# scientific reports

Check for updates

OPEN

# Latent negative precipitation for the delineation of a zero-precipitation area in spatial interpolations

Taesam Lee[1] & Ju-young Shin[2]✉

The spatial interpolation of precipitation has been employed in a number of fields, including by spatially downscaling the Global Circulation Model (GCM) to a finer scale. Most precipitation events become more sporadic when the coverage area increases (i.e., a portion of the points experience zero precipitation). However, spatial interpolations of precipitation generally ignore these dry areas, and the interpolated grids are filled with certain precipitation amounts. Subsequently, no delineation of dry and wet regions can be made. Therefore, the current study suggested a novel approach to determine dry areas in spatial interpolations of precipitation events by assigning latent negative precipitation (LNP) to points with observed precipitation values of zero. The LNP-assigned points are then employed in a spatial interpolation. After that, the dry region can be determined using the negative region (i.e., points with zero precipitation). The magnitude of LNP can be defined by multiplying the precipitation values of neighboring stations by a tuning parameter. The LNP method and the tuning parameter are tested on weather stations covering South Korea. The results indicate that the proposed LNP method can be suitable for the spatial interpolation of precipitation events by delineating dry and wet regions. Additionally, the tuning parameter plays a special role in that it increases in value with longer precipitation durations and denser networks. A value of 0.5–1.5 can be suggested for the tuning parameter as a rule of thumb when high accuracy for final products of interpolated precipitation is not critical. For future studies, the LNP model derived herein can be tested over much larger areas, such as the United States, and the model can also be easily adopted for other variables with spatially sporadic values.

The spatial interpolation of precipitation plays a critical role when mapping sparse point precipitation data from weather stations to obtain input data for distributed hydrologic models such as Vflo[1,2], evaluating weather prediction models[3], and assessing the impacts of climate change on local watersheds through spatial downscaling[4–7]. A number of spatial interpolation schemes have been proposed in the literature[8], such as Thiessen polygons[9], inverse distance weighting (IDW)[10], linear (and nonlinear) spatial interpolations[11], and geostatistical kriging[12,13].

In recent years, these spatial interpolation techniques have been popularly used to statistically downscale hydrometeorological variables when conducting climate change assessments[14–20]. A number of statistical downscaling methods have been used to simulate precipitation data for hydrologic and agricultural model inputs[18,20–22], such as bias correction spatial downscaling[23,24], constructed analogs[6], and conditional simulations[25].

In implementing a spatial interpolation of precipitation data, a portion of the observed points (or grids) often have observed precipitation values of zero. All target interpolating grids are filled with a certain amount of precipitation in most spatial interpolation methods, such as IDW, linear spatial interpolation, and kriging, unless all the considered stations have precipitation values of zero. The interpolated precipitation thus results in a high frequency of low precipitation values. These results might not be favorable for applications in water resource management and distributed hydrologic models since they reduce dry regions and increase outflows.

[1]Department of Civil Engineering, ERI, Gyeongsang National University, 501 Jinju-daero, Jinju, Gyeongnam 660-701, South Korea. [2]High-Impact Research Department, National Institute of Meteorological Sciences, Jeju, South Korea. ✉email: jyshin83@korea.kr

Therefore, a novel approach for spatially interpolating precipitation regarding the determination of wet and dry regions is proposed in the current study. The applicability of the proposed model was tested using weather stations in South Korea.

## Mathematical Background

**Spatial Interpolation.** IDW is one of the most common spatial interpolation approaches and involves assigning weights to the observed values of neighboring stations according to the distances between the neighboring stations and the target point and taking weighted averages. The precipitation, $\hat{P}$, for a target point or grid can be calculated with IDW, when $P_s$, $s = 1,\ldots,S$, where $S$ is the number of neighboring stations, as follows:

$$\hat{P} = \frac{\sum_{s=1}^{S} P_s w_s}{\sum_{s=1}^{S} w_s} \tag{1}$$

where the weights are $w_s = 1/d_s^2$ and $d_s$ is the distance between the target point and station $s$. IDW is a simple and efficient interpolation method[26], and more complex methods such as multiple linear regressions and kriging interpolation can perform better but require sufficient data densities[27].

The simplest approach for the spatial interpolation of rainfall is to assign the same value to the target point as that of the closest station, called the nearest-neighbor (NN) method, such that the following expression is true.

$$\hat{P} = \min(d_s, s \in \{1, \ldots, S\})P_s \tag{2}$$

Note that if precipitation is interpolated with this NN method, especially in gridded cases, Thiessen polygons can be constructed[28].

**Proposed model description.** To spatially interpolate precipitation, including precipitation in wet regions (i.e., precipitation > 0) and dry regions (i.e., precipitation = 0), an appropriate method must be proposed. To the best of the authors' knowledge, no appropriate method exists. In the current study, the latent negative precipitation (LNP) method is proposed. First, a negative precipitation amount is assigned before interpolation to stations with precipitation values of zero with the same magnitude as that of a precipitated neighbor station. This negative value is latent since it is not measured, but it indicates how much driving force is required to make a neighboring station dry.

The physical justification is the following. When a station has a high precipitation value, often accompanied by low pressure, a neighboring station that is dry may be affected by strong high pressure based on meso and macro scales. The detailed extensive description of the LNP method is explained with the diagram in Fig. 1.

A case with two weather stations is shown in Fig. 1. Station $S_1$ on the left side has a $P$ amount of rain, while station $S_2$ on the right side has no precipitation. For the spatial interpolation of precipitation with two stations separated by a distance of $D$, the interpolation values between the two stations can be linearly interpolated, as shown by the solid blue line in the top panel (a), or interpolated with IDW, as shown in the bottom panel (b). The interpolation results have no zero-precipitation areas except at station $S_2$. This drawback can be problematic when interpolated data are used for drought analyses, water resource management, or flood analyses. In recent decades, the spatial downscaling of global climate model (GCM) outputs has been popularly employed with spatial interpolation techniques[4,5] requiring adequate delineations of zero-precipitation areas.

