Improving Rainfall Fields in Data-Scarce Basins: Influence of the Kernel Bandwidth Value of Merging on Hydrometeorological Modeling

Nicolás Duque-Gardeazabal[1] and Erasmo A. Rodríguez[2]

Abstract: Accurate rainfall fields are important for several hydrological and meteorological applications. In poorly instrumented basins, the lack of rain gauges heavily affects the spatial rainfall estimation, yet neither remote sensed nor climate model estimates are good enough for management applications. To tackle this problem, we investigate the impacts of combining both sources of information by varying the kernel bandwidth value of the double smoothing merging algorithm and analyzing the error of the rainfall fields. We explored the correlation between rain gauge density and bandwidth and compared the results against classical geostatistical interpolation methods based solely on in-situ measurements. Propagation of rainfall error into hydrological modelling is usual, and therefore we evaluated the influence of the bandwidth in streamflow simulations implementing two hydrological models. The hydrological evaluation considered the analysis of hydrological signatures rather than just performance metrics. We found that there is a clear correlation between kernel bandwidth and monitoring network density and that the bandwidth also affects hydrological performance. Simple bilinear downscaling did not produce a significant difference in meteorological or hydrological errors, and rain gauge network configuration also impacts the error of the field. We conclude that merging outperforms the results of classical interpolation methods, in some cases by 20% or 50%, suggesting the suitability of the method for being applied in data-scarce domains. **DOI: 10.1061/JHYEFF.HEENG-5541.** *© 2023 American Society of Civil Engineers.*

Practical Applications: Reliable rainfall fields are important for estimating water resource availability and other related applications. However, using just rain gauges can misrepresent the estimated spatial variability of the field. Around the world, the number of rainfall gauges has been decreasing, making this task more challenging. Satellite and reanalysis data can help to overcome these problems, but their direct use is also insufficient. Merging the two mentioned data sources (in-situ and reanalysis data) is the option the authors implemented in this work. The authors focused on evaluating the advantages and disadvantages of the merging by comparing both rainfall and streamflow observations with estimates simulated by a hydrological model. The authors chose the double smoothing algorithm because it is oriented to enhance rainfall estimates in areas with very sparse rain gauge data, yet we evaluate its behavior in two basins with different monitoring conditions. The authors conclude that the merging outperforms the results of classical interpolation methods, in some cases by 20% or 50%, suggesting the suitability of the method for being applied in data-scarce domains to improve hydrological modelling estimates and use their results in water resource management and planning, in climate classification, and in the study of droughts and floods, among others.

Introduction

Rainfall is recognized as one of the most important drivers of the water cycle over land and a key variable for studying water availability on the surface and subsurface (Chow et al. 1988). Several disciplines, including hydrology, meteorology, agriculture, risk assessment, ecology, and others (García et al. 2016), apply quantitative estimates of rainfall to acquire insights into the water cycle and other variables. For instance, in hydrological and land surface modeling, distributed rainfall fields are essential inputs (Arnaud et al. 2002; Gupta et al. 1999; Refsgaard 1997). In addition to hydrological modelling, rainfall field estimates are required to perform climate classification, calculate ecological and drought indices (Beck et al. 2018; Rodríguez et al. 2020), and also perform analysis of the spatiotemporal dynamics of precipitation (Espinoza Villar et al. 2009).

Traditionally, rainfall has been measured by rain gauge networks located on land following the World Meteorological Organization guidelines (WMO 2008). Monitoring networks must also be continuously adjusted to properly quantify the temporal and spatial scales of the investigated hydrometeorological phenomena (Blöschl and Sivapalan 1995). From these point observations, rainfall fields are created by applying interpolation methods (Webster and Oliver 2007), using deterministic (Hurtado-Montoya and Mesa-Sánchez 2014; Ly et al. 2011) or geostatistical algorithms (Grimes and Pardo-Iguzquiza 2010). However, maintaining and expanding the monitoring networks requires large amounts of resources to guarantee records in distant or almost inaccessible environments. The high demand for resources and development of modern technologies to sense precipitation remotely have caused a decrease in the number of rain gauges around the globe (García et al. 2016). Satellite remote-sensed observations and meteorological reanalysis also provide estimates of precipitation, though they

[1]Researcher, Grupo de Investigación en Ingeniería de los Recursos Hídricos, Universidad Nacional de Colombia, Av. Carrera 30 # 45-03 - Ciudad Universitaria Edificio 409, Lab. Hidráulica, Bogotá 111321, Colombia (corresponding author). ORCID: https://orcid.org/0000-0002-2631-8573. Email: nduqueg@unal.edu.co

[2]Associate Professor, Grupo de Investigación en Ingeniería de los Recursos Hídricos, Universidad Nacional de Colombia, Av. Carrera 30 # 45-03 - Ciudad Universitaria Edificio 409, Lab. Hidráulica, Bogotá 111321, Colombia. ORCID: https://orcid.org/0000-0003-4303-6460. Email: earodriguezs@unal.edu.co

Note. This manuscript was submitted on June 15, 2021; approved on February 7, 2023; published online on April 18, 2023. Discussion period open until September 18, 2023; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Hydrologic Engineering***, © ASCE, ISSN 1084-0699.

use other technologies to measure the amount of water that falls over land or sea (Barrett 1975; Dee et al. 2011). Although these alternative data sources can help to tackle the problem of reduction in the number of stations, they suffer from temporal errors and usually exhibit a consistent bias (Dinku et al. 2010; Zulkafli et al. 2014). At the basin scale, the measurement errors do not allow for a direct application on, for instance, hydrological modelling and forecasting (Duncan et al. 1993; Muñoz et al. 2015; Serrat-Capdevila et al. 2014).

To equalize the patterns of the remotely sensed information with the actual rainfall field, reduce its bias, and enhance the applicability options, downscaling techniques have been applied to the raw estimates. The procedures normally use ancillary variables such as vegetation indices (Immerzeel et al. 2009), physiographic variables from digital elevation models (Long et al. 2016), and combinations of these (Ceccherini et al. 2015; López López et al. 2018). Several downscaling methods use spatially constant relationships, which ignore that the variation between the precipitation and alternative variables is nonstationary (Park 2013). Ceccherini et al. (2015) provide a framework called geographic weighted regression (GWR) to overcome this limitation and linearly combine several variables.

Furthermore, the "ground truth" provided by the rain gauges is another alternative to improve those remotely sensed or reanalysis background fields (background, i.e., a field that has not been produced by the estimates of rain gauges). The merging of these sources has great potential to produce accurate rainfall fields and patterns (López López et al. 2018; Nerini et al. 2015; Serrat-Capdevila et al. 2014). The techniques used to make the correction can be underpinned by inverse variance-based weightings for coefficients or linear combinations of them (Woldemeskel et al. 2013), bias correction by adding or multiplying coefficients (Vila et al. 2009), conditional merging (Ehret et al. 2008), spectral analysis combination (Heidinger et al. 2012), and Bayesian conditioning (Todini 2001). The background estimates can be used to improve the spatial structure estimation done by the variogram in geostatistical methods or as a covariate in the cokriging method (Chappell et al. 2012), and the point residuals between the background field and the in situ estimates can feed interpolation methods used later to merge the data sources (Li and Shao 2010; López López et al. 2018). Nerini et al. (2015) investigated and compared the performance of several techniques used to improve the rainfall error in the upper Peruvian Amazon Basin and recommended Bayesian conditioning and double smoothing (DS) algorithms (Li and Shao 2010; Todini 2001), highlighting the cases where their properties can be advantageous (considering the assumptions, limitations of the methods, and available density of measurement).

The double-kernel smoothing merging algorithm was proposed by Li and Shao (2010) as an improved statistical approach based on nonparametric statistics and data assimilation techniques. It emphasizes the blending of data sources on the spatial correction for sparse rain gauge networks on discontinuity correction, and more importantly, it does not follow Gaussian and second-stationary assumptions bound to geostatistical interpolation methods (Cressie 1994; Li and Shao 2010). The method has been used in several studies (Long et al. 2016; Nerini et al. 2015), and the error has also been assessed in complex high-relief terrain (Duque-Gardeazábal et al. 2018); however, the influence of the kernel bandwidth value and bilinear downscaling on the performance of hydrological modelling applications has not been determined, as far as we know.

Therefore, the objective of this study was to evaluate the characteristics of the error and influence of the kernel bandwidth values and trend produced by the DS methodology, especially on surface hydrology. We performed the evaluation in two phases. First, we analyzed the influence of rain gauge density in a watershed with a dense rain gauge network to determine the influence of station density on the error and trend in the bandwidth parameter; second, the lessons gained from the first phase were applied and further studied through hydrological modelling in a data-scarce basin. Finally, we discuss the results and draw conclusions from them.

Study Area and Data

River Basins and Local Hydrometeorological Data

The watersheds considered in the analysis were the Sogamoso and the Casanare River basins in Colombia, delimited until the Tablazo (24067010) and Cravo Norte (36027050) streamflow gauges, respectively. They correspondingly have areas of 20,777 and 14,740 km^2 (see Fig. 1). Both basins have their headwaters in the Andes mountain range, with the Sogamoso basin entirely located in a complex high-relief terrain, whereas the Casanare basin is composed mainly of savannas in the middle and lowlands. Moreover, both basins are located in the intertropical convergence zone, which causes many convective cloud systems with heavy precipitation to form over them (Poveda et al. 2006); also, the Andes mountain range and the highly variable topography enrich the land–atmosphere interactions. As a result, both basins have highly complex rainfall fields.

The Sogamoso river basin consists of the confluence of the Suarez and the Chicamocha rivers. The aridity index of the Suarez basin indicates a moderate surplus of water, according to IDEAM (2019). On the other hand, the Chicamocha river watershed is mainly classified as a deficit basin, with the upstream area of the basin having a moderate water surplus. The precipitation on the headwater catchments registers the lowest rainfall values of about 600 mm/year, and the maximum precipitation is approximately 4,000 mm/year; the basin presents a bimodal rainfall regime with higher values in April–May and October–November (IDEAM 2019). The watershed is highly influenced by the El Niño–Southern Oscillation (ENSO) climate driver, which produces a large interannual variability (Poveda et al. 2006).

