19]. The disaggregation of a region into similar subregions serves to reduce the errors generated when transferring hydrological storm data from one basin to another [20]. The techniques employed to ascertain the most appropriate homogeneous regions include residual analysis [21,22], time series statistics analysis [23] and multivariate techniques. It is essential to identify the significant variables or characteristics of the region to be studied prior to implementing any procedure [24,25]. The Surfer and Statistica software were employed for the purposes of cartography and multivariate analysis, respectively.

## 2.5. Construction of regional equations

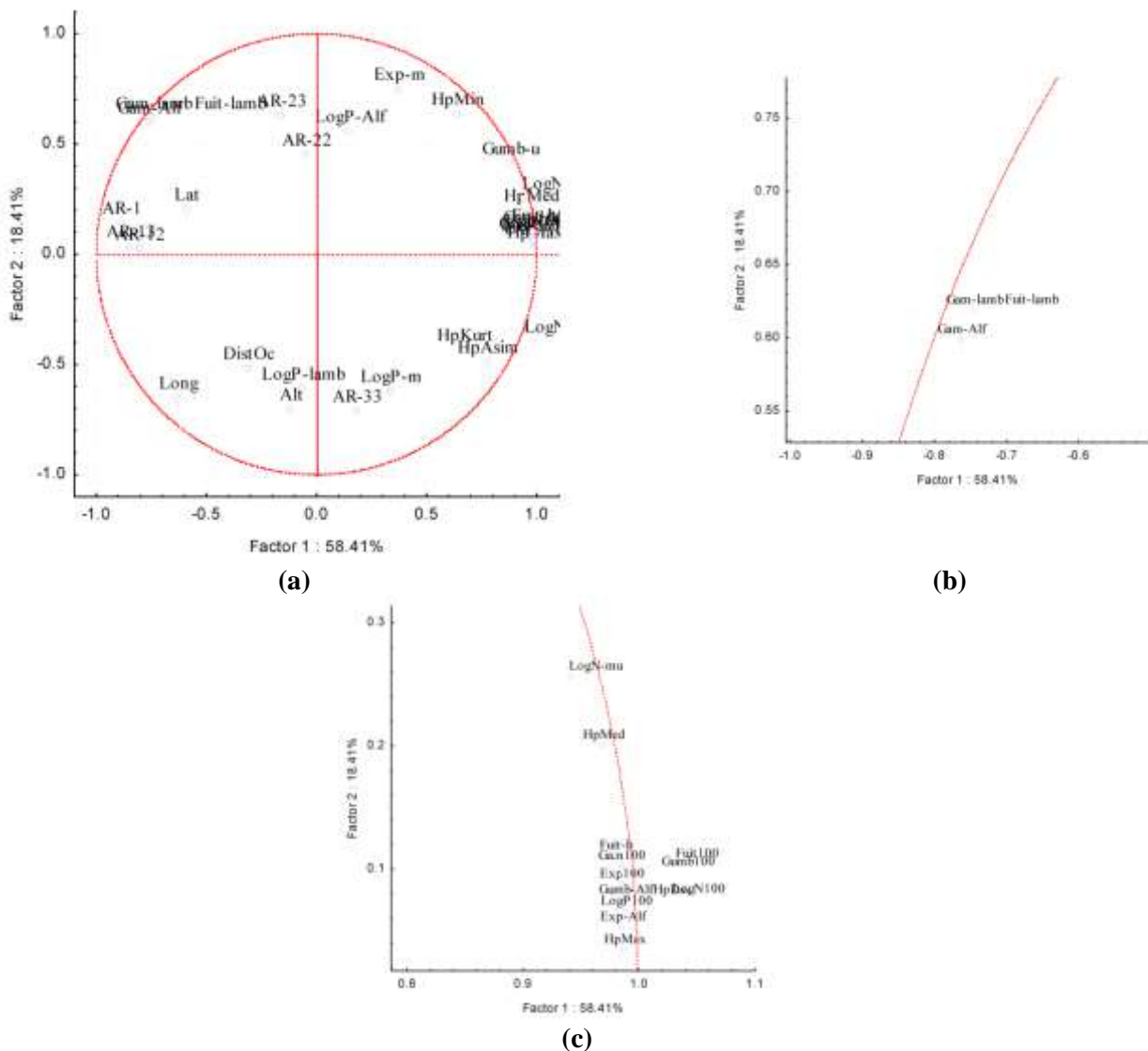
The most frequently employed methodology for establishing regional relationships is multiple correlation analysis. It is conventional practice to correlate the events to be predicted, for example maximum flows or rainfall, with the physiographic characteristics of the basin. A typical practice is to make use of the results of frequency analysis, relating the parameters of the fitting distribution to the physiographic, climatic or environmental characteristics of the region [26–28].

