

Characterizing the Uncertainty and Assessing the Value of Gap-Filled Daily Rainfall Data in Hawaii

RYAN J. LONGMAN

East-West Center, Honolulu, Hawaii

ANDREW J. NEWMAN

Research Applications Laboratory, National Center for Atmospheric Research, Boulder, Colorado

THOMAS W. GIAMBELLUCA

Water Resource Research Center, and Department of Geography and Environment, University of Hawai'i at Mānoa, Honolulu, Hawaii

MATHEW LUCAS

Department of Geography and Environment, University of Hawai'i at Mānoa, Honolulu, Hawaii

(Manuscript received 8 January 2020, in final form 14 May 2020)

ABSTRACT

Almost all daily rainfall time series contain gaps in the instrumental record. Various methods can be used to fill in missing data using observations at neighboring sites (predictor stations). In this study, five computationally simple gap-filling approaches—normal ratio (NR), linear regression (LR), inverse distance weighting (ID), quantile mapping (QM), and single best estimator (BE)—are evaluated to 1) determine the optimal method for gap filling daily rainfall in Hawaii, 2) quantify the error associated with filling gaps of various size, and 3) determine the value of gap filling prior to spatial interpolation. Results show that the correlation between a target station and a predictor station is more important than proximity of the stations in determining the quality of a rainfall prediction. In addition, the inclusion of rain/no-rain correction on the basis of either correlation between stations or proximity between stations significantly reduces the amount of spurious rainfall added to a filled dataset. For large gaps, relative median errors ranged from 12.5% to 16.5% and no statistical differences were identified between methods. For submonthly gaps, the NR method consistently produced the lowest mean error for 1- (2.1%), 15- (16.6%), and 30-day (27.4%) gaps when the difference between filled and observed monthly totals was considered. Results indicate that gap filling prior to spatial interpolation improves the overall quality of the gridded estimates, because higher correlations and lower performance errors were found when 20% of the daily dataset is filled as opposed to leaving these data unfilled prior to spatial interpolation.

1. Introduction

Serially complete daily rainfall time series are necessary for numerous applications in the fields of hydrology, resource management, and ecosystem modeling. Many agencies rely on direct analysis (e.g., from daily to monthly aggregation) or model output derived from these time series to inform management decisions. Despite the importance of complete records, almost all instrumental time series are affected by some amount of

missing values (Simolo et al. 2010). Missing data occur for a number of reasons including the failure of observation stations, the removal of erroneous values during quality control, inaccurate manual data entry, loss of data that is due to poor data management, and defective storage technologies (Tannenbaum 2009; Kim and Ryu 2016). Estimates of missing data can be used as a means of constructing a serially complete time series.

Identifying an appropriate method for filling gaps in a daily rainfall time series for broad applications is a challenge. A number of statistical and interpolation gap-filling techniques have been used with varying levels of

Corresponding author: Ryan J. Longman, rlongman@hawaii.edu

DOI: 10.1175/JAMC-D-20-0007.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](https://www.ametsoc.org/PUBSReuseLicenses) (www.ametsoc.org/PUBSReuseLicenses).

complexity ranging from rather simple to extremely complex. The simplest gap-filling techniques involve the imputation of an observation stations climatological median or mean based on the available historical record at the gauge (e.g., McCuen 2004) or using a prorated value (e.g., Longman et al. 2015b) calculated as the mean of observations over a discrete period of time (e.g., all available observations in a given month). Another simple approach is to directly impute an observation from a predictor station or stations into a missing field at a target station based on the predictor stations' positive correlation (e.g., Eischeid et al. 2000) or proximity (e.g., Vicente-Serrano et al. 2009) to a target station. Other approaches include gauge mean estimator, which takes an average of rainfall values at a select number of nearby stations (e.g., McCuen 2004); inverse distance weighting, where observations from a select number of stations are weighted by reciprocal distance to a target station (e.g., Lu and Wong 2008) or by correlation with a target station (e.g., Teegavarapu and Chandramouli 2005); normal ratio method (e.g., Longman et al. 2018), which uses the ratio of the target station mean to the mean of a predictor station as adjustment factors in the estimation procedure; linear regression and time series analysis methods (e.g., Vicente-Serrano et al. 2009), where statistical relationships between a target and predictor stations are established based on a period of overlapping data and then used to make a prediction (e.g., Salas 1993); and quantile mapping (e.g., Newman et al. 2015), where estimates are derived from a target stations cumulative distribution based on the quantile associated with an observation in the cumulative distribution at a predictor station. More sophisticated methods for gap filling daily rainfall include multiple linear regression (e.g., Kagawa-Viviani and Giambelluca 2020), nonlinear deterministic and stochastic variance dependent interpolation techniques (e.g., kriging; Adhikary et al. 2016), the use of artificial neural networks (Teegavarapu and Chandramouli 2005), the use of autoregressive conditional heteroscedasticity models (e.g., Gao et al. 2018), or a combination of many of the abovementioned methods (see Teegavarapu et al. 2018).

Another approach to handling data gaps is to exclude periods or individual observation stations with missing values, or to ignore the gap completely during an analysis. Ignoring gaps in the data may disregard valuable information and can induce biases in an investigation (Simolo et al. 2010). This can become especially problematic when daily rainfall data are aggregated to the monthly time step because of the fact that leaving a gap in a daily record is equivalent to setting the day to zero. The inclusion of additional zero rainfall days can potentially add negative bias to a monthly total. On the

other hand, if incomplete stations are omitted from the monthly record, both sample size and area of coverage will be reduced.

The majority of the gap-filling methods mentioned above involve some type of parameterization decision prior their execution. However, finding the most appropriate parameter selection process is not always straightforward. For some methods, statistical relationships between predictor stations and a station targeted for gap filling must first be established. Once statistical relationships are established, the station(s) with the highest correlation is (are) selected as the candidate predictor station(s) used to fill a gap. Typically, a correlation threshold is set to avoid using predictor stations with weak statistical relationships with a target station. However, both the amount of overlapping station data needed to establish a statistical relationship and the selection of a correlation threshold are tunable parameters in the gap-filling process. The appropriate number of predictor stations to use in an estimate must also be decided for methods using more than one station. Other tunable parameters include the area in which predictor stations are selected (neighborhood of influence) and weights given to predictor stations as well as other statistical thresholds that are set to include or exclude predictions.