National hydrometeorological data are gathered and managed by the Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM). The institute performs some data validation analysis before publicly delivering them through a web portal (IDEAM 2022); information is classified by variables from which we retrieved the precipitation and streamflow time series for the region. Within the Sogamoso basin considered area and surroundings, there are 228 rain gauges, from which 156 were selected for the study because they had more than 50% of the data in the period 1980–2012; consequently, the rain gauge density on the basin is 7.56 gauges/1,000 km^2, which is high enough to allow performing experiments with the rain gauge density in artificially randomized monitoring networks (see section "Evaluation of Rainfall Fields against Rain Gauges").

The Casanare River basin, which is adjacent to the Sogamoso watershed, is a typical example of a data-scarce basin (see Fig. 1). Within it, there are just six rain gauges (there are others in the surroundings that were included in the analysis, to make 16 rain gauges, but the rain gauge density of the network is still approximately 0.4 gauges/1,000 km^2). Precipitation intraannual variability is very high, with a dry period between December and March and highest rainfall occurring from June to September (mean annual precipitation is approximately 2,250 mm/year with a monomodal regime); rainfall in this basin is also affected by the ENSO

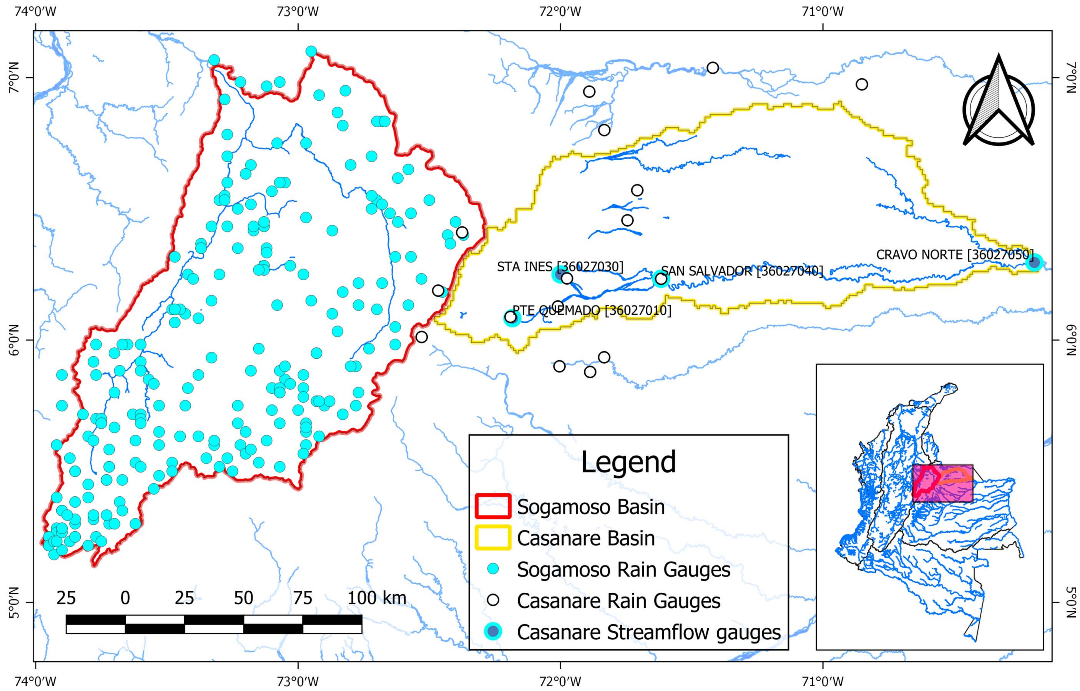

Fig. 1. Location of the Sogamoso and Casanare River basins and the hydrometeorological monitoring networks.

climate driver. From four available streamflow gauges, two of them have less than 1 year of record, and just one has more than 10 years; analysis of possible outliers was performed on the time series, retrieving those values that were impossible or incongruous considering the Oceanic Niño Index (ONI). The Cravo Norte [36027050] and San Salvador [36027040] streamflow gauges (see Fig. 1) were selected for this study because their records are within the period 1996–2012; the others were discarded due to the high number of missing values.

Alternative Estimates of Hydrometeorological Variables

We used multisource weighted-ensemble precipitation (MSWEP) version 1.1 as the rainfall background field for the DS merging algorithm. MSWEP is a combination of several precipitation data sets that includes gauges, satellite, and reanalysis products (Beck et al. 2017b). It is based on the disturbance of the CHPclim long-term precipitation climatology, which was replaced or corrected in certain parts of the world with more accurate regional data sets and inverse water balance estimates based on streamflow measurements. The disturbances are determined by the anomalies of the gauge, satellite, and reanalysis data sets at several time scales (e.g. monthly, daily and 3-hourly, when available); that temporal variability is averagely weighted depending on the regional availability of rain gauges and the comparative performance of the remote sensed and reanalysis products. The anomalies are applied first to the long-term climatology at monthly time scale using the satellite and reanalysis data sets; afterward, using gauge data, daily weights are calculated and applied to the monthly precipitation, and so forth. The data sets used for its creation include two based on rain gauge interpolation observations (CPC Unified and GPCC), three satellite remotely sensed precipitation products (CMORPH, GSMaP-MVK, and TMPA 3B42), and two atmospheric reanalyzes (ERA-Interim and JRA-55). The MSWEP data set v1.1 has a 0.25° spatial and daily temporal resolution.

To support hydrological modelling, we used evapotranspiration estimates from a product developed in the framework of the eartH2Observe project (Schellekens et al. 2017). The downscaled Penman-Monteith ECMWF v1 evapotranspiration at a spatial resolution of 0.083° was developed using the WFDEI data (Weedon et al. 2011), and the procedure to increase its resolution was based on the correlation with topographical variables such as elevation, aspect, slope, and the position of the light and shadows over the hillslope (Sperna Weiland et al. 2015). The data set is directly used in the model as meteorological forcing. That research used the 90-m–resolution DEM from the Shuttle Radar Topography Mission for calculations that required the topographical values. They also applied a temperature lapse rate correction and air pressure based on hydrostatic pressure equation and altitude.

Methods

Considering the objective of evaluating the influence of the kernel bandwidth on the DS method, we performed experiments with artificially created rain gauge networks. The results from the experiments revealed trends in the characteristics of the technique that were complemented by insights from a hydrological modelling application in a data-scarce basin. In the following section, we describe the methods and tools used to investigate the performance of the DS merging algorithm.

Double Kernel Smoothing

The DS technique is a residual-based interpolation method that is adapted to sparse gauge networks (Li and Shao 2010). The residuals are calculated between a background field and point observations; the field may come from any source (remotely sensed, simulated, etc.), and its purpose is to help to preserve the main spatial patterns of the phenomenon in vast regions. The gauges are locally used to adjust the amount and the pattern of the final rainfall field using kernel density estimation functions. Li and Shao (2010) warn that a coarse background field can induce a significant boundary bias and advise performing a previous downscaling of the product that generates a smoothed background field. Other authors have performed complex types of downscaling, previously applying the residual interpolation techniques (Long et al. 2016; López López et al. 2018), yet the original DS method recommends doing one based on area proportions or a bilinear downscaling. We compared the performance of both strategies, with bilinear and without smoothing of the background field (the latter consists of just resampling the field using the nearest neighbor estimate). Eqs. (1)–(4) show the mathematics of the DS algorithm

$$X_M(\mathbf{s}) = X_B(\mathbf{s}) - \hat{\mu}_B(\mathbf{s}) \tag{1}$$

$$\hat{\mu}_B(\mathbf{s}) = \frac{\sum_{i=1}^{n} K_2(\|\mathbf{s}-\mathbf{s}_i\|/h_2)D(\mathbf{s}_i) + \sum_{j=1}^{m} K_2(\|\mathbf{s}-\mathbf{s}_j\|/h_2)\hat{D}(\mathbf{s}_j)}{\sum_{i=1}^{n} K_2(\|\mathbf{s}-\mathbf{s}_i\|/h_2) + \sum_{i=1}^{m} K_2(\|\mathbf{s}-\mathbf{s}_j\|/h_2)} \tag{2}$$

$$\hat{D}(\mathbf{s}_j) = \frac{\sum_{i=1}^{n} K_1(\|\mathbf{s}-\mathbf{s}_i\|/h_1)D(\mathbf{s}_i)}{\sum_{i=1}^{n} K_1(\|\mathbf{s}-\mathbf{s}_i\|/h_1)} \tag{3}$$

$$D(\mathbf{s}_i) = X_B(\mathbf{s}_i) - X_O(\mathbf{s}_i) \tag{4}$$

where X_M, X_B, and X_O = estimates of the merged field, background field, and rain gauge estimates, respectively; $\hat{\mu}$ = estimated expected value of the background error field; $\mathbf{s}$ = position vector (longitude and latitude coordinates); D and $\hat{D}$ = residual and pseudoresidual; K_1 and K_2 = kernel functions; h_1 and h_2 = kernel bandwidths; i and n = index and total number of rain gauges; and j and m = another index and total number of points or cells considered.

Kernel density estimates are effective methods for recovering uni–multivariate probability distributions that pretend not to fall into the problems of parametric density estimation (Lall et al. 1993). They can be applied to discrete or continuous variable observations to estimate the variation of frequency or probability of a random variable, and there exist various function estimators. The estimator functions depend on a fixed parameter called kernel bandwidth, which is irrespective of the location of the value in the distribution (Rajagopalan et al. 1997). The value of the bandwidth in probability density estimation can be obtained by minimizing the estimated average mean integrated square error (AMISE) or with methods based on the sample variance of the random variable. However, the objective of using kernel functions in the DS technique is to interpolate the point residuals between the background field and rain gauge observations, and it is not related to the estimation of the frequency distribution of the variable.