## 3. Results

### 3.1. Multivariate analysis

The selection of variables was based on EOF analysis, as this multivariate method obtains the valuable capacity to determine the spatial dimensionality of the data, thereby allowing the dimensional space of the population to be reduced. The application of EOF to the characteristics matrix yielded the result that the first two principal components together explain 76.82% of the total variability. The first component explains 58.41%, while the second explains 18.41%. The variables were selected on the basis of their position, as indicated by their normalized correlation value, within the quadrants of the correlation circle, with the axes corresponding to the first two principal components (see **Figure 2**). On the other hand, the parameters related to all design events are grouped in the same zone. This means that the

probability distribution used for design events is not very significant for regionalization purposes. It can be postulated that the variables which are proximate to each other in the multivariate space will tend to demonstrate similar behaviour. It is further postulated that variables with a greater projection on the axis of the first principal component are of more considerable importance. Furthermore, variables located in opposite quadrants generally exhibit inverse relationships. Likewise, it can be seen how the parameters related to the probability distribution of Fuites and Gamma have similar behaviors. That is, Fuit-lamb and Gam-lamb are placed in the same place, which makes sense considering the distributions have the same origin as extreme distributions and the parameters are similar.

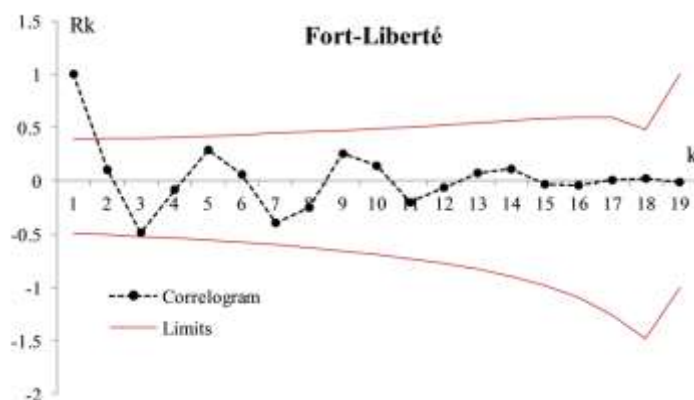


**Figure 2.** Correlation circle with EOF results for components 1–2 of all characteristics used in the study. Total representation 76.82%.

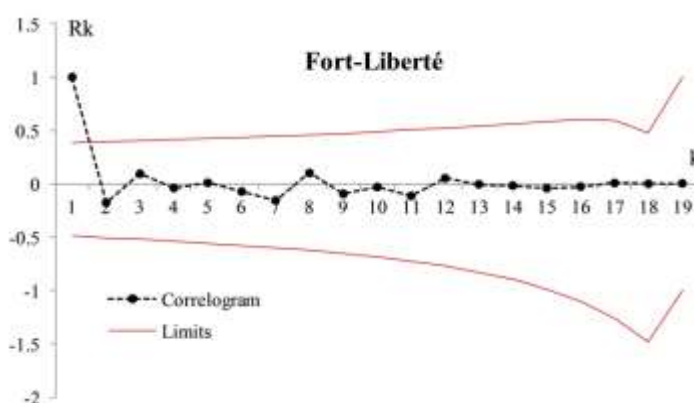
### 3.2. Independence of time series

As previously stated, the autocovariance function and the correlogram permit the verification of the independence of the time series. **Figures 3** and **4** illustrate the evolution of the correlogram over time for the same data series, in this case, for the

annual maximum precipitation recorded at the Fort-Liberté station. From an examination of both figures, it can be concluded that the AR (1) model represent not an appropriate choice for this sample of Fort-Liberté data, given that the correlogram displays variability near the confidence limits. This phenomenon occurs at lag times for  $k = 3$ ,  $k = 5$ ,  $k = 7$  and  $k = 9$ . Nevertheless, an AR (2) model, as illustrated in **Figures 3** and **4**, reveals that only lag time  $k = 1$  exhibits a minor discrepancy, while the remaining series is consistent in its independence across all time lags.