Therefore, a special remedy that allows spatial interpolations with the delineation of zero-precipitation areas is proposed in the current study. The dry severity of these regions is as high as the precipitation amount of the neighboring weather stations since the wet weather system is blocked by the dry system and a severely wet system must be blocked by a severely dry system to have an area with zero rainfall. Therefore, it is suggested that a dry weather system can be represented with LNP for stations with zero precipitation.

The procedure can be summarized as follows: (1) LNP is assigned to the stations with zero precipitation; (2) the target area is spatially interpolated with LNP values and other nonzero precipitation values; and (3) the negative-value area is set as a zero-rainfall area. In this procedure, the dry region is defined according to the strength of LNP. For example, when LNP is equal to -2P, two-thirds of the area (presented with $D$) has a value of zero in the linear interpolation, as shown in the top panel (a) of Fig. 1, and less than two-thirds of area (D) has a value of zero in IDW, as shown in the bottom panel (b); the situation is reversed in the case of -1/2P.

The severity of dryness represented with LNP can differ with each storm event and area. Even though the LNP severity of each storm event cannot be modeled, the overall regional effect can be considered using a tuning parameter (i.e., represented with -1/2 ~ -2 in Fig. 1). This parameterization is discussed following a detailed description of spatial interpolation using LNP.

Assuming that $M$ weather stations have precipitation values of zero among $S_i$, ($i = 1,\ldots,S$ where $S$ is the number of neighboring stations), the weather stations with zero precipitation are described as $S_m^*$ ($m = 1,\ldots,M$). The objective is to spatially interpolate a certain gridded area (say $N_L = N_C \times N_R$ where $N_C$ and $N_R$ are the numbers of columns and rows, respectively) containing the considered weather stations $S_i$, ($i = 1,\ldots,S$) with some zeros. The detailed procedure is as follows for weather stations $S_i$, $i = 1,\ldots, S$.

(1) Calculate the distances between stations and indicate the stations that have zero precipitation. Here, assume M stations have zero precipitation.
(2) Estimate LNP for the stations ($S_m^*$; $m = 1,\ldots,M$) with zero precipitation using spatial interpolation techniques such as NN or IDW.

2
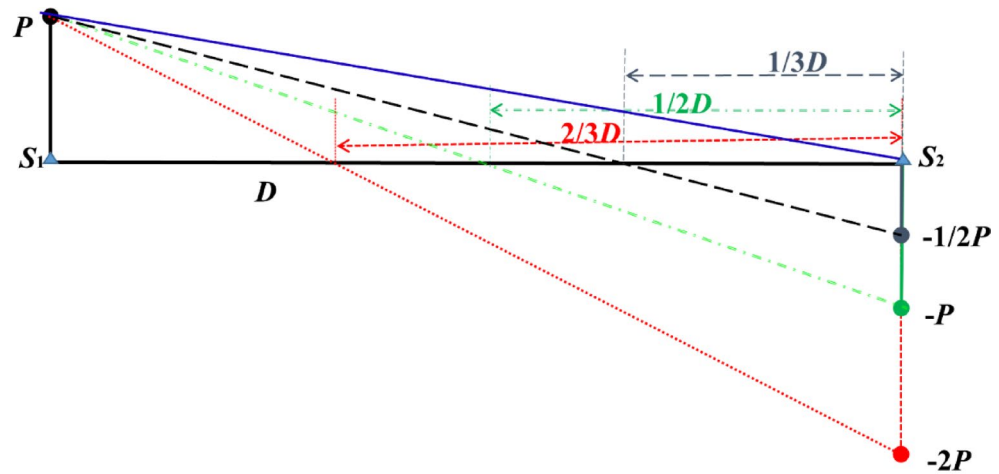
(a) Schematic illustration of the proposed NLP
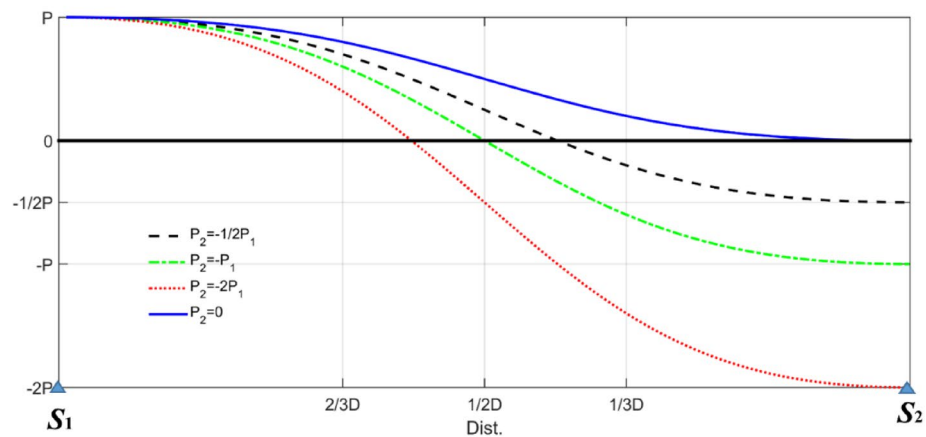


(b) Precip. Estimate with IDW

**Figure 1.** Schematic illustration of negative latent precipitation (**a**) and spatial interpolation performed with IDW for LNP (**b**). Note that the left side point ($S_1$) indicates a weather station that records a precipitation amount of $P$, while the weather station on the right side ($S_2$) records no precipitation. Three different negative values, $-2P$, $-P$, and $1/2P$, are assigned instead of a value of zero as latent precipitation. For example, when $-2P$ is assigned, the $2/3D$ area among $D$ is eventually a zero-precipitation area, as shown in the top panel (**a**). The precipitation estimates obtained with IDW in the bottom panel (**b**) were calculated using the simple equation $P_{IDW} = (a^2 + b)/(a^2 + 1)P$, where $a = D_2/D_1$ and $b = P_2/P_1$. Note that $a$ and $b$ are the ratios of the distance and precipitation, respectively, for site 1 and site 2.

$$\hat{P}_m^* = \frac{\sum_{j=1}^{S-M} P_j\left(1/d_j^2\right)}{\sum_{j=1}^{S-M} 1/d_j^2} \quad \text{for IDW} \tag{3}$$

$$\hat{P}_m^* = \min\left(d_j, j \in \{1, ..., S-M\}\right)P_j \quad \text{for NN} \tag{4}$$

$$LNP_m = -\hat{P}_m^* \tag{5}$$

(3) Spatially interpolate the target gridded area with the observed precipitation values, $P_j$ ($j = 1,...,S\text{-}M$), and the estimated $LNP_m$ ($m = 1,...,M$). Let the interpolated precipitation values be represented as $\tilde{P}_l$ ($l = 1,..., N_l$; the number of grids).