Regarding the type of kernel for merging, we used the Gaussian kernel because it has infinite support and has also been proved not to increase the estimation error, compared with the Epanechnikov or the bisquare kernel, which mathematically gives the lowest mean square error (Rajagopalan et al. 1997). To evaluate the effect of rain gauge density on the estimation of the error and bandwidth, we calibrated the bandwidth values (both h_1 and h_2) for each artificial gauge network using the shuffled complex evolution algorithm (Duan et al. 1992). This is a heuristic method for global optimization that combines the concepts of competitive evolution of communities within a population, controlled random search for evolution of those communities, and complex shuffling of them. Zulkafli et al. (2015) alternatively fixed the value of the bandwidth based on the distance between observations, similarly to what is used in kernel density estimation using the Silverman rule of thumb (Silverman 1986). Yet, there are other equations that minimize the AMISE assuming a parametric reference distribution, such as normal (PR-Normal) or exponential (PR-E). We calculated the value of the kernel using these rules to compare the values obtained with those from the calibration.

Interpolation Methods

We tested the performance of traditional interpolation methods to compare the skill of the DS merging with these interpolation techniques. Even though the selection of an appropriate method depends on the characteristics of local rainfall, morphological features, and rain gauge network configuration (Grimes and Pardo-Iguzquiza 2010), we investigated the most common geostatistical methods, including ordinary kriging (OK), universal kriging (UK), kriging with external drift (KED), and regression kriging (K.Reg); a description of their variations can be found in Rogelis and Werner (2013). From deterministic methods, we chose inverse distance weighted (IDW) because it is the most applied method in environmental sciences (Webster and Oliver 2007) and also applied the DS merging for interpolation, assuming that the background field is lacking and only gauge information is available, called DSg from here onward for comparison.

All kriging methods are based on the same basic linear estimator that is influenced by the spatial structure, represented by an experimental and afterward an adjusted model variogram (Cressie 1994). Furthermore, all kriging methods seek to minimize the variance of the estimation under the constraint of zero bias estimation. The spatial structure must ensure that the covariance of the spatial stochastic process depends exclusively on the separation of the realization points, and the correlation between the observations must reduce with the increase of the separation (Cressie 1994). Moreover, the mean of the process must be constant. These assumptions are known as second-order stationarity. To explore the fulfilment of the second stationary assumptions, we constructed dispersion diagrams and scatter plots on selected days showing that in several of them, removal of the trend should be applied; the UK, KED, and K.Reg focus on that extraction, yet we applied the OK to assess its performance regardless. The gauges' geographical position and elevation were used as covariates to retrieve the trend (see Rogelis and Werner 2013) for details of the trend removal and their advantages). From the several models of variograms, the literature research revealed that the Gaussian model is the one that best adjusts the interpolation of rainfall (Wagner et al. 2012).

Regarding deterministic interpolation methods, we used IDW with an exponent of distance weighting equal to 2 ($\beta = 2$) because it is the most used exponent in environmental studies (Webster and Oliver 2007). The other deterministic interpolation method considered was the DSg. This method is influenced by the kernel bandwidth values that are also calibrated; we only focused on the final performance error of this technique and not on determining the variation of the kernel bandwidth values against the gauge density.

Evaluation of Rainfall Fields against Rain Gauges

To account for the stations' spatial distribution, which can be present in other basins, we created randomized rain gauge networks from the Sogamoso river basin set of stations available. Five rain gauge densities were defined by sampling the 156 stations with different percentages; also, each density was sampled three times to consider the differences in results associated with the network configuration (see details in Table 1; the bandwidth values are the average of the individual values of each rain gauge configuration network; because h_2 depends on the distance of pseudoresiduals its value is constant for all densities). The gauges that were sampled in each network were used for interpolation or merging and to perform a one-time cross-validation (CV) error evaluation.

Table 1. Rain gauge network densities with their corresponding number of stations; h_1 and h_2 are the kernel bandwidths of merging

Percentage of network density	Merging parameter	100%	70%	50%	30%	10%
Rain gauges		156	109	78	46	15
Density (gauges/1,000 km^2)		7.69	5.38	3.85	2.27	0.74
Number of artificial network configurations		1	3	3	3	3
Rain gauges in IV[a] methodology		0 (not apply)	47	78	110	141
Silverman's rule (km)	h_1	4.75	5.13	5.42	5.99	8.16
	h_2	3.42	3.42	3.42	3.42	3.42
PR-normal (km)	h_1	9.55	10.36	10.92	12.05	16.39
	h_2	6.87	6.87	6.87	6.87	6.87
PR-exponential (km)	h_1	8.84	9.58	10.10	11.14	15.16
	h_2	6.36	6.36	6.36	6.36	6.36

[a]IV = independent validation.

Those gauges that were not used in the merging or interpolation were used for independent validation (IV). The error was measured with the root mean square error (RMSE) and the percent bias (PBias).

Hydrological Modelling Performance Evaluation

From the experiments performed in the data-rich basin (Sogamoso River), we derived the trend and value of both kernel bandwidth parameters, which were used to create the fields in the data-scarce basin (Casanare River). Specifically, the value was transferred from the data-rich basin based on the network density in the data-scarce basin to create one rainfall field that was tested with hydrological modelling.

Furthermore, in the data-scarce Casanare River basin, we tested the correspondence of rainfall fields with the streamflow daily observations in two gauges (see section 2.1). Other studies have evaluated precipitation data sets through hydrological modelling, in which rainfall is transformed into simulated streamflow (using process-based or conceptual models) and compared with in-situ data (Beck et al. 2017c). Although the modelling of the near-surface part of the hydrological cycle is not perfect, the peak flows generated by properly estimated rainfall field must correspond with the observed streamflow. On the contrary, a deficient rainfall field would not be able to represent the observed hydrographs. We followed the recommendation of Beck et al. (2017b), which highlights that a recalibration of the hydrological model for each data set is needed to reduce the uncertainty in the modelling and achieve the best possible performance related to meteorological forcing. Therefore, every precipitation data set created for the Casanare basin was used to force each hydrological model that was also recalibrated for each data set. The list of rainfall fields investigated can be found in Table 2. We created several rainfall fields from the DS merging algorithm by varying the kernel bandwidth to evaluate their influence on the hydrological model performance because it is one of the aims of the research; we equalized the values of h_1 and h_2 because they were shown to have similar sensitivity patterns (Duque-Gardeazábal et al. 2018).

For hydrological modelling, we used the distributed open-streams wflow-hbv model (Schellekens 2020) and the semidistributed TOPMODEL version available in R (Beven et al. 1995; Buytaert 2018). We chose both classes of models (semidistributed and distributed) to see the effect of spatially aggregating the rainfall field in the hydrological model performance. Both tools represent conceptual hydrological models based on water balance calculated in diverse tanks. wflow-hbv is based on the HBV-96 model (Lindström et al. 1997), but its structure is applied on each cell in the grid; it has three components (precipitation and snow, soil moisture, and runoff response) and a kinematic wave routing scheme that is used in conjunction with the accumulation of runoff to find river discharge (Schellekens 2020). On the other hand, TOPMODEL uses the reduction of spatial complexity through a topographic wetness index that classifies the terrain response to precipitation in classes; its hydrological structure is based on the assumptions that the processes in the saturated zone follow a steady-state scenario and hydraulic conductivity decreases exponentially with depth (Beven et al. 1995).

Both model parameters were calibrated with the dynamical dimension search (DDS) algorithm (Tolson and Shoemaker 2007). Because the Cravo Norte gauge has the longest streamflow record, it was designated as the boundary condition for calibration of the model's parameters in the period 1997–2006 (hereafter called boundary conditioning station), and all the results reported (used for comparison between methods) are from the validation period that goes from 2007 to 2012. The San Salvador gauge was used as a blind gauge to assess the robustness of the modelling exercise, the appropriateness and correspondence between the rainfall field and the streamflow observations in a subcatchment of the Casanare basin (hereafter called blind point).

We measured the robustness of the hydrological simulations in terms of the Nash–Sutcliffe efficiency (NSE) coefficient (Nash and Sutcliffe 1970). We chose this metric because the NSE is focused on the evaluation of peak flows generated by peaks in precipitation and because it is also sensitive to long-term bias, which is related to rainfall bulk volume of estimation. Moreover, we further evaluate the simulated and the observed hydrographs by comparing the

Table 2. List of rainfall fields used to force the hydrological models in the Casanare basin

Method	Characteristics
DS	Variation of the kernel bandwidth between 10 and 50 km with increments of 5 km and from 50 to 100 km with increments of 10 km: 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 km, Silverman's rule
DSg	Kernel bandwidth assigned by calibration against the 16 rain gauges in the Casanare basin
IDW	Interpolation exponent $\beta = 2$
MSWEP	Raw and bilinear downscaled data sets

hydrological signatures of the simulated and observed time series. The chosen signatures (Beck et al. 2017a) include the higher segment of the flow duration curve (FDC) represented by the value of discharge exceeded 2% of the time; the peak flow value, defined by the 1% of the FDC; and the variance of the streamflow record.

Results

Evaluation of Rainfall Fields against Rain Gauges (In-Situ Data)

The error was evaluated using daily precipitation estimations from all methods in the period 1994–1996 (CV is computationally intensive). We chose this time interval because positive, neutral, and negative phases of the ENSO phenomenon occurred in Colombia and affected the Sogamoso basin. Fig. 2 shows the error reduction that is achieved by blending the background estimates with the local rain gauges. It is also clear that whereas the network density increased, the error decreased, and for low densities, the DS was closer to the background field performance. The error also depends on the network configuration, and in some cases the differences can be significant (1 mm/day for the density of 4 gauges/1,000 km^2). The effect of applying the bilinear downscaling to the background MSWEP field was hardly noticeable (just 0.25 mm/day), but it must be highlighted that in some cases, bilinear downscaling raised the error rather than reducing it. The differences in the methodologies of evaluation (CV and IV), shown in Fig. 2 for the density of 0.74 gauges/1,000 km^2, are because the CV methodology was calculating the error with a low number of gauges.

Regarding the variation of bandwidths, both kernel bandwidth values (h_1 and h_2) increased when the network density was reduced either if downscaling was performed or if it used the raw background field (see Figs. 3 and 4). The same trend was found when using the aforementioned rules of thumb, yet the calibrated values were higher than those listed in Table 1 for the corresponding network density. The downscaling did not change the calibrated value of the merging a significant amount (more clearly depicted in the right-side panel for the IV methodology), and the values of kernel bandwidth tended to stabilize on a value of 30 km when the density was below 3 gauges/1,000 km^2.