**Figure 3.** Correlogram for station Fort-Liberté AR (1) series of annual maximum rainfall.

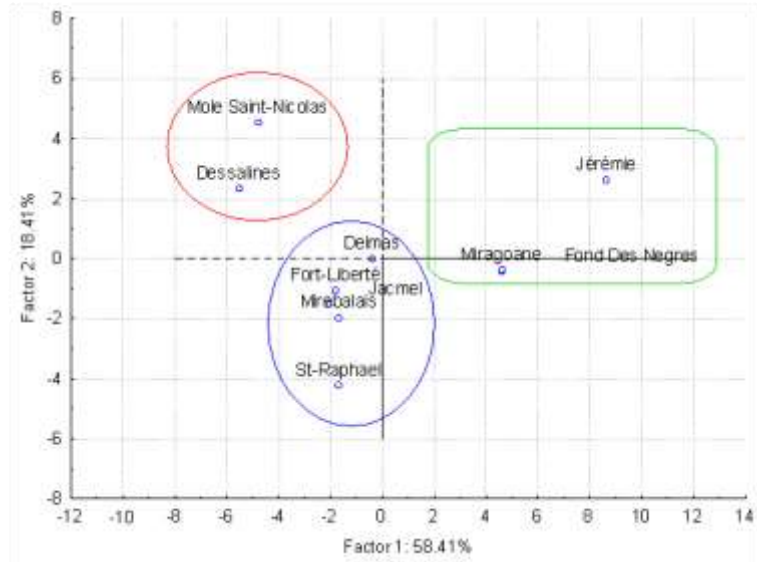


**Figure 4.** Correlogram for station Fort-Liberté AR (2) series of annual maximum rainfall.

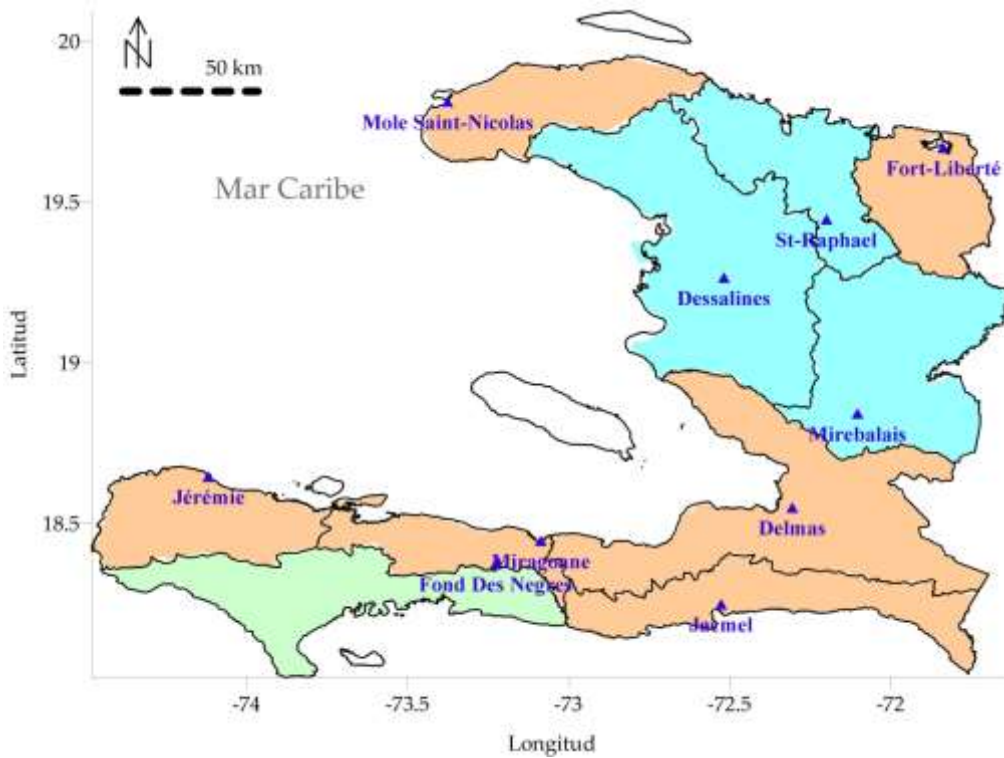
### 3.3. Hydrologically homogeneous regions