Parameterization decisions for individual gap-filling methods are typically addressed by researchers on an individual basis either by performing some type of method optimization on a sample dataset (e.g., Eischeid et al. 2000), by following guidelines published by others in the scientific literature (e.g., Longman et al. 2018), or by arbitrary choice (Teegavarapu and Chandramouli 2005). Research has shown that the use of optimized exponents in distance and correlation-based weighing methods, classifiers for rain or no-rain conditions, and an optimal number of predictor stations improves estimates of missing data (Teegavarapu 2014; Teegavarapu et al. 2018). Despite their importance, these steps are time consuming to complete and not always straightforward; therefore, oftentimes they are not executed correctly or completely overlooked (e.g., Caldera et al. 2016).

Identifying the best gap-filling method and most appropriate parameterization of that method can be a difficult task because the quality of the predictions varies in both space and time. Therefore, it is important to identify a method that works across the spatial and temporal domain of interest. This depends on several factors including topography, climatic conditions, and the density of the rain gauge network. Guidance relating to these factors in designing a suitable gap-filling strategy is limited.

An equally important aspect of gap filling that is often overlooked is the characterization of uncertainty within a gap-filled dataset. Oftentimes, studies reduce uncertainty estimates to a single value that holds limited practical information about performance. This is especially true if the results of an error analysis are not presented in absolute units or if the mean is used as a measure of central tendency in a nonnormal distribution. In addition, limited understanding exists of the uncertainty associated with aggregating daily gap-filled rainfall to a monthly time step. Specifically, how many days in a given month are acceptable to fill with a given method and what is the relationship between gap size and the magnitude of error? Last, there is lack of information pertaining to the value of gap filling missing point data prior to the creation of a gridded surface (e.g., monthly rainfall map). While serially complete data are necessary for many application models, it is not a requirement for the execution of many spatial interpolation techniques. Interpolating only available data with variable station density in time versus serially complete data with static station density in time will impact the final interpolation field in different ways.

The primary goals of this study are to 1) identify the most appropriate technique for gap filling daily rainfall in Hawaii, 2) characterize the uncertainty associated with gap filling data across several different size temporal gaps, and 3) determine the value of gap filling prior to spatial interpolation. To accomplish these goals, several research objectives were identified. First, five easily executable gap-filling methods—normal ratio (NR), linear regression (LR), inverse distance weighting (ID), quantile mapping (QM), and single best estimator (BE)—are described as a reference for the community. Next, these gap-filling methods are optimized for our study area and evaluated based on their ability to fill a multiyear gap in a time series. Then, these five methods and two additional gap-filling methods that are not appropriate for filling long temporal gaps, the prorating (PR) and “equals zero” (EZ) approaches, are used to characterize the uncertainty associated with filling submonthly data gaps in daily records on monthly aggregations. Last, the value of gap filling daily point data prior to the creation of a monthly gridded surface is examined.

2. Methods

a. Study site and data

The climate of Hawaii Island is characterized by extreme spatial gradients in rainfall and temperature due to its complex topography, persistent northeast trade

wind flow, and the presence of persistent temperature inversion (~90% of the year) found at an average base height of 2150 m (Giambelluca et al. 2013; Longman et al. 2015a; Frazier et al. 2016). Midelevation areas on the windward side of the island receive up to 7600 mm of rainfall annually, with measurable rainfall on nearly 90% of the days (Giambelluca et al. 2013; Longman et al. 2019). On the lee side of the islands semiarid conditions prevail, with less than 250 mm of annual rainfall with measurable rainfall occurring on 20% of the days. The highest elevation observation site on the island (3410 m) receives ~400 mm annually, with rainfall occurring on fewer than 1 in 10 days (Giambelluca et al. 2013; Longman et al. 2019). These features make Hawaii Island a compelling scientific testbed in which to evaluate gap-filling methods.

The rainfall data utilized in this analysis were drawn from a quality-controlled 25-year (1990–2014) observed daily rainfall dataset for the Hawaiian Islands (see Longman et al. 2018). A total of 75 stations located on Hawaii Island were selected on the basis of their completeness of record during the study period (Fig. 1). This dataset is utilized in several ways to address the needs of this study. First, all available observations between 1990 and 2012 were used to establish statistical relationships between stations. Next, all available observations between 2013 and 2014 were used to validate daily rainfall predictions derived from the NR, LR, ID, QM, and BE gap-filling approaches. Last, all of the 31-day months (January, March, May, July, August, October, and December) with complete data during the 2013–14 period were identified and were used to characterize the uncertainty of filling smaller gaps (between 1 and 31 days missing). We use Spearman's rank correlation (“correlation” throughout) to define station correlations because it is more appropriate for the assessment of precipitation data.

b. Gap-filling approaches

In this section, the seven gap-filling approaches used in this analysis (NR, LR, ID, QM, BE, PR, and EZ) are described. Some theoretical thoughts are provided at the end of the section.

1) NR

The NR method (Paulhus and Kohler 1952) uses the ratios of the mean rainfall at a target station to the means of the highest correlated predictor stations as adjustment factors in the estimation of missing rainfall at the target station. Eischeid et al. (2000) used the NR method in conjunction with other methods to fill daily gaps for both rainfall and temperature time series in the western United States. In Hawaii, the NR method has