(4) Set the negative values as zero for the spatially interpolated grids, as shown below.

$$\hat{P}_l = \begin{cases} \tilde{P}_l & \text{if } \tilde{P}_l > 0 \\ 0 & \text{otherwise} \end{cases} \qquad (6)$$

In Fig. 1, weather station $S_1$ on the left side has $P$ amount of rain, while $S_2$ has dry conditions. By assigning the negative value of $S_1$ as –P to the dry site ($S_2$), half of the distance (1/2$D$) is negative by linear interpolation (the dash-dotted green line in Fig. 1). By setting the negative area to zero, this area becomes a nonprecipitated area. Note that the NN method results in the same 1/2$D$ area.

However, the dry area can differ when assigning different LNP values. For example, an LNP value of −2$P$ enforces the zero-precipitation area as 2/3$D$ (see the dotted red line in Fig. 1), while a value of -1/2$P$ enforces the zero-precipitation area as 1/3$D$ (the dashed black line in Fig. 1). Even though an appropriate LNP value length can differ among each precipitation event, overall adjustment can be made by parameterizing the proportion as follows:

$$LNP_m = -\lambda \hat{P}_m^* \qquad (7)$$

where $\lambda$ plays the role of a tuning parameter.

*Estimation of the tuning parameter.*    To estimate the tuning parameter $\lambda$, K-fold cross validation (KFCV) was employed. With the presumed tuning parameter $\lambda$, KFCV can be performed by the following procedure: (1) dividing the dataset into $k$ subsets, (2) fitting a model with the data excluding each subset, (3) predicting the excluded subset and calculating the cross-validation errors, and (4) repeating this $K$ times. The cross-validation error is calculated as follows:

$$E_K(\lambda) = \sum_{k=1}^{K} \sum_{i=1}^{n_k} \left\{ P_i - \hat{P}_i^{-k}(\lambda) \right\}^2 \qquad (8)$$

where $\hat{P}_i^{-k}(\lambda)$ is the predicted precipitation with the LNP interpolation method and the data excluding each $k$th subset and $n_k$ is the number of data points in the $k$th subset. The tuning parameter $\lambda$ is taken as the minimum value of $E_K(\lambda)$. Note that the fitting model includes spatial interpolation models such as NN and IDW, including the LNP for zero-precipitation areas as suggested. Also, root-mean-square-error (RMSE) was estimated to show the performance of the proposed LNP model and is denoted as:

$$RMSE = \sqrt{1/n \sum_{i=1}^{n} \left\{ P_i - \hat{P}_i \right\}^2} \qquad (9)$$

where $\hat{P}_i$ is the predicted precipitation.

In this study, ten iterations were adopted for KFCV. The first to ninth iterations had seven stations, and the tenth iteration had ten stations. Eight durations, 1, 2, 3, 4, 6, 8, 12, and 24 h, were employed for the precipitation event duration. To investigate the impacts of the spatial coverage of a precipitation event on the interpolation methods, various numbers of wet stations, such as 10, 20, 30, 40, 50, and 60, were selected, for which the precipitation depths were higher than 0 mm. For the case of ten wet stations, any precipitation event with more than ten wet stations during the event was defined as a precipitation event. In addition, precipitation events were omitted in the cross-validation if the total depth of precipitation for all employed stations was lower than 10 mm to exclude very light precipitation events and to attenuate errors in the observations.

**Study area.**    To validate the suggested model, 73 stations over South Korea were applied; these stations are run by the Korea Meteorological Agency (KMA) and are shown in Fig. 2a and Table S1 of the Supplementary Material for their latitude and longitude. South Korea illustrates serious rainfall deviation, as most of the annual rainfall falls during the rainy season, leading to vulnerability to floods. Extreme rainfall events also often occur from tropical cyclones or severe thunderstorms, and events are reportedly increasing[29]. Assessing future extreme rainfall events is critical, and the spatial downscaling of GCM outputs is crucial since the resulting areas affected by rainfall events are too small to obtain by directly applying GCM outputs.

Figure 2b presents the number of studied precipitation events along with the number of precipitated stations. The frequency indicates that a large number of partial precipitation events occur. For example, precipitation events with only one station occurred approximately 750 times over the study period, while 10–50 stations experience approximately 100 precipitation events each. Additionally, events in which all the stations recorded precipitation occurred approximately 700 times.

## Results

Figure 3 illustrates one of the precipitation events that occurred over South Korea by interpolating the country with IDW and enhancing the LNP method to indicate dry regions. The top left panel of the figure presents the constant altitude plan position indicator (CAPPI) of the radar product provided by KMA for August 8, 2015, at 17:40 (downloaded from the site: https://www.kma.go.kr/eng/weather/images/radar.jsp), and the other panels of the figure illustrate the interpolated precipitation obtained using the rainfall that occurred during 17:00–18:00 on August 8, 2015 as measured by the 73 weather stations with different lambda values. Note that (1) CAPPI products do not directly indicate the amount of precipitation because they measure the amount of water vapor