The delineation of homogeneous regions was conducted through the utilization of EOF analysis. As illustrated in **Figure 5**, three regions were delineated within the country. The first region is located in the northern part of the country and is defined by the meteorological data from the stations Mole Saint-Nicolas and Dessalines. These regions exemplify the humid conditions characteristic of the Caribbean Sea. The second region is constituted by the stations of Jeremie, Fond Des Negres and Miragone. Situated in the southern region of the country, this area is distinguished by its elevated susceptibility to extreme meteorological events, including hurricanes. The remaining stations constitute the third hydrologically homogeneous region, which represents the dry condition of the country. The results of the configuration of

the hydrologically homogeneous regions, created by the hierarchical clustering technique (dendrograms) and the EOF analysis, respectively, are presented in **Figures 6 and 7**.



**Figure 5.** Results of EOF for homogeneity of regions.



**Figure 6.** Results of the dendrogram analysis for homogeneity of regions.



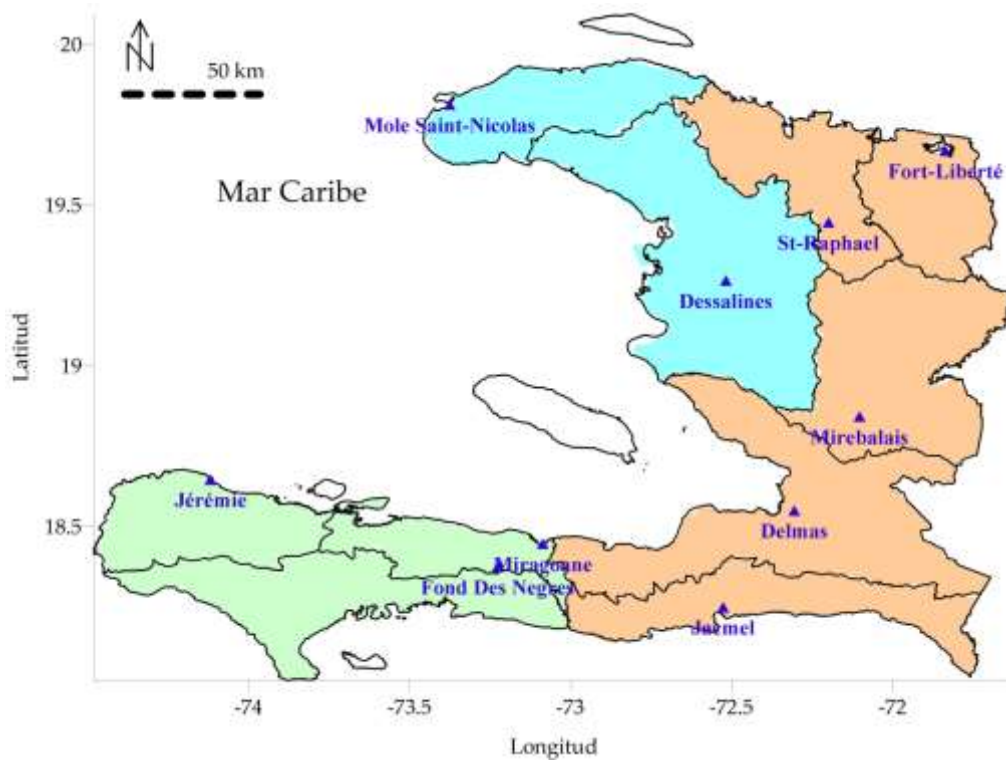


Figure 7. Results of EOF for homogeneity of regions.