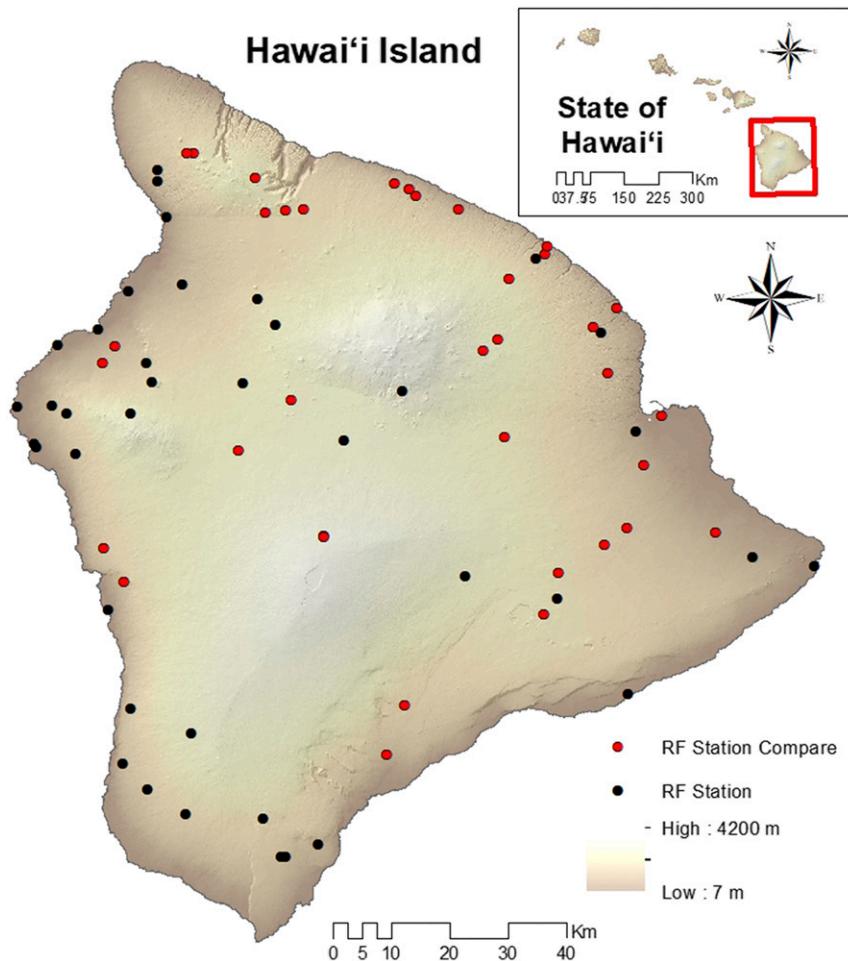


FIG. 1. Elevation map of Hawaii Island with locations of rainfall observation stations; “RF station” indicates all available rainfall stations on Hawaii Island, and “RF station compare” indicates the stations used to directly compare gap-filling approaches.

been used previously to fill both monthly (Giambelluca et al. 2013) and daily (Longman et al. 2018) rainfall time series. To execute this method, first, station means are calculated at both the target and predictor stations using all available data between 1990 and 2012. An alternative approach is to use individual monthly climatologies, which we did not do in this analysis due to the short period of record at some stations (e.g., 3 years). Next, the ratio between the target and predictor station is calculated. Last, the daily observed rainfall at the highest correlated predictor station(s) is multiplied by the corresponding ratio. When multiple predictor stations are used, the mean of the predictor station estimates is used as the fill value at the target station. The NR method can be expressed with the following equation:

$$\widehat{P}_x = \frac{1}{n} \left(\sum_{i=1}^n P_i \frac{N_x}{N_i} \right), \quad (1)$$

where \widehat{P}_x is the predicted rainfall at a target station x , N_x is the average rainfall at station x , N_i is the average rainfall at the i th-best correlated station, and P_i is the rainfall at the i th-best correlated station-predictor station. One major limitation of the NR method in this form is that the method may fail to take into account redundant information from spatially clustered stations when multiple predictor stations are used (McCuen 2004).

2) LR

With the LR method, gaps are filled on the basis of the linear regression relationship between two stations applied to observed data. One station is considered to be the explanatory station (x) and the other is considered to be the dependent station (y). To execute this method, the relationships between a target station y and all potential predictor stations x are established using a user defined number of overlapping data points, which is

typically an arbitrary choice. Next the slope and the intercept of this linear relationship were used to estimate missing data at the target station from a predictor station. The LR method can be expressed as

$$\widehat{P}_x = \widehat{b}P_i + \widehat{a}, \tag{2}$$

where \widehat{P}_x is the predicted value at the target station, P_i is the observation from a predictor station, \widehat{b} is the slope of the line, and \widehat{a} is the intercept. Note that in some applications of this method the regression line with a target station is forced to pass through the origin to avoid negative values and retain the zeros (Vicente-Serrano et al. 2009). In this study, during an initial test period, it was determined that forcing the regression through the origin did not improve the quality of the predictions. Therefore, a conditional statement is included to set all negative predictions to zero.

3) ID

The ID method is one of the most commonly used approaches for filling data gaps in hydrological studies (Teegavarapu et al. 2009). The ID method estimates rainfall at a target station using a linear combination of surrounding predictor station values weighted by an inverse function of the distance between each predictor station and the target station (Li and Heap 2008). The ID method can be expressed with the following equation:

$$\widehat{P}_x = \sum_{i=1}^n P_i w_i, \tag{3}$$

where \widehat{P}_x is the predicted value at the target station x , P_i are the observations from predictor stations, n is the number of stations used in the interpolation, and w_i are the weights, defined as

$$w_i = \frac{1/d_i^\lambda}{\sum_{i=1}^n 1/d_i^\lambda}, \tag{4}$$

where λ is the weighting parameter and d_i is the distance between the target station x and predictor station i .

4) QM

For a given target station, missing rainfall values are filled using the QM method (Panofsky and Brier 1968), which consists of the following steps: (i) predictor stations are ranked from highest to lowest on the basis of their correlation with a target station, (ii) cumulative distribution functions (CDFs) are computed for some user-defined overlapping data period at both target and predictor stations, (iii) the cumulative probability of an

observation at a predictor station is computed from the predictor station CDF, and that cumulative probability is used in the target station CDF to determine the fill value at the target station.

5) BE

The BE approach consists of imputing the concurrent rainfall value at the highest-correlated predictor station into the missing field at a target station (Eischeid et al. 1995). This is not necessarily the best possible estimate using a well-defined estimation method, for example, the best linear unbiased estimate (BLUE).