**(a)**



**(b)**
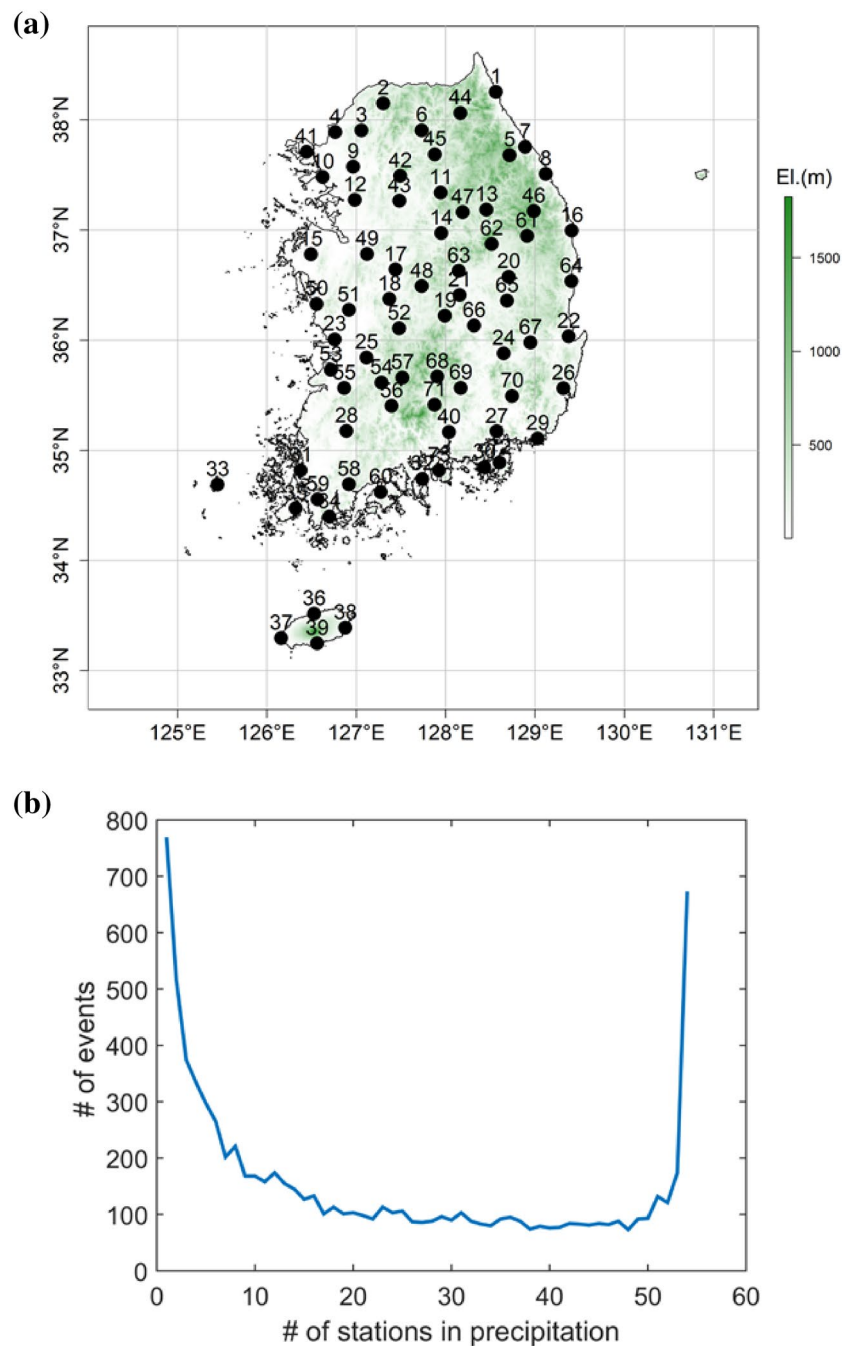


**Figure 2.** (**a**) Locations of weather stations employed and (**b**) number of events corresponding to the number of stations with precipitation. For example, the number of events in which just one station recorded precipitation is approximately 750 among 73 stations in South Korea.

present before it falls to the ground, and the time interval is different from that of the right side of the figure; (2) the estimated optimal $\lambda$ for this case was 0.51 that are close to the one of the left and second-row panel of the figure. The spatial rainfall distributions of the radar and the one close to the optimal one (the left and second-row panel) coincides well with each other. The dry area over South Korea is also represented well with the LNP model. The top right panel presents the case when $\lambda = 0$, and no dry area can be shown since the LNP method was not adopted. In contrast, the dry area increases as $\lambda$ increases as shown in the second and third-row panels. This indicates that $\lambda$ must be adequately estimated to be used effectively and illustrates that the proposed LNP model can be acceptable for use in spatial interpolations, especially regarding the delineation of dry regions.