#### 4. Discussion

As previously stated, the construction of regional equations is achieved through the application of multiple correlations. In this instance, our focus is on the equations used to forecast the maximum 24-hour rainfall. Initially, all significant variables that have already been analyzed are employed (Equation (2)). It is crucial to exclude the data from the Delmas station, as this allows us to utilize genuine data to compare the values obtained with the regional equations (Table 2). The equation for the entire region, used to forecast the maximum 24-hour rainfall, is:

$$\begin{aligned} \text{HpMax} = & (6.647208129 * \text{Lat}) + (2.540808455 * \text{Long}) - (0.019672487 * \text{Alt}) - \dots \\ & - (4.8897344 * \text{HpMin}) + (3.724475655 * \text{HpKurt}) + (5.903582493 * \text{Gumbu}) \\ & + (0.643863993 * \text{Gumb100}) + (0.535285418 * \text{Fuitlamb}) \end{aligned} \quad (2)$$

The same procedure is used to construct the regional equations for the other homogeneous regions. The equation (Equation (3)) for the region in orange (region A) (Figure 7), for the forecast of the maximum 24-hour rainfall is:

$$\begin{aligned} \text{HpMax} = & -(0.00714296 * \text{Alt}) + (2.95093\text{E}^05 * \text{Dist0c}) + (2.028189568 * \text{HpKurt}) \\ & + (0.778523875 * \text{Gumb100}) \end{aligned} \quad (3)$$

The equation (Equation (4)) for the green region (region B) (Figure 7), for the forecast of the maximum 24-hour rainfall is:

$$\text{HpMax} = (0.664288316 * \text{Alt}) - (4.472\text{E}^{-06} * \text{Dist0c}) + (0.670126744 * \text{Gumb100}) \quad (4)$$

The equation (Equation (5)) for the blue region (region C) (Figure 7), for the forecast of the maximum 24-hour precipitation height is:

$$HpMax = -(0.466551856 * Long) + (4.483E^{-05} * DistOc) \quad (5)$$

The results facilitate the acquisition of events at sites with scarce or no data. In the conventional approach, a multiple correlation is employed to establish relationships between totally the variables and the event to be estimated (Equation (2)). Nevertheless, the methodology detailed herein enables the verification that the utilisation of all variables from the whole stations produces high-quality estimated values [13]. However, given the unique circumstances of Haiti, this research develops the following steps in light of that country’s specific context: (i) Identify the variables that describe the phenomenon to be estimated [4]; (ii) verify the independence of the time series [7]; (iii) cartography of the hydrologically homogeneous regions [9]; (iv) construction of regional equations for hydrological information transfer; and (v) Verification of the applicability of the regional equations [11].

**Table 2.** Overview of the values estimated with the regional equations for the Delmas station. Maximum 24-hour rainfall (mm) and error\*.

Real rainfall height (mm)	Equation (2) Regional Complete	Equation (3) Region A	Equation (4) Region B	Equation (5) Region C
69.89	59.53	70.83	131.31	34.93
	10.36*	0.94*	61.42*	34.95*

## 5. Limitations

The rigorous methodology implements several techniques (EOF, clustering, cross-validation) to ensure robust results. However, the quality is relatively limited by the scarcity and inconsistency of historical data in Haiti, which may limit precision. No statistics can measure the performance of hydrological regionalization. The only way to do this is to cross-validate, i.e. to remove a site’s records from the beginning of the procedure and then verify with the regional equations the precision of the estimation of events at that site (possibly with a quadratic error). The main limitation is obviously the availability of and access to data. However, a complete execution of each of the steps mentioned here, with little data (<20 years), can provide absolutely reliable estimates in sites without records. A fundamental aspect is some researchers confuse regionalization with calculations that are carried out on a routine basis, such as [4]:

- Hydrologic regionalization is not: calculating and drawing iso-lines of some hydrologic variable; this is called cartography.
- Hydrologic regionalization is not: using different physiographic characteristics of a region; this is called spatial ponderation.
- Hydrologic regionalization is not: using probability distributions in several stations of a region; this is called multiple frequency analyses.
- Hydrological regionalization is not: using some method such as Inverse of distance, Kriging, Splines, etc.; this is called spatial interpolation.
- Hydrological regionalization is not: estimating envelope curve function of hydrological events; this is called analysis of extremes.