6) PR

The PR approach is used much like the imputation of the climatological mean; however, in this analysis we calculate the mean on the basis of all available data at a target station within a given month/year. For example, if a 10-day gap is being filled in a 31-day month, the prorated value is calculated as the average rainfall occurring over the 21 days with observations. This average value is then used to fill each day in the 10-day gap. In other words, daily average derived from the available data within a given month is multiplied by the number of days in the month to get the estimated monthly total. The prorating approach could be applied (albeit with caution) with as little as one daily observation being available in a given month.

7) EZ

In some cases, researchers aggregating daily data to the monthly time step may choose to set the monthly total to the total of the available data, regardless of the number of missing days. This amounts to setting the missing values to zero.

Methods 1–3 and 5 (NR, LR ID, and BE, respectively) are all spatial interpolation methods that use information from the target and predictor stations. The NR method uses a ratio to rescale the predictor station (and not a bias correction), so the shape of the probability density function (PDF) of the predictor station is scaled to be narrower or wider if the ratio is respectively less than or greater than 1. For the LR method, both a slope and intercept term are used, so the PDF of the predictor station can be shifted and scaled (even when $N = 1$). For the ID method, if $N > 1$ the distributional change is a weighted combination of the predictor stations and translates and scales distributions. For the BE method when $N = 1$, the method just translates the PDF left or right (increase or decrease amounts) and does not change the shape of the distribution. Method 4 (QM) uses information from both target and predictor stations but preserves the distribution of the target station.

Method 6 (PR) is self-contained infilling using only information from the target station to fill missing data. Method 7 (EZ) replaces missing values with zeros, potentially causing drastic changes to the distribution.

c. Optimization method

Drawing on the 25-yr rainfall dataset described above, we selected a 2-yr sample (January 2013–December 2014), the period with the least number of missing observations within the dataset, for out-of-sample testing of the methods. For a given target station, we removed all observations and then filled the entire time series using each of the gap-filling approaches and under different optimization scenarios. The observed time series was then used to cross validate the results.

1) NUMBER OF PREDICTOR STATIONS FOR NR AND ID METHODS

For both the NR and ID methods, multiple predictor stations can be used to derive an estimate at a target station. The number of predictor stations to use must be decided, often in arbitrary fashion. For the NR method, [Eischeid et al. \(2000\)](#) suggested using no more than four predictor stations and that the use of more than four predictor stations can degrade the estimates. For the ID method, many researchers have suggested the use of three or four predictor stations (e.g., [Teegavarapu and Chandramouli 2005](#)). In this study, we test how the use of between 1 and 10 predictor stations can influence the results at a target station for both the NR and ID methods.

2) ESTABLISHING STATISTICAL RELATIONSHIPS BETWEEN STATIONS

The NR, LR, QM, and BE methods require that statistical relationships be established between stations. Both [Eischeid et al. \(2000\)](#) and [Tardivo \(2015\)](#) suggest that a stable estimation statistic can be derived from 10 years of overlapping data between stations. Because of missing data and relatively short periods of record, even using a dataset that spans 25 years, requiring a 10-year period of overlap in this present study would eliminate most of the available sample stations. Therefore, statistical relationships were established between stations when at least 1000 days of overlapping data between stations were available (note that this was choice and not an optimized parameter). Statistical relationships between stations were calculated using only data available during the 23-year period (1990–2012) so as to avoid use of data during the 2013–14 test period. As a subsequent step, the statistical significance of each station relationship was tested using the correlation between two stations and all stations relationships that do not have statistically significant ($p \leq 0.05$) positive correlation

with a given target station are not considered as potential predictor stations for that target station.

3) CORRELATION THRESHOLD

For the NR, LR, QM, and BE approaches, predictor stations are typically selected based on the strength of their correlation with a target station. The choice of this correlation threshold, however, is arbitrary. In prior studies, candidate predictor stations were selected using minimum Pearson correlation coefficient r thresholds of 0.35 (for several gap-filling methods; [Eischeid et al. 2000](#)), 0.5 (for the nearest-neighbor method; [Vicente-Serrano et al. 2009](#)), and 0.88 (for the NR method; [Giambelluca et al. 2013](#)). In this study, the threshold correlation between target and predictor stations was tested by comparing the errors associated with using predictor stations within six correlation ranges (0.3–0.4, 0.4–0.5, 0.5–0.6, 0.6–0.7, 0.7–0.8, and >0.8). First, all target stations that were significantly correlated with at least one predictor station within each threshold range are identified (i.e., a station must have a complete range of correlations with other stations). This allowed for an equal sample size within each threshold range. The four gap-filling approaches that use correlation as a criterion for selecting a predictor station were applied, and the errors produced at each threshold range were analyzed. For this analysis the ID method is executed without any modifications. Correlation thresholds (as opposed to correlation ranges) were also tested using the entire sample dataset where gap filling was done, and errors were calculated using all available predictor stations below four thresholds (≤ 0.4 , < 0.5 , < 0.6 , and < 0.7).

4) NEIGHBORHOOD OF INFLUENCE

The arbitrariness in the choice of the neighborhood from which predictor stations are selected has been identified as a distinct limitation to the ID method ([Teegavarapu and Chandramouli 2005](#)). In this study, the relationship between distance to the nearest predictor station and error (difference between predicted and observed rainfall) is assessed for all methods. The distance between each target station and all of the predictor stations was calculated, and six distance ranges (90–117, 65–90, 41–65, 21–41, 4–21, and <4 km) were selected such that an equal sample size was achieved within each range. Errors were calculated for gap filling done at select target stations using unique predictor stations within each range.

5) WEIGHTING PARAMETERS FOR THE ID METHOD

The arbitrary choice of a weighting parameter λ assigned to stations when using the ID method is another limitation

to this method (Teegavarapu and Chandramouli 2005). With the ID method, the exponent λ is typically set to 2 (e.g., Shepard 1964; Webster and Oliver 2001). Here we test how the use of different values of the exponent ($\lambda = 1, 1.5, \text{ and } 2$) influences the results.