In Fig. 4, the optimal $\lambda$ value was found by assigning LNP amounts obtained with IDW and NN, as in Eq. (3) and Eq. (4). Both cases show a similar optimal value of $\lambda = 0.5$. However, IDW shows a lower RMSE than NN, indicating that assigning LNP using the weighted averages of neighboring stations according to inverse distances
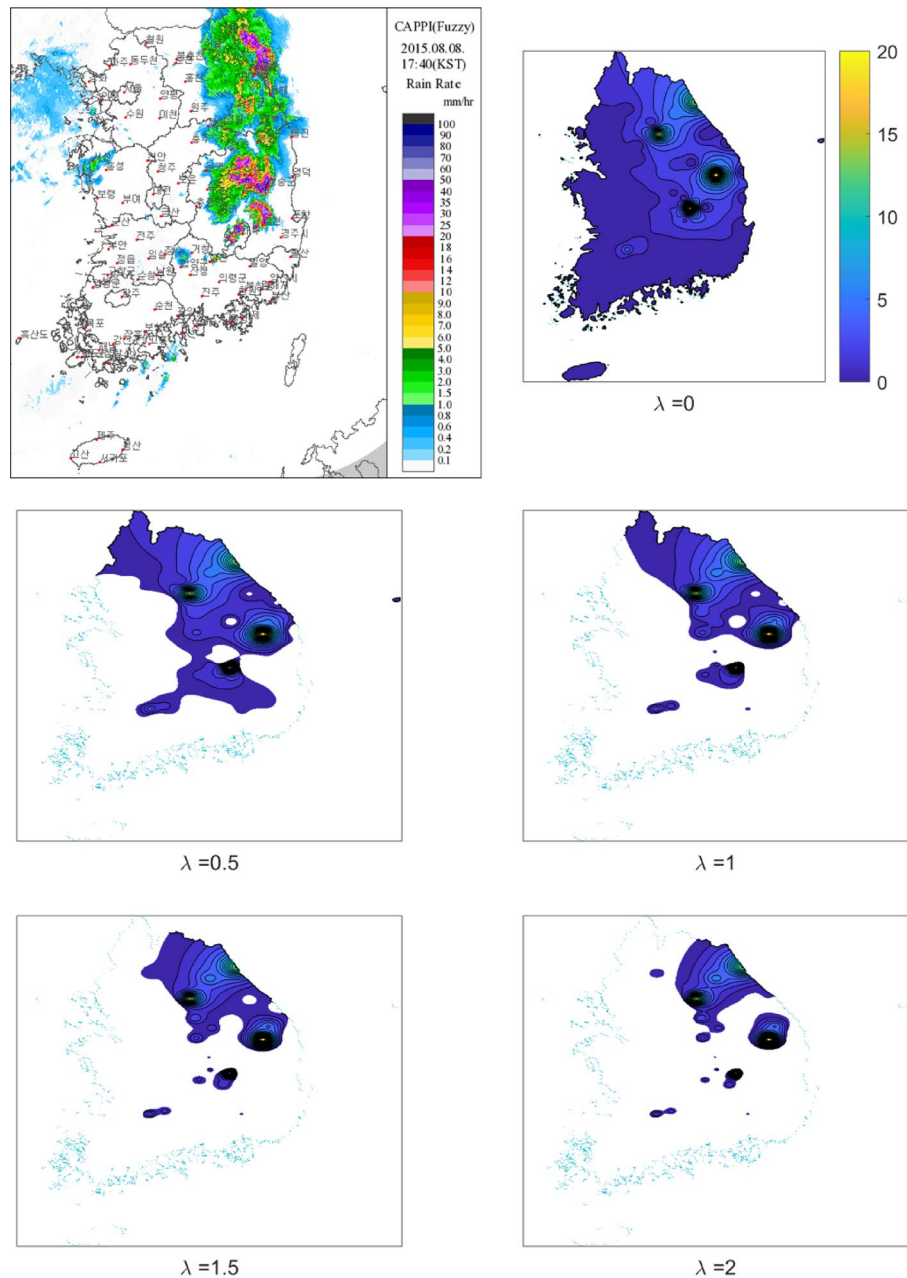
**Figure 3.** Example of the radar image for the precipitation (mm/hr) on August 8, 2015 at 17:40 (top left panel) and the proposed spatial interpolation LNP and IDW technique with different lambdas. Note that the estimated optimal value of lambda is 0.51 for this case. The figures were created by MATLAB 2021a (https://www.mathw orks.com/products/matlab.html).

(i.e., IDW) might be better than using nearest-neighbor estimations (see Table 1). By estimating the lambda value, a precipitation event can be interpolated by separating the wet and dry regions, as shown in Fig. 3.

Further experiment and test were made to provide the statistical test especially for comparing the interpolation with and without the LNP model. In the experiment, 60 stations were employed to interpolate while the other 13 stations were assumed not to have observed precipitation data. The precipitation at 13 stations was interpolated using IDW with and without the LNP model. The interpolated precipitation data are used for the Kolmogorov–Smirnov test (KS test) with observed precipitation data. KS test is a nonparametric test that can be used to compare two samples or a sample with a reference probability distribution (e.g. observed distribution). Smaller of KS statistic between the distribution from the interpolated data and the one from the observed data indicate better agreement with the observed data. The KS test statistics for two interpolated data sets (i.e. with and without the LNP model) are presented in Table S2 of the Supplementary Material. The precipitation data interpolated with LNP provide a smaller KS value than the other. The interpolated data with LNP can provide a more robust outcome than without the LNP model. For visual inspection, the histograms of observed and interpolated precipitation for three durations are presented in Figure S1 of the Supplementary Material. As shown in
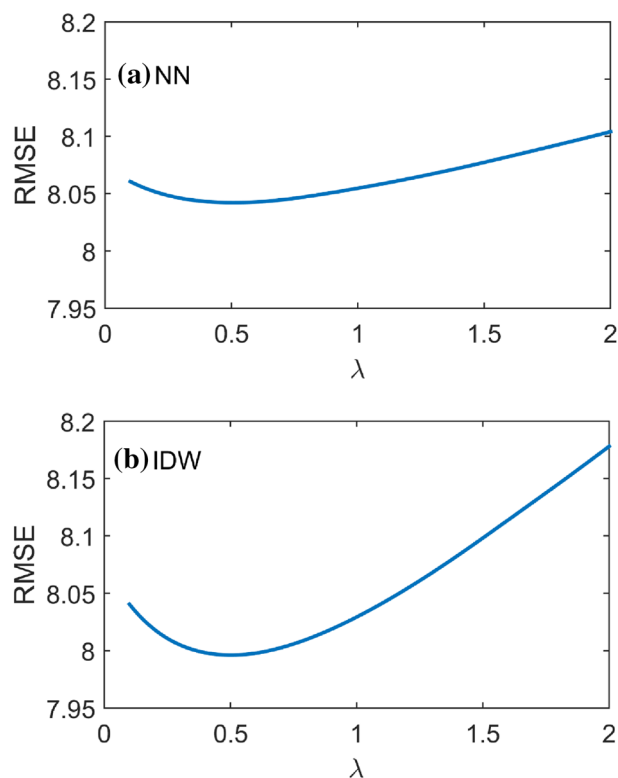
**Figure 4.** RMSEs of k-fold cross-validation outputs obtained with nearest-neighbor (NN) and inverse distance weight (IDW) interpolations for the LNP method when estimating the best λ.

| lambda | NN | IDW |
|--------|-------|-------|
| 0.0 | 8.343 | 8.343 |
| 0.5 | 8.042 | 7.996 |
| 1.0 | 8.055 | 8.029 |
| 1.5 | 8.077 | 8.098 |
| 2.0 | 8.104 | 8.178 |
| 2.5 | 8.133 | 8.263 |
| 3.0 | 8.162 | 8.351 |

**Table 1.** RMSEs of the estimated precipitation with nearest-neighbor (NN) and inverse distance weights for different lambda values.

this figure, the histogram of precipitation with LNP is more similar to the histogram of observed precipitation than the one without the LNP model.