From the results of the geostatistical interpolation methods, we found that regression kriging was the one with the lowest error. This is because a regression that includes the coordinates of gauges on a polynomial of second order and the elevation as covariates gives the method better skill to perform the interpolation (see Fig. S1). The KED and the KU reported good performance but not as good as K. Reg because they include the position and the elevation covariates only as a linear trend.

Fig. 5 depicts the error of the K.Reg method against the density of rain gauges, jointly with the DS error (without bilinear

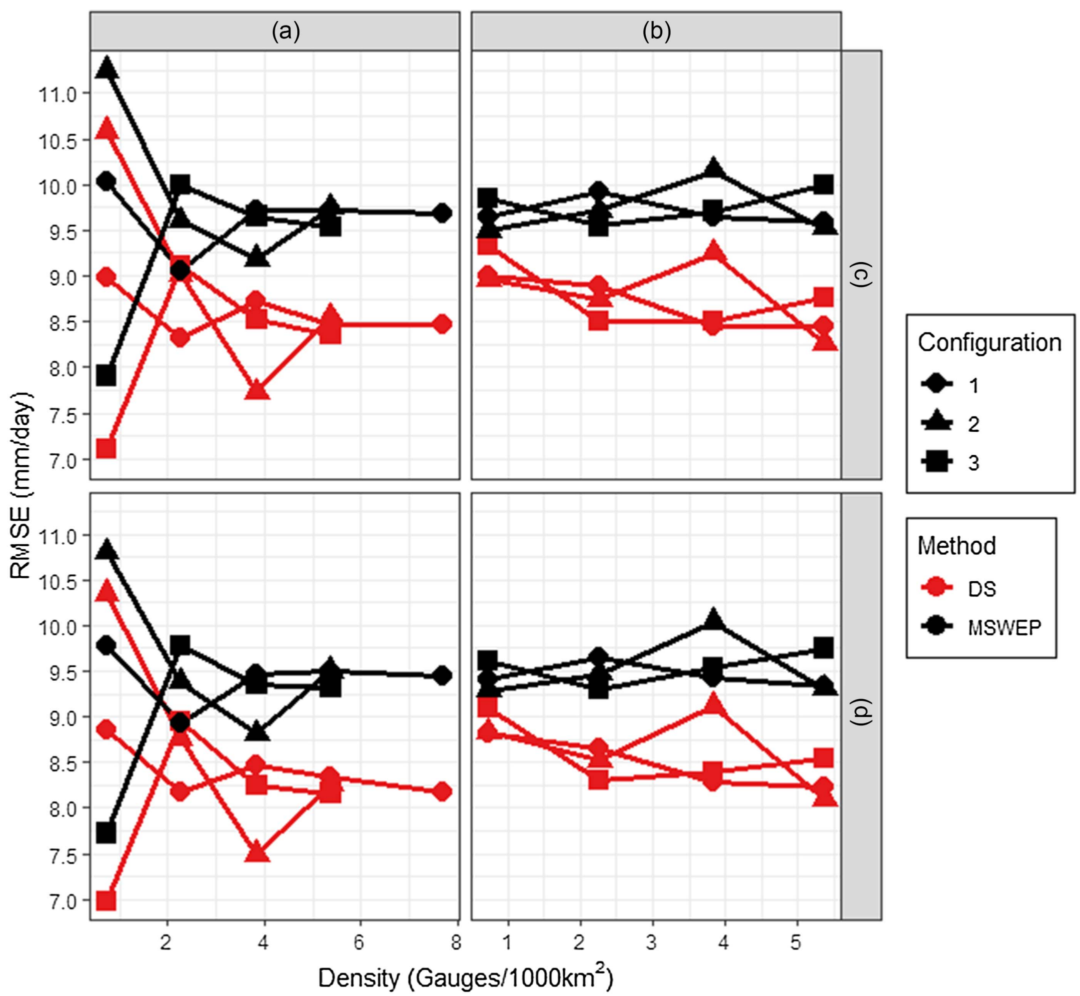

Fig. 2. Variation of RMSE of the DS merging and MSWEP. Results for (a) cross-validation; (b) independent validation; (c) raw MSWEP; and (d) bilinear downscaled MSWEP.

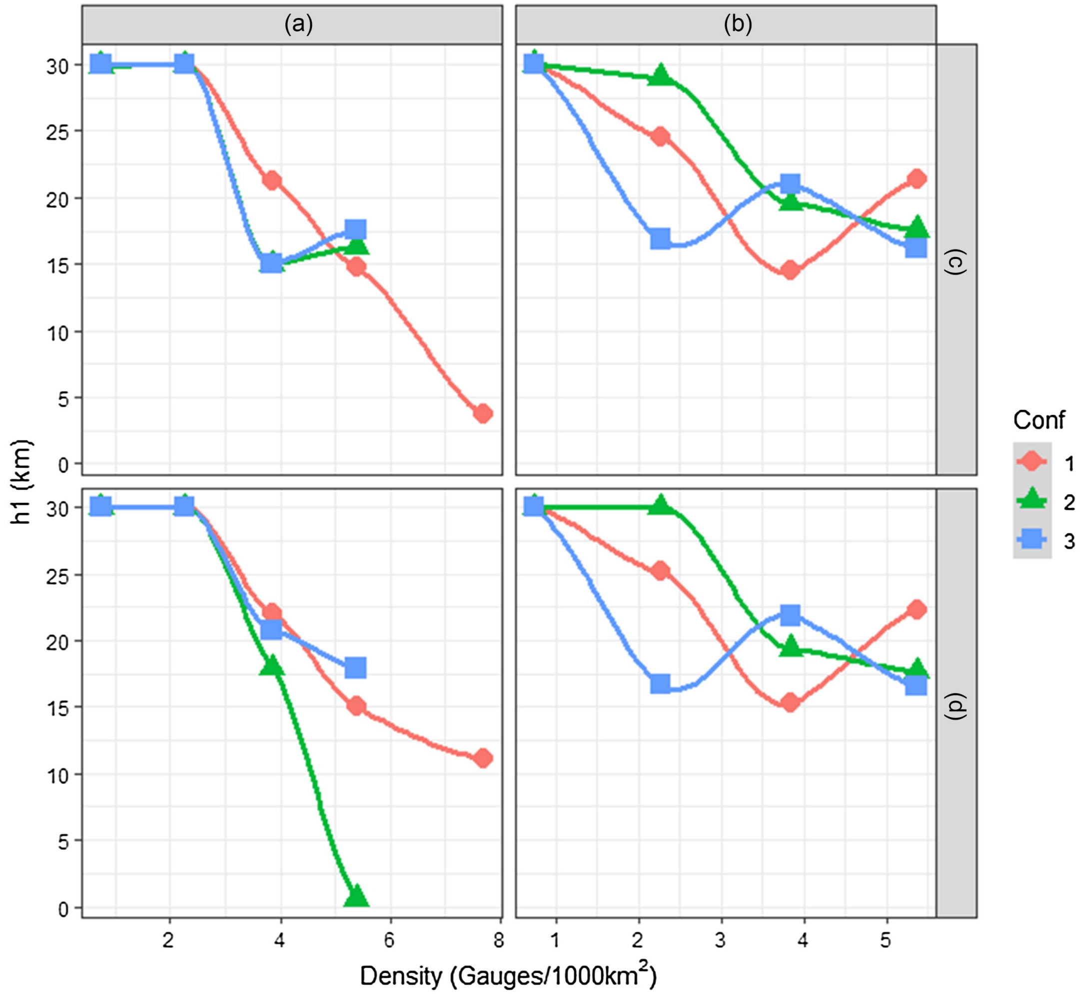

Fig. 3. Variation of the kernel bandwidth h1 with the density of rain gauges. (a) Cross-validation; (b) independent validation; (c) raw MSWEP; and (d) merging with the bilinear downscaled MSWEP.

downscaling), the IDW, and the error of the interpolation with the DSg. For the highest rain gauge density (7.69 gauges/1,000 km^2), the DS merging reported the lowest error, closely followed by K. Reg, DSg, and IDW. Using IV methodology, in all densities, the DSg and the DS merging registered the lowest error, which confirms its ability to produce an adequate rainfall field for the basin. K.Reg's performance significantly weakened when the density was reduced, and in rather low densities, the error surged, depending on the rain gauge network configuration (we trimmed the scale values shown in Fig. 5 because the network with configuration number 3 reported the highest error of 67.3 mm/day, which was far worse than the others); the main cause for this behavior is the lack of information to properly represent the spatial correlation and correctly remove the trend for fulfilling the second-order stationary assumptions. As expected, these results confirm that the kriging methods based only on gauge data should not be applied in data-scarce basins.

It must also be highlighted that DS merging also reported a lower bias compared to other methods because its strategy is to correct the rainfall field near the location of the rain gauges (see Fig. 6). In the highest density, K.Reg obtained a lower bias, which is produced by the effect of the covariates in helping the interpolation, yet PBias still confirms that K.Reg seriously suffered in low densities. Even though IDW registered a similar error to DSg and DS (Fig. 5), the reported bias was more erratic, depending on the network configuration, and its PBias was slightly higher (Fig. 6).