6) RATIO THRESHOLD FOR THE NR METHOD

Normal ratios are calculated as the long-term mean of a target station divided by the long-term mean of a predictor station with means based only on overlapping data periods between stations. Following Longman et al. (2018), a ratio range from 0.3 to 2.7 is used to identify candidate predictor stations.

7) RAIN/NO-RAIN BIAS CORRECTION

One limitation that applies to all of the gap-filling approaches is that they typically do not represent rain/no-rain percentages accurately. For gap-filling methods using multiple predictor stations, the number of rain days is generally overpredicted. With multiple predictor stations, if any of them has nonzero rainfall on a given day, then the estimated rainfall at the target station will be nonzero, leading to a bias in the rain day frequency (Hasan and Croke 2013). In addition, the possibility of overestimation of rainfall using the LR and QM methods that use a single predictor station is also present as rainfall is predicted based on a statistical relationship that could produce daily rainfall > 0 at a target station when zero rainfall is observed at a predictor station. Over estimation as a result of gap filling creates the need for a rain/no-rain correction to be included in the rainfall prediction. Corrections to rainfall estimates, based on rain/no-rain characteristics at the closest station (based on Euclidean distance) or the highest correlated station have been shown to be effective (Teegavarapu et al. 2018). In this work we tested rain/no-rain corrections that were based on both distance and correlation for all of the methods and compare the results with gap filling in the absence of the rain/no-rain correction.

d. Gap filling at various temporal resolutions

Once the optimal parameterization was established for each method, the same 2-year daily time series (2013–14) was gap filled and the methods were assessed on their ability to predict rainfall at each target station during this long artificial gap. The 2-year reconstructed time series was tested against the observed time series and errors assessed. Errors associated with gap filling small gaps in station record (between 1 and 31 days) are also tested using the following steps. For all complete 31-day station/months in the sample dataset ($n = 321$), a random day is removed from each monthly time series and the gap is filled using each of the seven methods

(NR, LR, ID, QM, BE, PR, and EZ). The filled daily data were then aggregated to monthly totals and compared with the observed monthly data. This process is repeated 31 times for each data-gap scenario (between 1 and 31 days missing). For example, to test the effect of filling a 10-day gap in a given month, 31 different samples were created, each with 10 daily values randomly removed, and then filled using each method. Errors (departures from the actual observed values) are then calculated for each of the filled values and the average of both the collective mean and median deviations between the gap filled and observed monthly rainfall are calculated. Note that for the PR approach gaps are filled for up to 30 days because at least one observation is needed for executing this method. Results are only compared in instances where all of the gap-filling methods can be employed. While this reduces the total number of samples, it allows for a homogeneous sample comparison between methods. For each gap size, error metrics are derived based on a combined sample of 308 481 combinations.

e. The added value of gap filling daily data

Serially complete data are important for a number of applications including the spatial interpolation of point data to a gridded surface. It can be assumed that the denser the station coverage the better the gridded surface will be. But does this assumption hold for partially gap-filled data? Here, this assumption was tested using sample months from the monthly dataset. First, the complete daily station records of fourteen 31-day months were aggregated to monthly totals. Next, two additional monthly datasets were created after 20% of the daily station records were randomly removed from the time series and another in which the removed stations were gap filled first and then aggregated to monthly. Then, gridded surfaces were created using a climatologically aided ordinary kriging approach (see Frazier et al. 2016), using three different monthly rainfall datasets: the monthly observed dataset, the monthly dataset with 20% of the stations removed, and the monthly dataset with 20% of the stations filled. Last, the gridded surfaces created with those three datasets are compared.

f. Performance criteria metrics and statistical tests

Several statistics are used to measure the effectiveness of the gap-filling approaches. The mean bias error (MBE) is used as measure of the mean over/under prediction of a method. The mean absolute error (MAE) is used as measure to see how close simulated values are to observed values assuming the data are normally distributed. The root-mean-square error (RMSE) is used to measure the overall spread of the errors in a normal

distribution. The median absolute error (MED) is used as a measure of central tendency in a nonnormal distribution. The coefficient of determination R^2 is used to measure the linear correlation between observed and filled data. These metrics among others are widely used for the performance of comparison of gap-filling methods (Kim and Ryu 2016).

Comparative analysis of the errors associated with both the optimization and execution of the individual gap-filling methods are carried out using both parametric and nonparametric statistical tests depending on the shape of the distribution. In instances where error distributions are normal and variances are equal, a one-way analysis of variance (ANOVA) was used to determine the degree to which mean errors differ between the gap-filling methods. When significant differences were identified, the Tukey's honestly significant difference (HSD) test was used to test for significant differences in each of the pairwise comparisons. For nonnormally distributed data, the Wilcoxon rank-sum test was used to identify significant differences in median errors. For both parametric and nonparametric tests, the null hypothesis H_0 is that differences between observed and estimated rainfall are from the same continuous distribution for any two gap-filled datasets. The alternative hypothesis H_a is that these two datasets are from different continuous distributions. The hypothesis tests were carried out at a statistical significance level of 5% ($\alpha = 0.05$).

g. Probability of precipitation

Daily probability of precipitation (PoP) can be used to quantify the frequency of spurious nonzero rainfall introduced during the gap-filling process (e.g., Newman et al. 2015). The differences in climatological mean PoP as a result of gap filling are evaluated for all of the various gap sizes. First, the PoP at each station is calculated as the number of observed rain days [days with > 0.15 mm of rainfall (RF)] divided by the total number of days in the time series. PoP is then calculated for the predicted datasets and the difference between the predicted and observed datasets is used as measure of prediction bias.

3. Results

a. Method optimization

Of the 75 observation stations examined in this study, only 36 met the criteria necessary for the execution of all of the gap-filling methods. On average, station records in the 2-year sample were 93% complete (range 59%–100%); therefore, error statistics are based on a sample of approximately 24 000 station-days.

1) NUMBER OF PREDICTOR STATIONS

For both the NR and the ID methods, ANOVA results showed no significant advantage to using more than one predictor station to gap fill rainfall at a target station. The optimal number of predictor stations to use varied by metric. However, for the NR method the lowest MAE and RMSE were obtained when three stations were used and for the ID method when 10 stations were used. Subsequent iterations using these methods were parameterized accordingly.