## 6. Conclusion

The creation of regional equations based on the delineation of homogeneous regions permitted the estimation of precipitation values at the Delmas station. The data obtained at the Delmas station were excluded from the regional analysis. Nevertheless, cross-validation demonstrated that if the complete regional equation had been employed, the resulting error would have been 10.36 mm. However, if the regional equation for the Delmas area had been utilized, the error would have been reduced to a mere 0.94 mm. The regional equations facilitate the reconstruction of rainfall fields and estimation of events at sites without historical records in Haiti. In the absence of a dense network of climatological stations in Haiti, the regional equations, derived from the cartography of hydrologically homogeneous basins, represent a valuable initial tool that can be enhanced with the incorporation of additional physiographic characteristics.

Hydrological regionalization has been identified as an exceedingly effective tool for the reconstruction of precipitation fields, like that of historical maximum rain in Haiti. The results obtained from this reconstruction allow for the estimation of rainfall values, which can then be used in the calculation of water balances. This initial regionalization of Haiti represents a significant advancement in the fact-finding field. Marking the beginning of a research sequence aimed at improving the availability of records and reliable historical series for the management of water resources; in a country confronted with such challenging social and political circumstances. Subsequent research will centre on constructing hydrological transfer equations, with the adjustment parameters of the probability distributions being a key priority. This approach will facilitate the estimation of not only a point value, but also a probability distribution at the analysis site.

**Author contributions:** Conceptualization, CG and AGL; methodology, CG and AGL; validation, AGL; formal analysis, CG and AGL; investigation, AGL; resources, CG; data curation, CG; writing-original draft preparation, CG; writing-review and editing, AGL. All authors have read and agreed to the published version of the manuscript.

**Acknowledgments:** For the support of the Consejo Nacional de Humanidades, Ciencias y Tecnologías (Conahcyt) by means of the postgraduate scholarship.

**Conflict of interest:** The authors declare no conflict of interest.

## References

1. UNICEF & OMS. Progress on household drinking water, sanitation and hygiene 2000-2020: Five years into the SDGs. UNICEF & OMS; 2021.
2. Gonel J. Studying the potential of surface water in Haiti to address drinking water shortages (French). Available online: <https://espace.inrs.ca/id/eprint/428/#> (accessed on 2 June 2024).
3. FAO. Aligned National Action Program to Combat Desertification (French). FAO; 2015.
4. Gutiérrez-López A, Aparicio J. Las seis reglas de la regionalización en hidrología. *Aqua-LAC*. 2020; 12(1): 81-89. doi: 10.29104/phi-aqualac/2020-v12-1-07
5. Hlaing PT, Humphries UW, Waqas M. Hydrological model parameter regionalization: Runoff estimation using machine learning techniques in the Tha Chin River Basin, Thailand. *MethodsX*. 2024; 13: 102792. doi: 10.1016/j.mex.2024.102792