The effects of the density of the weather station network and precipitation duration on the tuning parameter were further analyzed. The density of the network can substantially affect the tuning parameter because a longer interpolation area range must be estimated with a lower density network. To test this characteristic, an experiment was performed by randomly selecting stations with varying network densities from 40 to 100 km/station. A further range was not feasible due to the size of the tested area. Using KFCV, the optimal values were estimated and averaged. The detailed results of each case are presented in Figs. S2–S7 of the Supplementary Material. Additionally, the effects of different precipitation durations on the tuning parameter (λ) were also tested.

The characteristics of the tuning parameter as affected by the density network are well presented in the bottom panel of Fig. 5. A network density higher than 50 km/station (i.e., each station has an average distance of 50 km from another station) presents a tuning parameter value of less than one, and the parameter value converges to 0.5 with increasing distance. With a smaller network density, the optimal tuning parameter presents a higher value. This result might be induced by the fact that more accurate interpolated precipitation outputs are possible with a denser network. This intensifies the magnitude of the tuning parameter since dry regions are well-established in high-density networks.
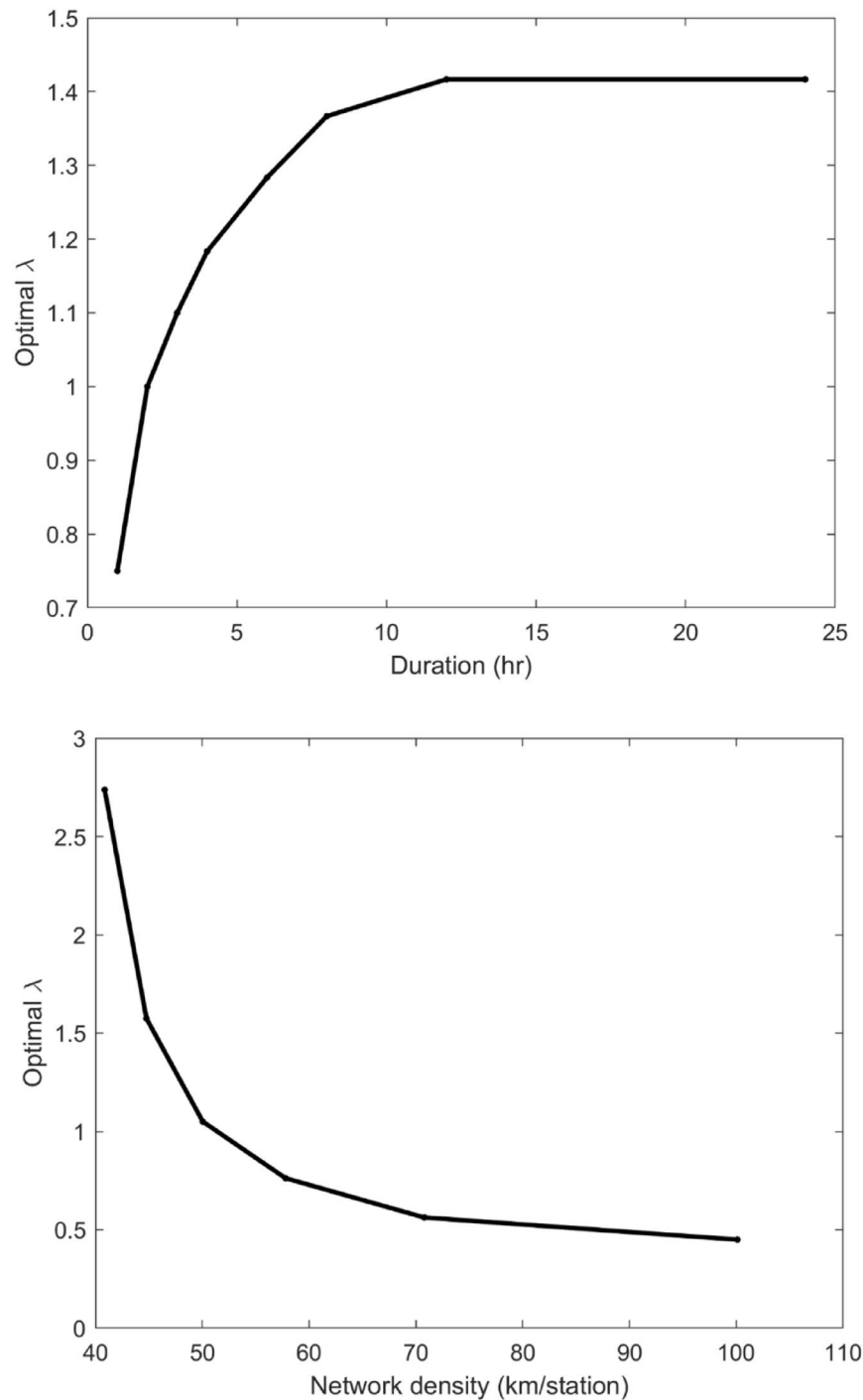
**Figure 5.** Variation in the optimal lambda along with the precipitation duration (top panel) and network density (bottom panel). Note that the average distance to the nearest station in the current study is 30 km.

The estimated tuning parameter increases up to 1.4 as the precipitation duration increases, as shown in the top panel of Fig. 5. Note that a precipitation event with a longer duration is likely to have a larger wet area than one with a shorter duration since the area is exposed longer to the storm system. In this case, wet areas in precipitation events with longer durations illustrate less sparsity than wet areas in shorter precipitation events. To obtain a small dry area in a large wet area, a high optimal parameter value is needed.