2) CORRELATION THRESHOLD

Thirteen target stations with at least one predictor station in each of the six correlation threshold ranges were identified. Results show no differences in the MED error between the NR, ID, and QM methods for predictor stations above the $r = 0.5$ – 0.6 range (Figs. 2a,b). Below this range, the ID method (which was executed without any modifications) produces significantly lower MED errors than all of the other methods. MAE errors are significantly lower for the ID method below the $r = 0.6$ – 0.7 range. Error was also assessed using MAE and MED for sets of all available stations at four correlation thresholds ($r < 0.4$, < 0.5 , < 0.6 , and < 0.7). Results show that, when compared with the other methods, the ID method produces significantly lower MED errors for predictor stations below the $r = 0.5$ threshold and significantly lower MAE errors for stations below the $r = 0.6$ threshold (Fig. 3). For subsequent portions of this analysis and for methods requiring the selection of predictor stations based on correlation, the $r = 0.6$ threshold is used to identify candidate predictor stations for the four gap-filling methods that use the correlation as a criterion for selecting candidate predictor stations.

3) NEIGHBORHOOD OF INFLUENCE

The ability to fill daily rainfall using predictor stations at six different distance thresholds was tested. On average the closest predictor station is 5.5 km away from a given target station and the average distance to the next closest station is 8.9 km with high variance across the dataset. In total, 14 target stations having a nonzero sample of predictor stations in each of the six distance threshold ranges (roughly 23-km bins) were identified for this test. Note that for the methods that require established statistical relationship with another station, the rank correlation was only considered when multiple predictor stations were available at a given distance threshold. Results show no statistical differences in the MAE or MED between methods at any of the distance thresholds analyzed (Figs. 2c,d). In general, both the

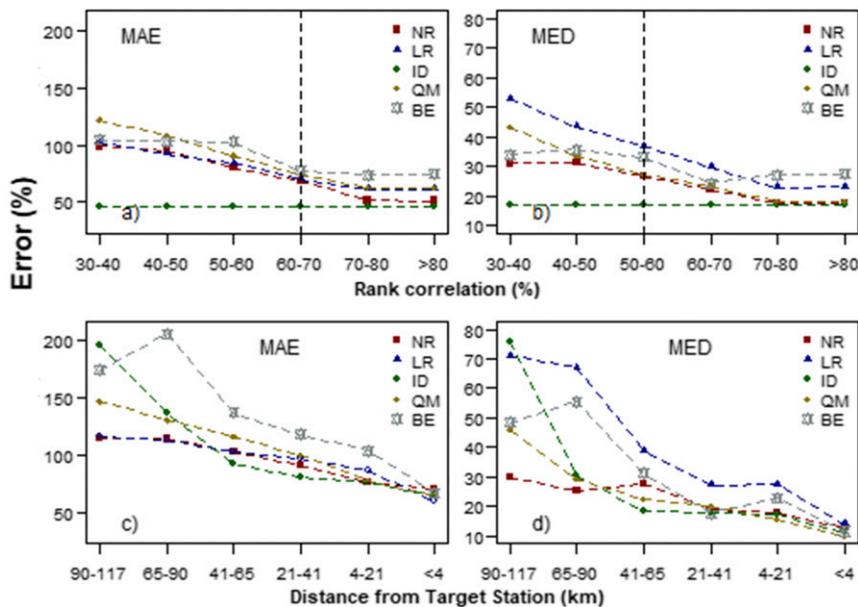


FIG. 2. (left) MAE and (right) MED associated with gap filling daily rainfall using predictor stations (a),(b) within six unique correlation ranges ($n = 13$ stations) and (c),(d) at six unique distance ranges ($n = 14$); the vertical dashed line in (a) and (b) marks the threshold at which the error produced when using the ID method is significantly lower than the error produced using all of the other approaches.

MAE and MED errors were higher at further distances from a target station.

4) WEIGHTING PARAMETERS FOR THE ID METHOD

The ID approach was tested using three different weighting approaches ($\lambda = 1, 1.5, \text{ and } 2$). No statistical differences were identified between the distributions of errors for the different approaches. The lowest errors were obtained when the standard exponent of two ($\lambda = 2$) was utilized. Therefore, the weight of $\lambda = 2$ was used in the execution of the ID method in subsequent analyses.

5) RAIN/NO-RAIN BIAS CORRECTION

The bias correction methods were evaluated based on the difference between observed and estimated PoP calculated after removing 20% of the data and gap filling. Figure 4, shows the deviation in PoP for predictions made with the five gap-filling methods with the distance (DIST) and correlation (COR) bias corrections and without a bias correction (NO). For the NR, LR, ID, and QM gap-filling approaches, results show that the inclusion of either a distance or correlation based bias correction significantly reduces the inclusion of spurious rainfall during gap filling. The distribution means, derived from Fig. 4, were closest to zero when the COR method was used for the NR, LR, and QM methods,

while the DIST correction produced the least bias for the ID method. For all of the methods, the range of errors was the smallest when using the COR correction. In subsequent portions of this analysis, the COR bias correction was applied to the NR, LR, and QM methods, and the DIST correction was applied to the ID method.

b. Evaluation of methods

1) OPTIMAL METHOD COMPARISON (LONG GAPS)

Five gap-filling approaches were assessed on their ability to accurately fill a large (~2 year) gap in a time series. This “long” 2-year gap allows for metrics to be calculated representing some seasonal and interannual RF variability in Hawaii. On average, 663 ± 79 days of data were filled at each of the 36 stations available for comparison. Results were fairly similar among methods over this long gap (Fig. 5). Comparative statistics for all performance metrics are presented in Table 1. No statistical difference was identified in the error distributions between any of the five gap-filling methods for any of the eight performance metrics considered. The lowest normalized MAE (62.2%) was obtained using the LR method, and the lowest MED ($12.5\% \pm 8.8\%$) was obtained using the NR method. The highest R^2 (62%) was obtained using the ID method. MED errors ranged from 12.5% to 16.5% across the methods suggesting that all methods do a fairly reasonable job of gap filling longer