6. Dasgupta R, Das S, Banerjee G, et al. Revisit hydrological modeling in ungauged catchments comparing regionalization, satellite observations, and machine learning approaches. *HydroResearch*. 2024; 7: 15-31. doi: 10.1016/j.hydres.2023.11.001
7. Araujo ACS de, Frizzone JA, Camargo AP de, et al. Discharge sensitivity of collapsible drip tapes to water temperature. *Revista Brasileira de Engenharia Agrícola e Ambiental*. 2021; 25(1): 3-9. doi: 10.1590/1807-1929/agriambi.v25n1p3-9
8. Ye K, Liang Z, Chen H, et al. Regionalization Strategy Guided Long Short-Term Memory Model for Improving Flood Forecasting. *Hydrological Processes*. 2024; 38(10). doi: 10.1002/hyp.15296
9. Neelam TJ, Steinschneider S, Woodward DE, et al. Improved Regionalization of the CN Method for Extreme Events at Ungauged Sites across the US. *Journal of Hydrologic Engineering*. 2024; 29(6). doi: 10.1061/jhyeff.heeng-6180
10. Goovaerts P. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of Hydrology*. 2000; 228(1-2): 113-129. doi: 10.1016/S0022-1694(00)00144-X
11. Guo Y, Zhang Y, Zhang L, et al. Regionalization of hydrological modeling for predicting streamflow in ungauged catchments: A comprehensive review. *WIREs Water*. 2020; 8(1). doi: 10.1002/wat2.1487
12. Finck LF, Leite IR, Almeida AK, et al. A streamflow regionalization method using hydrological data and geoprocessing tools—a Brazilian midwest analysis. *Journal of South American Earth Sciences*. 2024; 133: 104695. doi: 10.1016/j.jsames.2023.104695
13. Hazin LS. Working together for more efficient management of drinking water and sanitation services in Haiti (French). *Publication des Nations Unies*; 2005.
14. Ebisemiju FS. An objective criterion for the selection of representative basins. *Water Resources Research*. 1979; 15(1): 148-158. doi: 10.1029/wr015i001p00148
15. Nathan RJ, McMahon TA. Identification of homogeneous regions for the purposes of regionalisation. *Journal of Hydrology*. 1990; 121(1-4): 217-238. doi: 10.1016/0022-1694(90)90233-N
16. Caratti JF, Nesser JA, Lee Maynard C. Watershed classification using canonical correspondence analysis and clustering techniques: a cautionary note1. *JAWRA Journal of the American Water Resources Association*. 2004; 40(5): 1257-1268. doi: 10.1111/j.1752-1688.2004.tb01584.x
17. Kanishka G, Eldho TI. Streamflow estimation in ungauged basins using watershed classification and regionalization techniques. *Journal of Earth System Science*. 2020; 129(1). doi: 10.1007/s12040-020-01451-8
18. Hu C, Xia J, She D, et al. Parameter Regionalization With Donor Catchment Clustering Improves Urban Flood Modeling in Ungauged Urban Catchments. *Water Resources Research*. 2024; 60(7). doi: 10.1029/2023wr035071
19. Smithers JC, Schulze RE. A methodology for the estimation of short duration design storms in South Africa using a regional approach based on L-moments. *Journal of Hydrology*. 2001; 241(1-2): 42-52. doi: 10.1016/S0022-1694(00)00374-7
20. Leviandier T, Lavabre J, Arnaud P. Rainfall contrast enhancing clustering processes and flood analysis. *Journal of Hydrology*. 2000; 240(1-2): 62-79. doi: 10.1016/S0022-1694(00)00315-2
21. Bhaskar NR, O'Connor CA. Comparison of Method of Residuals and Cluster Analysis for Flood Regionalization. *Journal of Water Resources Planning and Management*. 1989; 115(6): 793-808. doi: 10.1061/(ASCE)0733-9496(1989)115:6(793)
22. Hall MJ, Minns AW, Ashrafuzzaman AKM. The application of data mining techniques for the regionalisation of hydrological variables. *Hydrology and Earth System Sciences*. 2002; 6(4): 685-694. doi: 10.5194/hess-6-685-2002
23. Kachroo RK, Mkhathi SH, Parida BP. Flood frequency analysis of southern Africa: I. Delineation of homogeneous regions. *Hydrological Sciences Journal*. 2000; 45(3): 437-447. doi: 10.1080/02626660009492340
24. Berger KP, Entekhabi D. Basin hydrologic response relations to distributed physiographic descriptors and climate. *Journal of Hydrol*. 2001; 247(3-4): 169-182. doi: 10.1016/S0022-1694(01)00383-3
25. Burn DH, Elnur MAH. Detection of hydrologic trends and variability. *Journal of Hydrology*. 2002; 255(1): 107-122. doi: 10.1016/S0022-1694(01)00514-5
26. Wiltshire SE. Identification of homogeneous regions for flood frequency analysis. *Journal of Hydrol*. 1986; 84(3-4): 287-302. doi: 10.1016/0022-1694(86)90128-9
27. Burn DH. Catchment similarity for regional flood frequency analysis using seasonality measures. *Journal of Hydrol*. 1997; 202(1-4): 212-230. doi: 10.1016/S0022-1694(97)00068-1
28. Castellarin A, Burn DH, Brath A. 2001. Assessing the effectiveness of hydrological similarity measures for flood frequency analysis. *Journal of Hydrol*. 2001; 241(3-4): 270-285. doi: 10.1016/S0022-1694(00)00383-8