It is obvious that interpolated precipitation estimates in wet areas will decrease compared with the case that the LNP approach is not applied because the LNP approach assigns a negative value (but eventually zero values) to the stations with no precipitation. In order to investigate this behavior further, the average of the mean areal precipitation for all rainfall events using IWD with and without the LNP during the studied period over South Korea was calculated and presented in Table 2. As shown in Table 2, the IDW with the LNP is smaller than IDF

| Duration (hr) | IDW without LNP (mm) | IDW with LNP (mm) | RD(%) |
|---|---|---|---|
| 1 | 1.071 | 0.809 | 24.463 |
| 2 | 1.656 | 1.292 | 21.981 |
| 3 | 2.149 | 1.729 | 19.544 |
| 4 | 2.569 | 2.121 | 17.439 |
| 6 | 3.392 | 2.887 | 14.888 |
| 8 | 3.987 | 3.463 | 13.143 |
| 12 | 5.22 | 4.667 | 10.594 |
| 24 | 7.603 | 7.075 | 6.945 |

**Table 2.** Average of mean areal precipitation for all rainfall events during the studied period from 2003 to 2017 across South Korea. Note that RD = (IDW w/o LNP—IDW w LNP)/(IDW w/o LNP)*100.

without the LNP as assumed. The difference between rainfall amounts from IDW with and without the LNP is around 0.45 mm for all employed durations and about 25 to 7 percent. The difference increases as the duration become longer. For rainfall events of 1- and 24-h duration, the difference of rainfall amount is 0.26 mm and 0.55 mm, respectively.

With the overall assessment of the tuning parameter, it is concluded that the parameter has a critical role in determining dry areas via spatial interpolation. The parameter increases as the duration of a precipitation event increases and as the network density of weather stations increases. In a general case, the tuning parameter value can be set between 0.5 and 1.5 as a rule of thumb if the accuracy of the interpolated area is not significantly important.

The physical assumption of the delineation of the dry area due to pressure difference was thoroughly tested. Mean sea level (MSL) pressure data from ERA5 (5th Generation ECMWF atmospheric ReAnalysis of the global climate) was used to calculate the pressure at the points of interest. The MSL pressure near Korean peninsula during precipitation events in 2016 was matched to the grid points of precipitation data using IDW, and then characteristics of MSL pressure for dry and wet points were estimated. Figure S8 presents the spatial distribution of MSL pressure and precipitation data on April 4, 2016. There is a good agreement between pressure and wet/dry area.

In the current study, GPM (Global Precipitation Measurement) IMERG (Integrated Multi-satellitE Retrievals) Final Precipitation Level 3 data was used to illustrate how well the proposed LNP model can produce the spatial distribution of a rainfall event. The data was downloaded from https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGDF_06. The GPM data set is the gridded precipitation data produced by merging multiple satellite-based microwave data with calibration using ground truth data. The precipitation data by IDW with LNP using observed precipitation data and GPM are presented in Figure S9. The overall good agreement of dry area between two precipitation data sets was observed. However, the satellite-based precipitation data contains strong uncertainty[30–32]. The non-zero precipitation that appeared in observed precipitation on the exact location of weather stations are true value since these results were produced based on the ground truth. Thus, the satellite-based precipitation leads to false estimation on some stations. Nevertheless, the result supports that the LNP approach is useful and can be applied to calibrate microwave-based precipitation rate estimation.

## Summary and conclusions

When spatially interpolating a precipitation event, dry areas with zero precipitation have historically been ignored. Water management and impact assessments of climate change resulting from precipitation events require determining wet and dry regions appropriately. The current study proposed a novel approach to delineate wet and dry regions by assigning LNP to stations with precipitation values of zero. The proposed LNP method was tested with weather stations covering South Korea.

The overall results indicate that the LNP method can adequately delineate wet and dry regions with the appropriate tuning parameter. Additionally, the tuning parameter plays a special role in that its value increases with longer precipitation durations and denser networks. It is further suggested that a tuning parameter value of 0.5 – 1.5 can be used as a rule of thumb when the accuracy of the interpolation output is not critical.

The proposed LNP can be further tested over a much larger area, such as the United States, to enhance the spatial interpolation outputs. Additionally, the suggested LNP method can be employed for another variable, such as snow cover, for spatial interpolations over regions with sparse measurements. Also, the proposed LNP model cannot consider temporal dependency and physical features such as horizontal heterogeneity since it statistically interpolates the dry regions. Further model development can be made in the future study by taking the temporal dependency and physical features into account.