Evaluation of the Hydrological Performance

The evaluation of the performance of the DS merged rainfall fields in reproducing streamflow series is plotted in Fig. 7. Comparing the left-side with the right-side panels reveals that the application of bilinear downscaling did not enhance the models' results either in the boundary conditioning station or in the blind gauge. However, the rainfall field that produced the maximum performance in the forcing aggregated TOPMODEL was the one created with a slightly lower bandwidth value (40 to 45 km), compared with the field used to force the distributed wflow model (45 to 50 km). Comparing the NSE calculated for both models, it is clear that the distributed wflow model achieved better and more consistent results after a certain kernel bandwidth value (around 40–45 km and beyond). TOPMODEL also attained its best performance in the range from 40 to 45 km, yet the quality of simulations reduced with the increase of the kernel bandwidth over that value. This behavior was caused by an improper estimation of the rainfall fields produced with high bandwidths, which were inputs to the hydrological models, and in the case of the lumped forcing, it affected the bulk volume of rain in the field and consequently caused lower NSE values. The more complex structure of the distributed wflow model made a

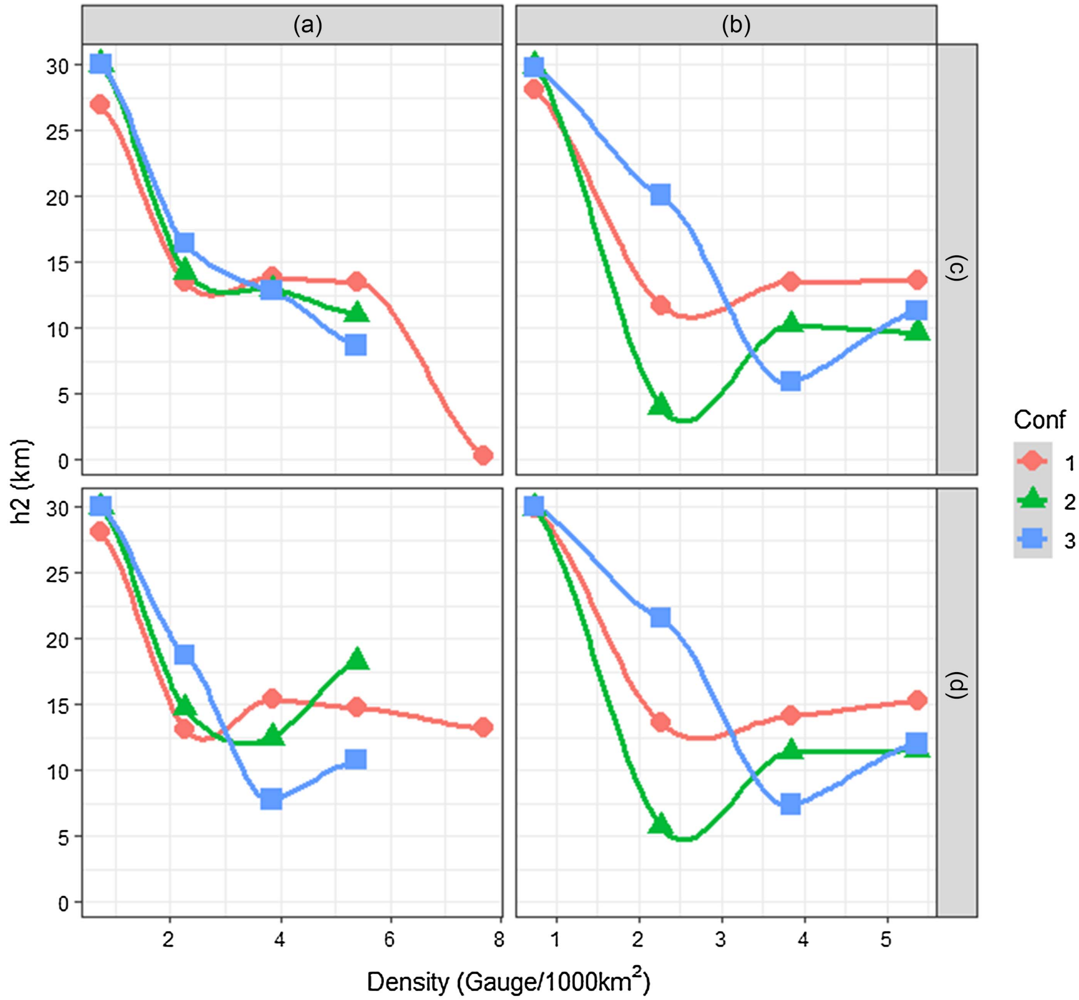

Fig. 4. Variation of the kernel bandwidth h2 with the density of rain gauges. (a) Cross-validation; (b) independent validation; (c) raw MSWEP; and (d) merging with the bilinear downscaled MSWEP.

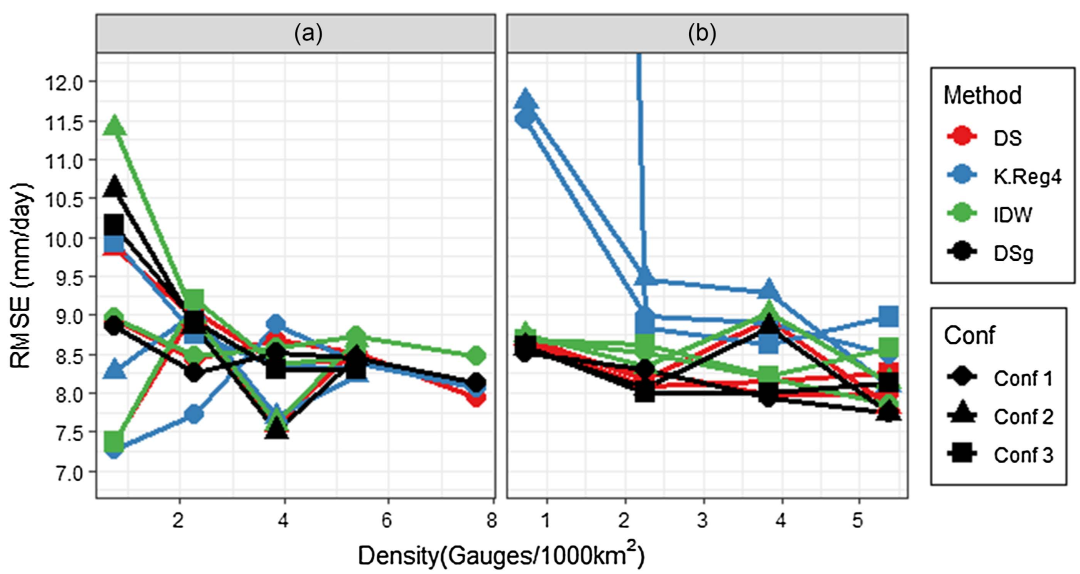

Fig. 5. RMSE of selected interpolation methods and the DS technique for several rain gauge densities. "Conf" stands for the artificial network created by randomly sampling the available gauges: (a) cross-validation; and (b) independent validation.

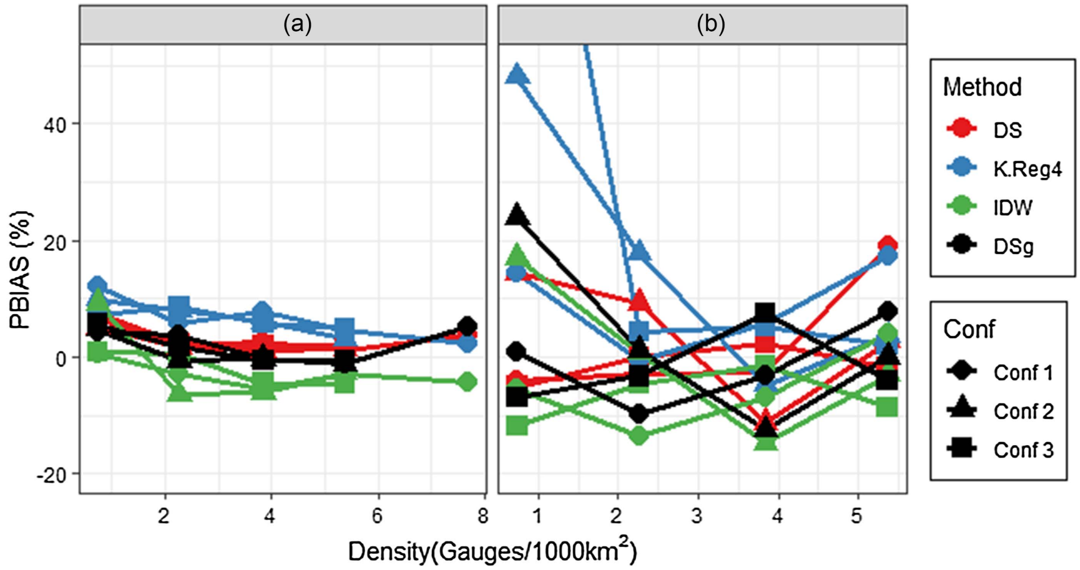

Fig. 6. PBIAS of interpolation methods and the DS technique for several rain gauge densities. "Conf" stands for the artificial network created by randomly sampling the available gauges: (a) cross-validation; and (b) independent validation.