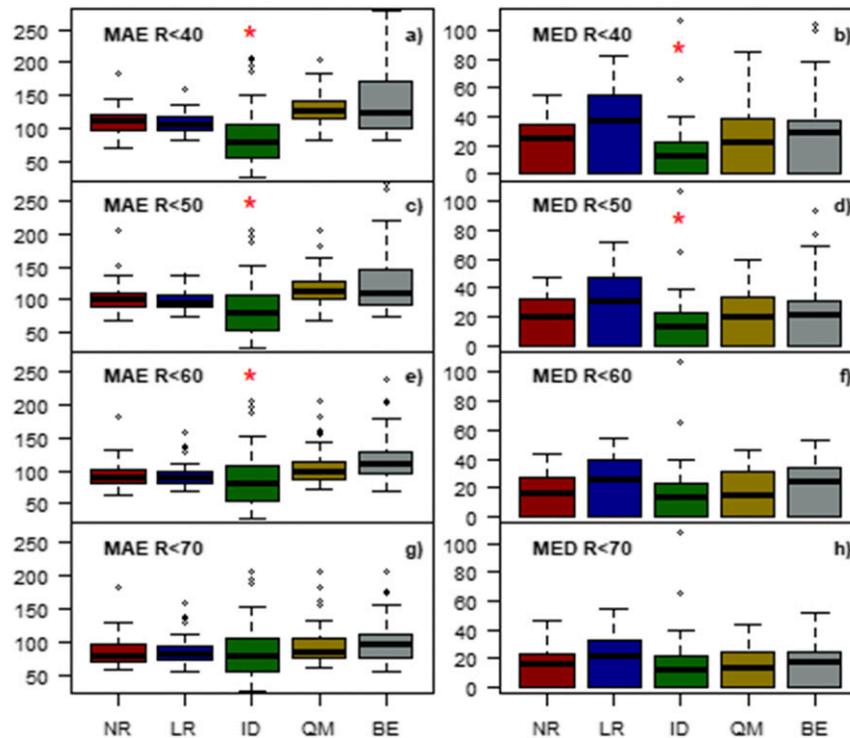


FIG. 3. (left) MAE and (right) MED for five gap-filling approaches including NR, LR, ID, QM, and BE and under four correlation thresholds: (a), (b) ≤ 0.4 , (c), (d) ≤ 0.5 , (e), (f) ≤ 0.6 , and (g), (h) ≤ 0.7 . The red star indicates where the errors produced from the ID method are significantly lower than the errors produced from all of the other methods; data points exceeding the maximum error [75th percentile (top of box) + $1.5 \times$ the interquartile range (whisker)] are plotted as outliers of the data, and those outliers exceeding the y-axis bounds on the graph are not shown.

records. In general, the MAE was high (62.2%–71.9%) for all methods, which suggests that large outliers are present in all of the gap-filled datasets. This is especially true for the BE method, which had the highest MAE (71.9%). The difference in PoP errors were generally negative but small for all of the methods. This suggests that while the rain/no-rain correction does well at reducing of spurious rainfall created during the filling process it also produces artificial zero values in instances where a positive rainfall value should have been predicted.

2) DATA AGGREGATION ERROR (SHORT GAPS)

Daily data were gap filled at each station and for each sample month using the same five gap-filling methods that were previously analyzed as well as the PR and EZ methods described earlier. In total, 231 complete station-months were used. First daily data were aggregated to a monthly time step and then the difference between observed and gap-filled data were compared. Median and mean deviations between observed and gap-filled monthly rainfall totals for 1–31-day data gaps are presented in Fig. 6. The distribution in errors was

determined to be nonnormal; therefore, nonparametric tests were used to identify statistical differences between the distributions and for each of the 31 different gap sizes. Results indicated no consistent statistical differences between the NR, LR, ID, and QM methods for any gap size tested. No significant differences were found between the ID and BE methods, and all methods were significantly better than the PR and EZ methods. Figure 7 shows the distribution of errors for 1-, 15-, and 30-day gaps and the frequency for which a single method outperformed the other methods as based on minimum deviation between observed and gap-filled monthly rainfall totals (Table 2). For 1-day gaps, mean errors ranged from 2.1% to 4.0%, and the NR method produces both the lowest mean (2.1%) and median (1.8%) errors. For 15-day gaps, mean errors ranged from 16.6% to 48.3% and, again, the NR method had the lowest mean (16.6%) and median (13.1%) errors. For 30-day gaps, the NR method had the lowest mean error (27.4%), but the QM method had the lowest median error (20.3%). In general, the NR method most consistently performed as the optimal method for all gap sizes considered.

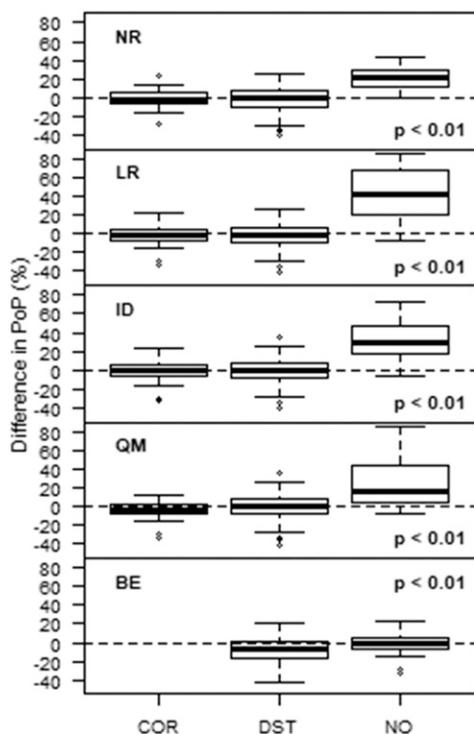


FIG. 4. Differences in the PoP between observed data and gap-filled data under three bias correction approaches— NO, COR, and DIST (middle column)—for five gap-filling methods: NR, LR, ID, QM, and BE. Data points falling outside the Q1–Q3 range are plotted as outliers of the data.