# References

1. Vieux, B. E. Finite element analysis of hydrologic response areas using geographic information systems. *A dissertation submitted in partial fulfillment of the degree of Doctor of Philosophy. Department of Agricultural Engineering. Michigan State University* (1988).
2. Vieux, B. E. *Distributed Hydrologic Modeling Using GIS.* (Kluwer Academic Publishers, 2004).
3. Palecki, M. A., Angel, J. R. & Hollinger, S. E. Storm precipitation in the United States. Part I: meteorological characteristics. *J. Appl. Meteorol.* **44**, 933–946. https://doi.org/10.1175/JAM2243.1 (2005).
4. Hwang, S. & Graham, W. D. Development and comparative evaluation of a stochastic analog method to downscale daily GCM precipitation. *Hydrol. Earth Syst. Sci.* **17**, 4481–4502. https://doi.org/10.5194/hess-17-4481-2013 (2013).
5. Wood, A. W., Leung, L. R., Sridhar, V. & Lettenmaier, D. P. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Clim. Change* **62**, 189–216 (2004).
6. Hidalgo, H. G., Dettinger, M. D. & Cayan, D. R. Downscaling With Constructed Analogues: Daily Precipitation and Temperature Fields Over the United States. 48 (California Energy Commission, California, 2008).
7. Lee, T. & Singh, V. P. *Statistical Downscaling for Hydrological and Environmental Applications.* Vol. 1 (CRC Press, 2018).
8. Chen, T. *et al.* Comparison of spatial interpolation schemes for rainfall data and application in hydrological modeling. *Water (Switzerland)* **9**, https://doi.org/10.3390/w9050342 (2017).
9. Thiessen, A. H. Precipitation averages for large areas. *Mon. Weather Rev.* **39**, 1082–1084 (1911).
10. Shekhar, S. & Xiong, H. *Inverse Distance Weighting.* (Springer, 2008).
11. Goodale, C. L., Aber, J. D. & Ollinger, S. V. Mapping monthly precipitation, temperature, and solar radiation for Ireland with polynomial regression and a digital elevation model. *Climate Res.* **10**, 35–49. https://doi.org/10.3354/cr010035 (1998).
12. Ly, S., Charles, C. & Degré, A. 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**, 2259–2274. https://doi.org/10.5194/hess-15-2259-2011 (2011).
13. Bayraktar, H. & Turalioglu, F. S. A Kriging-based approach for locating a sampling site-In the assessment of air quality. *Stoch. Environ. Res. Risk Assess.* **19**, 301–305. https://doi.org/10.1007/s00477-005-0234-8 (2005).
14. Murphy, J. An evaluation of statistical and dynamical techniques for downscaling local climate. *J. Clim.* **12**, 2256–2284 (1999).
15. Wilby, R. L. *et al.* Guidelines for use of climate scenarios developed from statistical downscaling methods. *Guidelines for Use of Climate Scenarios Developed from Statistical Downscaling Methods* (2004).
16. Wilby, R. L. & Wigley, T. M. L. Downscaling general circulation model output: a review of methods and limitations. *Prog. Phys. Geogr.* **21**, 530–548 (1997).
17. Phatak, A., Bates, B. C. & Charles, S. P. Statistical downscaling of rainfall data using sparse variable selection methods. *Environ. Model. Softw.* **26**, 1363–1371. https://doi.org/10.1016/j.envsoft.2011.05.007 (2011).
18. Timbal, B., Fernandez, E. & Li, Z. Generalization of a statistical downscaling model to provide local climate change projections for Australia. *Environ. Model. Softw.* **24**, 341–358. https://doi.org/10.1016/j.envsoft.2008.07.007 (2009).
19. Wilby, R. L., Dawson, C. W. & Barrow, E. M. sdsm—a decision support tool for the assessment of regional climate change impacts. *Environ. Model. Softw.* **17**, 145–157. https://doi.org/10.1016/S1364-8152(01)00060-3 (2002).
20. Lee, T. & Jeong, C. Nonparametric statistical temporal downscaling of daily precipitation to hourly precipitation and implications for climate change scenarios. *J. Hydrol.* **510**, 182–196. https://doi.org/10.1016/j.jhydrol.2013.12.027 (2014).
21. Hessami, M., Gachon, P., Ouarda, T. B. M. J. & St-Hilaire, A. Automated regression-based statistical downscaling tool. *Environ. Model. Softw.* **23**, 813–834. https://doi.org/10.1016/j.envsoft.2007.10.004 (2008).
22. Hashmi, M. Z., Shamseldin, A. Y. & Melville, B. W. Statistical downscaling of watershed precipitation using Gene Expression Programming (GEP). *Environ. Model. Softw.* **26**, 1639–1646. https://doi.org/10.1016/j.envsoft.2011.07.007 (2011).
23. Maurer, E. P. & Hidalgo, H. G. Utility of daily vs. monthly large-scale climate data: An intercomparison of two statistical downscaling methods. *Hydrol. Earth Syst. Sci.* **12**, 551–563 (2008).
24. Salathé, E. P. Jr., Mote, P. W. & Wiley, M. W. Review of scenario selection and downscaling methods for the assessment of climate change impacts on hydrology in the United States pacific northwest. *Int. J. Climatol.* **27**, 1611–1621 (2007).
25. Lee, T. & Ouarda, T. B. M. J. Conditional stochastic simulation model for spatial downscaling for assessing the effects of climate change on hydro-meteorological variables. *Clim. Change* **150**, 163–180. https://doi.org/10.1007/s10584-018-2276-1 (2018).
26. Ahrens, B. Distance in spatial interpolation of daily rain gauge data. *Hydrol. Earth Syst. Sci.* **10**, 197–208. https://doi.org/10.5194/hess-10-197-2006 (2006).
27. Eischeid, J. K., Pasteris, P. A., Diaz, H. F., Plantico, M. S. & Lott, N. J. Creating a serially complete, national daily time series of temperature and precipitation for the western United States. *J. Appl. Meteorol.* **39**, 1580–1591. https://doi.org/10.1175/1520-0450(2000)039%3c1580:CASCND%3e2.0.CO;2 (2000).
28. Diskin, M. H. On the computer evaluation of Thiessen weights. *J. Hydrol.* **11**, 69–78. https://doi.org/10.1016/0022-1694(70)90116-2 (1970).
29. Son, C., Lee, T. & Kwon, H. H. Integrating nonstationary behaviors of typhoon and non-typhoon extreme rainfall events in East Asia. *Sci. Rep.* **7**, https://doi.org/10.1038/s41598-017-04629-1 (2017).
30. Becker, A. *et al.* A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901-present. *Earth Syst. Sci. Data* **5**, 71–99. https://doi.org/10.5194/essd-5-71-2013 (2013).
31. Huffman, G. J. *et al.* The global precipitation climatology project (GPCP) combined precipitation dataset. *Bull. Am. Meteor. Soc.* **78**, 5–20. https://doi.org/10.1175/1520-0477(1997)078%3c0005:TGPCPG%3e2.0.CO;2 (1997).
32. Nicholson, S. E. *et al.* Validation of TRMM and other rainfall estimates with a high-density gauge dataset for West Africa. Part I: validation of GPCC rainfall product and pre-TRMM satellite and blended products. *J. Appl. Meteorol.* **42**, 1337–1354 (2003).

## Acknowledgements

## Author contributions

T.S. conceived the project and designed the study. J.Y. drafted the manuscript and programmed the method. All authors reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-99888-4.

**Correspondence** and requests for materials should be addressed to J.S.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.