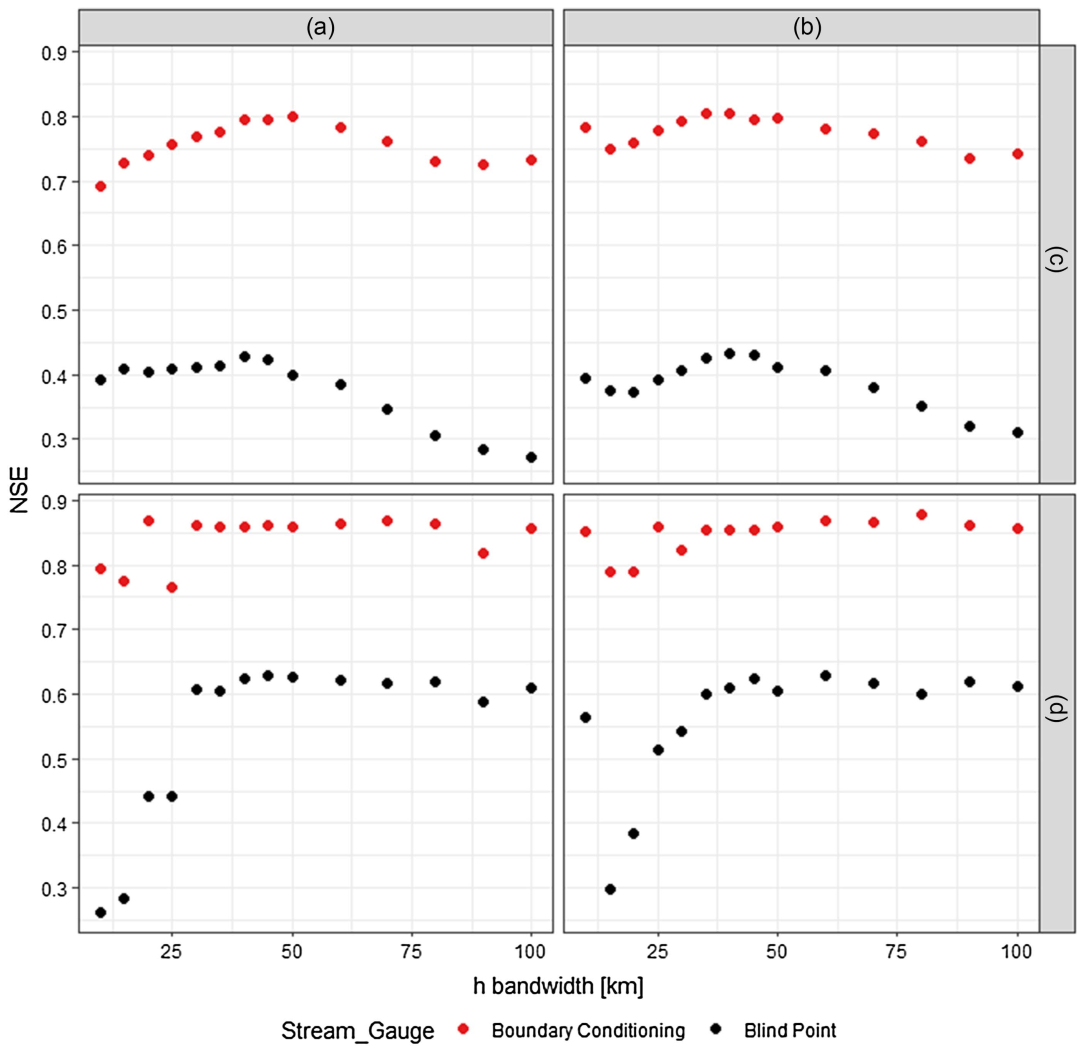

Fig. 7. Variation of the NSE for both hydrological models and for the type of downscaling: results for (a) models run with the DS merged fields without downscaling; (b) bilinear downscaling; (c) TOPMODEL; and (d) wflow. Top row of dots is the performance in the boundary conditioning gauge, and the lower row of dots refers to the blind to calibration gauge.

trade-off between the errors in the rainfall field and in the hydrological model configuration, maintaining a stable performance.

We also compared the best performance of the TOPMODEL and wflow model forced with the DS merging rain fields and fields created using some of the interpolation methods investigated (see Fig. 8). The observed and simulated hydrographs can be seen in the Supplemental Materials Fig. S2. The fields obtained through merging for different bandwidths are displayed with the prefix M (they considered $h_1 = h_2$; the names and bandwidth values used were M_30 km, M_45 km, and $M_$Silver); the MSWEP background was not downscaled with the bilinear method for the creation of these merged rainfall fields because this method did not produce a significant improvement in the results (see Fig. 8). The models forced with the fields created with the DS merging showed the best performances for both the boundary conditioning gauge and the blind point gauge, also in both models. The fields based solely on rain gauges (IDW and DSg) reported a NSE higher than the MSWEP in the calibration station, but their performance was generally lower in the blind point. This shows the robustness of the DS rainfall fields created with values of bandwidths between 30 and 50 km, which confirms what was found in section 4.1. Not surprisingly, the spatial aggregation of the rainfall field used with TOPMODEL not only caused a reduction in the representation of the streamflow but also generated rainfall fields that almost all reported the same performance in the small basin. In the blind gauge TOPMODEL, the reported NSE was nearly 0.4 for all inputs, and in the conditioning gauge, all were near 0.77 except the forcing created with Silverman's rule of thumb.

With the distributed wflow model, the difference in performance between inputs was appreciable, both in the boundary and blind points. For the blind spot, there was a large variation in NSE when comparing the rainfall fields, which is opposite that reported by the TOPMODEL. However, the IDW rainfall field was not able to produce NSE values higher than 0.3 in the blind gauge due to the low density of gauges in the Casanare basin. In the boundary conditioning, there were also discrepancies, as expected, but the performance was generally higher than those produced by TOPMODEL. Not only did the distributed model report better performance, but it also had the advantage of easily providing a higher amount of information such as simulations in ungauged locations. The semidistributed model considered the spatial variation of phenomena; however, to find estimates at ungauged points, the modeler must specify the subdivision and even build a specific model of the subcatchment in the case of TOPMODEL.

A revision of the hydrological signatures produced in the validation period by simulations at the Cravo Norte station, compared with the same signatures obtained from the observed streamflow, reveals the results shown in Table 3. Although the wflow simulations forced with the DS merging with 45-km bandwidth field reported the highest NSE (Fig. 8), the simulations associated with the 30-km bandwidth showed a better peak flow and a better reproduction of the high segment of the flow duration curve (FDC). The difference in those signatures produced with the 45-km bandwidth rainfall field was only slightly higher. Regarding the variance of the streamflow, the simulations forced with DSg interpolation performed best, closely followed by the mergings with 30-km and 45-km bandwidths.

Regarding the TOPMODEL hydrographs signatures in Table 3, the best results were achieved by merging but using the Silverman bandwidth value, with similar errors produced by the mergings with 30-km and 45-km bandwidths (in Q1% and Q2%). TOPMODEL with DSg forcing also produced the best variance of

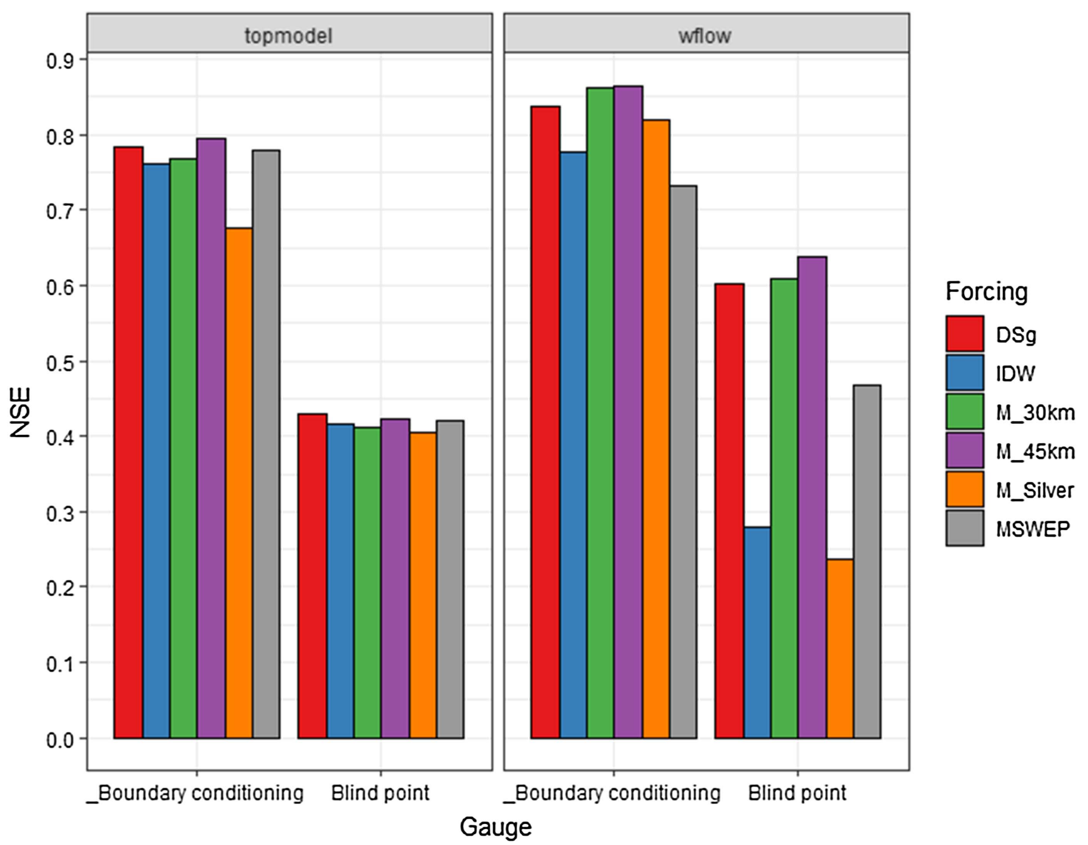

Fig. 8. Performance of the TOPMODEL and wflow model forced with interpolated, remotely sensed, and merged fields.

Table 3. Differences between observed and simulated hydrological signatures by wflow and TOPMODEL streamflow models results at Cravo Norte gauge

Hydrological signatures	Model	DSg (%)	IDW (%)	M 30 km (%)	M 45 km (%)	M silver (%)	MSWEP (%)
Q peak (Q1%)	Wflow	−13.4	−21.7	**−6.0**	−9.5	−18.3	−17.5
	TOPMODEL	6.8	6.6	4.7	6.0	**2.4**	−2.5
Q high segment (Q2%)	Wflow	−22.5	−35.5	**−13.1**	−15.4	−30.2	−30.7
	TOPMODEL	5.7	5.9	2.8	5.2	**0.8**	−6.5
Q variance	Wflow	**−26.1**	−41.9	−27.3	−29.5	−36.2	−39.7
	TOPMODEL	**−12.2**	−24.8	−23.3	−18.1	−31.8	−28.5

Note: Best values for each model and forcing are bolded.

streamflow, and surprisingly the IDW reported good results, which might be because the spatial errors of the rainfall field were averaged. Moreover, TOPMODEL produced better peak signature values with every forcing compared to wflow, which is opposite the NSE performance reported by the model (Fig. 8). The reason for this behavior is that TOPMODEL weakened the representation of the recession curves and other streamflow signatures to achieve the best possible NSE performance, a metric that focuses on the peak flow signatures; parameter overfitting might be the phenomenon leading to those results. The averaged input rainfall field also influenced that behavior and made the evaluation of the overall performance contradictory.

Discussion

In this section, we contrast the general insights from this research with similar research that sought the improvement of rainfall fields. Furthermore, the characteristics of bandwidth influence are analyzed.

In the Rainmerging master library, Zulkafli et al. (2015) fixed the value of the kernel bandwidth using the Silverman rule of thumb (Silverman 1986), and we also used other two methods (PR-N and PR-E). Through the results of this research, it is evident that the Silverman rule of thumb did not produce satisfactory results and that the bandwidth values differed significantly from those found in calibration. The values of the PR-N and PR-E methods not only resembled more those bandwidth values found by calibration in high network densities, they did not produce errors as low as those bandwidths obtained by adjustment. However, the values found by PR-N are a good seed to guide the calibration and define values, in conjunction with Figs. 3 and 4, if calibration is not possible. The reason might be that a rain gauge influence area is more associated with the density of the network and position of the gauges rather than with the simple number of gauges and variance of the distance.

Similar to the situation exposed by Long et al. (2016), a correction in the boundary of the rainfall field estimated by the background source or rain gauges does not necessarily improve the overall performance of the merging. They reported that a Boolean interpolation (indicator kriging) to include rainfall spatial intermittency does not improve the performance of the merging. Similarly, with the use of a simple bilinear downscaling, we found that there was not a clear improvement, yet from a theoretical point of view, it helped to improve the continuity and slope of the rainfall surface function (Li and Shao 2010). The reduction in error by bilinear downscaling was not appreciable in our meteorological and hydrological evaluations. More complex downscaling methods consider more meaningful covariates such as vegetation indexes (Ceccherini et al. 2015; Immerzeel et al. 2009) or topography (Jongjin et al. 2016), especially the aspect and wind direction; this has been highlighted by López López et al. (2018) and used by Long et al. (2016). However, those studies and the results reported here emphasize that mergings with gauges, like the DS and others, are methods that further improve the performance of the rainfall fields to have better correspondence with the hydrological response to rainfall.

Conclusions

This work found that variation in kernel bandwidth has an impact on both meteorological and hydrological performances when using the DS merging technique. The DS merging outperformed the traditional deterministic and geostatistical interpolation methods, in some cases by a considerable amount; in others, it performed slightly better. Although the impact of the bandwidth value can be considered to influence mainly the spatial distribution of the rainfall, we found that one significant possible outcome is an important variation in the bulk volume amount of rainfall, which, as a result, affects the hydrological simulations.

The rule-of-thumb equations for estimating the kernel bandwidth did not provide a good correction of the merged fields. The principal reason for this behavior might be that they were formulated for density estimation, whereas their use in DS correction should consider a different procedure, one more related to geostatistics. The KED and K.Reg methods consider correlations with geomorphological variables; therefore, an analysis of the spatial structure of the rainfall phenomenon is an asset when compared with the Silverman rule or the relation of the bandwidth value to geomorphological variables. Results prove that the variation of the bandwidth value affects both meteorological and hydrological evaluation, for example, through the effect in rainfall bulk volume and NSE performance, which is an asset of this research; in "Discussion," we suggested that calibration in the DS algorithm is advisable, but seeds can be previously calculated.

From the analysis of the randomized rain gauge network configurations, we can conclude that some stations are of importance due to their location. Some of the network configurations reported higher errors in comparison to the others, which is a consequence of not including those important gauges in the interpolations or the mergings. To analyze which of them are reporting essential information, the researcher can use methods based on entropy theory (Yang and Burn 1994), but this is beyond the scope of this research. Overall, the improvement of the DS can be by 20% over the results reported by some interpolation methods and 50% compared with others.

The importance of a proper estimation of rainfall field spatial variability is evidenced in the results achieved by the two hydrological models. Merging improved the performance of models by approximately 10% and in some cases by 20%. From the experiments, we can conclude that the variance of forcing performance was likely a consequence of the spatial differences in the inputs used in wflow. On the contrary, the equal but misleading behavior of TOPMODEL was probably produced by the aggregation of the inputs. Nevertheless, the results of this research must encourage meteorological and environmental agencies to improve the density of measurement in regions with network deficiencies.

Data Availability Statement

The data, code, and models that support the findings of this study are available on reasonable request from the corresponding author.

Acknowledgments

The authors would like to thank the six anonymous reviewers and the editorial board, whose comments helped to improve this work. Moreover, we would like to thank the Instituto de Hidrología, Meteorología y Estudios Ambientales -IDEAM- and the EU FP7 eartH2Observe project for providing the data for this research. ND would like to thank Joaquin Duque and Helena Gardeazábal for providing emotional and financial support when this research needed it and the fellow researchers of the GIREH group for the constructive discussions about the topic. The research reported in this paper has received funding from the European Union's Seventh Programme for Research Technological development and demonstration under Grant Agreement No. 603608. Moreover, the support from Universidad Nacional de Colombia through digital infrastructure and the financial support of Joaquin Duque & Helena Gardeazábal is gratefully acknowledged.

Supplemental Materials

Figs. S1 and S2 are available online in the ASCE Library (www.ascelibrary.org).

References

Arnaud, P., C. Bouvier, L. Cisneros, and R. Dominguez. 2002. "Influence of rainfall spatial variability on flood prediction." *J. Hydrol.* 260 (1–4): 216–230. https://doi.org/10.1016/S0022-1694(01)00611-4.

Barrett, E. 1975. "The estimation of monthly rainfall from satellite data." *Mon. Weather Rev.* 98 (4): 1109–1114.

Beck, H. E., A. I. J. M. Van Dijk, A. de Roo, E. Dutra, G. Fink, R. Orth, and J. Schellekens. 2017a. "Global evaluation of runoff from 10 state-of-the-art hydrological models." *Hydrol. Earth Syst. Sci.* 21 (6): 2881–2903. https://doi.org/10.5194/hess-21-2881-2017.

Beck, H. E., A. I. J. M. Van Dijk, V. Levizzani, J. Schellekens, D. G. Miralles, B. Martens, and A. de Roo. 2017b. "MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data." *Hydrol. Earth Syst. Sci.* 21 (1): 589–615. https://doi.org/10.5194/hess-21-589-2017.

Beck, H. E., N. Vergopolan, M. Pan, V. Levizzani, A. I. J. M. van Dijk, G. P. Weedon, L. Brocca, F. Pappenberger, G. J. Huffman, and E. F. Wood. 2017c. "Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling." *Hydrol. Earth Syst. Sci.* 21 (12): 6201–6217. https://doi.org/10.5194/hess-21-6201-2017.

Beck, H. E., N. E. Zimmermann, T. R. McVicar, N. Vergopolan, A. Berg, and E. F. Wood. 2018. "Present and future Köppen–Geiger climate classification maps at 1-km resolution." *Sci. Data* 5 (1): 180214. https://doi.org/10.1038/sdata.2018.214.

Beven, K., R. Lamb, P. Quinn, R. Romanowicz, and J. Freer. 1995. "TOPMODEL." In *Computer Models of watershed hydrology*, edited by V. P. Singh, 627–668. Highlands Ranch, CO: Water Resources Publications.

Blöschl, G., and M. Sivapalan. 1995. "Scale issues in hydrological modelling: A review." *Hydrol. Processes* 9 (3–4): 251–290. https://doi.org/10.1002/hyp.3360090305.

Buytaert, W. 2018. *TOPMODEL: Implementation of the hydrological model TOPMODEL in R*. CRAN.r-project: London.

Ceccherini, G., I. Ameztoy, C. Hernández, and C. Moreno. 2015. "High-resolution precipitation datasets in South America and West Africa based on satellite-derived rainfall, enhanced vegetation index and digital elevation model." *Remote Sens.* 7 (5): 6454–6488. https://doi.org/10.3390/rs70506454.

Chappell, A., L. Renzullo, and M. Haylock. 2012. "Spatial uncertainty to determine reliable daily precipitation maps." *J. Geophys. Res. Atmos.* 117 (Jan): 1–14. https://doi.org/10.1029/2012JD017718.

Chow, V., D. R. Maidment, and L. Mays. 1988. *Applied hydrology.* Singapore: McGraw-Hill Book.

Cressie, N. A. C. 1994. *Statistics for spatial data biometrics*. New York: Wiley. https://doi.org/10.2307/2533238.

Dee, D. P., et al. 2011. "The ERA–interim reanalysis: Configuration and performance of the data assimilation system." *Q. J. R. Meteorol. Soc.* 137 (656): 553–597. https://doi.org/10.1002/qj.828.

Dinku, T., F. Ruiz, S. J. Connor, and P. Ceccato. 2010. "Validation and intercomparison of satellite rainfall estimates over Colombia." *J. Appl. Meteorol. Climatol.* 49 (5): 1004–1014. https://doi.org/10.1175/2009JAMC2260.1.

Duan, Q., S. Sorooshian, and H. V. Gupta. 1992. "Effective and efficient global optimization for conceptual rainfall-runoff models." *Water Resour. Res.* 28 (4): 1015–1031. https://doi.org/10.1029/91WR02985.

Duncan, M. R., B. Austin, F. Fabry, and G. L. Austin. 1993. "The effect of gauge sampling density on the accuracy of streamflow prediction for rural catchments." *J. Hydrol.* 142 (1–4): 445–476. https://doi.org/10.1016/0022-1694(93)90023-3.

Duque-Gardeazábal, N., D. Zamora, and E. Rodríguez. 2018. "Analysis of the kernel bandwidth influence in the double smoothing merging algorithm to improve rainfall fields in poorly gauged basins." In *Proc., 13th Int. Conf. on Hydroinformatics*, 635–642. London: International Water Association. https://doi.org/10.29007/2xp6.

Ehret, U., J. Götzinger, A. Bárdossy, and G. G. S. Pegram. 2008. "Radar-based flood forecasting in small catchments, exemplified by the Goldersbach catchment, Germany." *Int. J. River Basin Manage.* 6 (4): 323–329. https://doi.org/10.1080/15715124.2008.9635359.

Espinoza Villar, J. C., J. Ronchail, J. L. Guyot, G. Cochonneau, F. Naziano, W. Lavado, E. De Oliveira, R. Pombosa, and P. Vauchel. 2009. "Spatio-temporal rainfall variability in the Amazon basin countries (Brazil, Peru, Bolivia, Colombia, and Ecuador)." *Int. J. Climatol.* 29 (11): 1574–1594. https://doi.org/10.1002/joc.1791.

García, L. E., D. J. Rodríguez, M. Wijnen, and I. Pakulski. 2016. *Earth observations for water resources management: Current use and future opportunities for the water sector.* Washington, DC: The World Bank.

Grimes, D. I. F., and E. Pardo-Iguzquiza. 2010. "Geostatistical analysis of rainfall." *Geogr. Anal.* 42 (2): 136–160. https://doi.org/10.1111/j.1538-4632.2010.00787.x.

Gupta, H. V., S. Sorooshian, and P. O. Yapo. 1999. "Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration." *J. Hydrol. Eng.* 4 (2): 135–143. https://doi.org/10.1061/(ASCE)1084-0699(1999)4:2(135).

Heidinger, H., et al. 2012. "TRMM rainfall correction over the Andean Plateau using wavelet multi- resolution analysis." *Int. J. Remote Sens.* 33 (14): 4583–4602. https://doi.org/10.1080/01431161.2011.652315.

Hurtado-Montoya, A. F., and Ó. J. Mesa-Sánchez. 2014. "Reanalysis of monthly precipitation fields in Colombian territory." *Dyna* 81 (186): 251. https://doi.org/10.15446/dyna.v81n186.40419.

IDEAM. 2019. *Estudio Nacional del Agua 2018*. Ideam, Bogotá: IDEAM.

IDEAM. 2022. "Consulta y Descarga de Datos Hidrometeorológicos [WWW Document]." Accessed November 14, 2020. http://dhime.ideam.gov.co/atencionciudadano/.

Immerzeel, W. W., M. M. Rutten, and P. Droogers. 2009. "Spatial downscaling of TRMM precipitation using vegetative response on the Iberian Peninsula." *Remote Sens. Environ.* 113 (2): 362–370. https://doi.org/10.1016/j.rse.2008.10.004.

Jongjin, B., P. Jongmin, R. Dongryeol, and C. Minha. 2016. "Geospatial blending to improve spatial mapping of precipitation with high spatial resolution by merging satellite-based and ground-based data." *Hydrol. Processes* 30 (16): 2789–2803. https://doi.org/10.1002/hyp.10786.

Lall, U., Y.-I. Moon, and K. Bosworth. 1993. "Kernel flood frequency estimators: Bandwidth selection and kernel choice." *Water Resour. Res.* 29 (4): 1003–1015. https://doi.org/10.1029/92WR02466.

Li, M., and Q. Shao. 2010. "An improved statistical approach to merge satellite rainfall estimates and raingauge data." *J. Hydrol.* 385 (1–4): 51–64. https://doi.org/10.1016/j.jhydrol.2010.01.023.

Lindström, G., B. Johansson, M. Persson, M. Gardelin, and S. Bergström. 1997. "Development and test of the distributed HBV-96 hydrological model." *J. Hydrol.* 201 (1–4): 272–288. https://doi.org/10.1016/S0022-1694(97)00041-3.

Long, Y., Y. Zhang, and Q. Ma. 2016. "A merging framework for rainfall estimation at high spatiotemporal resolution for distributed hydrological modeling in a data-scarce area." *Remote Sens.* 8 (7): 599. https://doi.org/10.3390/rs8070599.

López López, P., W. W. Immerzeel, E. A. Rodríguez, G. Sterk, and J. Schellekens. 2018. "Spatial downscaling of satellite-based precipitation and its impact on discharge simulations in the Magdalena River basin in Colombia: Impact of high spatial resolution precipitation on streamflow simulations." *Front. Earth Sci.* 6 (Jun): 68. https://doi.org/10.3389/feart.2018.00068.

Ly, S., C. Charles, and A. Degré. 2011. "Geostatistical interpolation of daily rainfall at catchment scale: The use of several variogram models in the Ourthe and Ambleve catchments, Belgium. Hydrol." *Earth Syst. Sci.* 15 (7): 2259–2274. https://doi.org/10.5194/hess-15-2259-2011.

Muñoz, E., P. Tume, and G. Ortíz. 2015. "Uncertainty in rainfall input data in a conceptual water balance model: Effects on outputs and implications for predictability." *Earth Sci. Res. J.* 18 (1): 69–75. https://doi.org/10.15446/esrj.v18n1.38760.

Nash, J. E., and J. V. Sutcliffe. 1970. "River flow forecasting through conceptual models part I—A discussion of principles." *J. Hydrol.* 10 (3): 282–290. https://doi.org/10.1016/0022-1694(70)90255-6.

Nerini, D., Z. Zulkafli, L.-P. Wang, C. Onof, W. Buytaert, W. Lavado-Casimiro, and J.-L. Guyot. 2015. "A comparative analysis of TRMM–rain gauge data merging techniques at the daily time scale for distributed rainfall–runoff modeling applications." *J. Hydrometeorol.* 16 (5): 2153–2168. https://doi.org/10.1175/JHM-D-14-0197.1.

Park, N.-W. 2013. "Spatial downscaling of TRMM precipitation using geostatistics and fine scale environmental variables." In Vol. 2013 of *Advances in meteorology*, 1–9. https://doi.org/10.1155/2013/237126.

Poveda, G., P. R. Waylen, and R. S. Pulwarty. 2006. "Annual and inter-annual variability of the present climate in northern South America and southern Mesoamerica." *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 234 (1): 3–27. https://doi.org/10.1016/j.palaeo.2005.10.031.

Rajagopalan, B., U. Lall, and D. G. Tarboton. 1997. "Evaluation of kernel density estimation methods for daily precipitation resampling." *Stochastic Hydrol. Hydraul.* 11 (6): 523–547. https://doi.org/10.1007/BF02428432.

Refsgaard, J. C. 1997. "Parameterisation, calibration and validation of distributed hydrological models." *J. Hydrol.* 198 (1–4): 69–97. https://doi.org/10.1016/S0022-1694(96)03329-X.

Rodríguez, E., et al. 2020. "Combined use of local and global hydro meteorological data with hydrological models for water resources management in the Magdalena–Cauca Macro Basin–Colombia." *Water Resour. Manage.* 34 (7): 2179–2199. https://doi.org/10.1007/s11269-019-02236-5.

Rogelis, M. C., and M. G. F. Werner. 2013. "Spatial interpolation for real-time rainfall field estimation in areas with complex topography." *J. Hydrometeorol.* 14 (1): 85–104. https://doi.org/10.1175/JHM-D-11-0150.1.

Schellekens, J., et al. 2017. "A global water resources ensemble of hydrological models: The eartH2Observe Tier-1 dataset." *Earth Syst. Sci. Data* 9 (2): 389–413. https://doi.org/10.5194/essd-9-389-2017.

Schellekens, J. 2020. "OpenStreams wflow documentation release 1.0RC1 [WWW Document]." Accessed April 26, 2017. http://wflow.readthedocs.org/en/latest/.

Serrat-Capdevila, A., J. B. Valdes, and E. Z. Stakhiv. 2014. "Water management applications for satellite precipitation products: Synthesis and recommendations." *J. Am. Water Resour. Assoc.* 50 (2): 509–525. https://doi.org/10.1111/jawr.12140.

Silverman, B. W. 1986. *Density estimation for statistics and data analysis*. London: CRC Press.

Sperna Weiland, F. C., P. Lopez, A. van Dijk, and J. Schellekens. 2015. "Global high-resolution reference potential evaporation." In *Proc., 21st Int. Congress on Modelling and Simulation*, 2548–2554. Canberra, ACT, Australia: Modelling and Simulation Society of Australia and New Zealand.

Todini, E. 2001. "A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements." *Hydrol. Earth Syst. Sci.* 5 (2): 187–199. https://doi.org/10.5194/hess-5-187-2001.

Tolson, B. A., and C. A. Shoemaker. 2007. "Dynamically dimensioned search algorithm for computationally efficient watershed model calibration." *Water Resour. Res.* 43 (1): 1–16. https://doi.org/10.1029/2005WR004723.

Vila, D. A., L. G. G. de Goncalves, D. L. Toll, and J. R. Rozante. 2009. "Statistical evaluation of combined daily gauge observations and rainfall satellite estimates over continental South America." *J. Hydrometeorol.* 10 (2): 533–543. https://doi.org/10.1175/2008JHM1048.1.

Wagner, P. D., P. Fiener, F. Wilken, S. Kumar, and K. Schneider. 2012. "Comparison and evaluation of spatial interpolation schemes for daily rainfall in data scarce regions." *J. Hydrol.* 464–465 (Sep): 388–400. https://doi.org/10.1016/j.jhydrol.2012.07.026.

Webster, R., and M. A. Oliver. 2007. *Geostatistics for environmental scientists, statistics in practice*. Chichester, UK: Wiley. https://doi.org/10.1002/9780470517277.

Weedon, G. P., S. Gomes, P. Viterbo, W. J. Shuttleworth, E. Blyth, H. Österle, J. C. Adam, N. Bellouin, O. Boucher, and M. Best. 2011. "Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century." *J. Hydrometeorol.* 12 (5): 823–848. https://doi.org/10.1175/2011JHM1369.1.

WMO (World Meteorological Organization). 2008. *Guide to hydrological practices*. WMO 168, Ginebra, Suiza: WMO.

Woldemeskel, F. M., B. Sivakumar, and A. Sharma. 2013. "Merging gauge and satellite rainfall with specification of associated uncertainty across Australia." *J. Hydrol.* 499 (Aug): 167–176. https://doi.org/10.1016/j.jhydrol.2013.06.039.

Yang, Y. J., and D. H. Burn. 1994. "An entropy approach to data-collection network design." *J. Hydrol.* 157 (1–4): 307–324. https://doi.org/10.1016/0022-1694(94)90111-2.

Zulkafli, Z., W. Buytaert, C. Onof, B. Manz, E. Tarnavsky, W. Lavado, and J.-L. Guyot. 2014. "A comparative performance analysis of TRMM 3B42 (TMPA) versions 6 and 7 for hydrological applications over Andean–Amazon river basins." *J. Hydrometeorol.* 15 (2): 581–592. https://doi.org/10.1175/JHM-D-13-094.1.

Zulkafli, Z., D. Nerini, and B. Manz. 2015. *Rainmerging: Pixel-point merging functions for precipitation time series (R package)*. San Francisco: GitHub. https://doi.org/10.5281/zenodo.18869.