Surprisingly, the EZ method was the optimal method 22 times for 1-day data gaps and produced significantly lower errors than using the PR method. This is explained by the fact that the EZ method does not introduce any additional rainfall; therefore, setting 1-day gaps to zero in dry months is often an accurate estimate. For gaps from 3 to 26 days, however, the PR method was a significantly better choice than the EZ method. Errors for both of these methods were significantly higher, and, therefore, they should not be employed regardless of the gap size.

3) ADDED VALUE OF GAP FILLING

The added value of gap filling daily rainfall prior to the spatial interpolation of monthly rainfall was assessed by creating a gridded surface with and without gap filling and then comparing these products to a gridded surface created with complete observations. Missing daily rainfall was gap filled using first the NR method when statistical thresholds were met and the ID method for all other instances. The combined use of these methods allows for the creation of the serially complete datasets A 14-month composite of the differences between gridded surfaces with missing and

filled data is shown in Fig. 8. Note that in Fig. 8a composite errors are a result of missing data and in Fig. 8b errors are a result of the gap filling. Errors were found in both composites, but the mean and median errors were higher in the composite with the missing station data. In general, in the missing data composite, there was a large overprediction on the windward side of the island, which was greatly reduced in the gap-filled composite. The individual grid cells produced with and without gap filling were also compared directly with observations (Figs. 9a,b). For the gap-filled product, the correlation between observed and predicted rainfall was higher and all errors were considerably lower for the product created with missing observations. The distribution of errors was much smaller in the composite with filled data (Fig. 9d) when compared to the distribution with missing data (Fig. 9c) and the range of both mean (Fig. 9e) and median errors was smaller for the gap-filled product as can be seen in comparison side-by-side comparisons in Fig. 9e and Fig. 9f, respectively. This comparison demonstrates the added value of gap filling monthly rainfall prior to spatial interpolation of point data. We do acknowledge that the location and magnitude of the error would undoubtedly vary depending on which stations are removed/filled prior to the interpolation and that, despite gap filling, errors are still present in the composite.

4. Summary and future work

The primary objectives of this study were to determine the performance of gap filling daily rainfall in Hawaii for data gaps of various size and to identify the value of gap filling daily rainfall prior to the spatial interpolation of monthly values. First, five gap-filling methods were optimized and then tested on a large (2 yr) data gap. Then, these same five methods and two additional methods were tested to determine the error associated with filling submonthly gaps (between 1 and 31 days missing) by calculating the error associated with aggregating gap-filled daily rainfall to monthly accumulations.

The optimization analysis highlighted the importance of tailoring a gap-filling approach to a specific area of interest, as the density of a station network, topographical setting, and area of influence will be unique for each study site. Optimization of the selected gap-filling approach is a critical step that should not be overlooked. Results in this study suggest that both the correlation and distance between a target and predictor station are important factors in producing a quality estimate. As expected, larger distance or lower correlations correspond to larger errors. More importantly, the

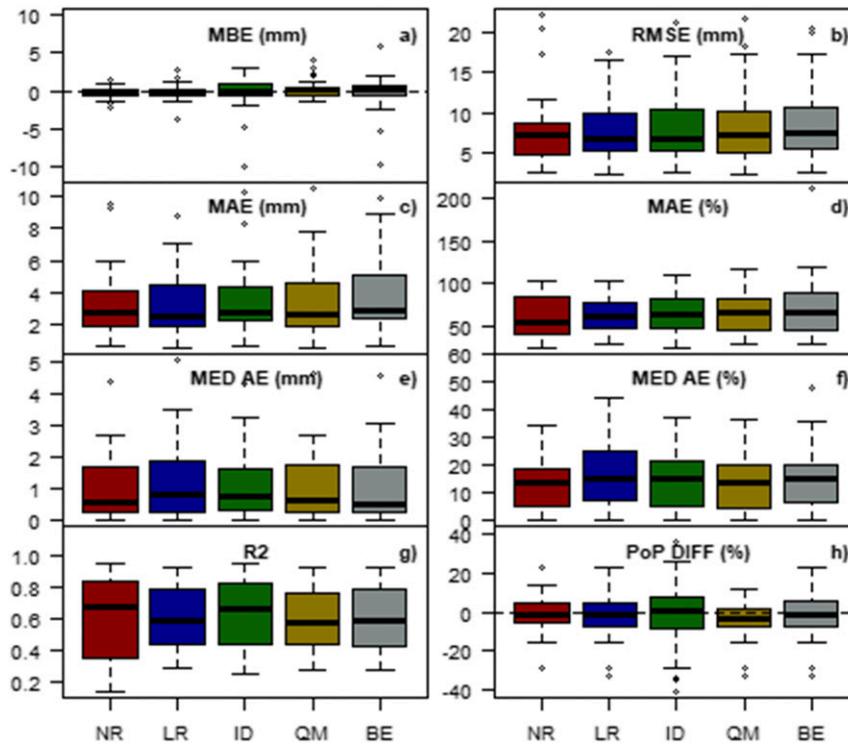


FIG. 5. Gap-filling errors across eight unique metrics—(a) MBE, (b) RMSE, (c) MAE, (d) relative MAE, (e) median absolute error (MED AE), (f) relative MED (MED AE), (g) R^2 , and (h) the difference in the PoP and gap-filled datasets—for five gap-filling methods: NR, LR, ID, QM, and BE. Data points falling outside the Q1–Q3 range are plotted as outliers of the data, and those outliers exceeding the y-axis bounds on the graph are not shown.

establishment of a correlation threshold allows for a decision to be made as to when to choose between different methods. To achieve serially complete data using the methods described in this study, the ID method must be used exclusively or in combination with another method due to the fact that the other methods have minimum statistical requirements that may not be met all of the time. To this end, identifying a correlation threshold is a way to determine when the ID method should be

used in place of another method. This of course is also dependent on the availability of stations used to properly execute the ID method. In some instances, where a network may be less dense than the one utilized in this study, it might be better to select a station with a low correlation as opposed to a station at greater distance or to use a completely different gap-filling method to address the issue. Because of the density of the stations used in this analysis, the area of influence was not a limiting factor.

TABLE 1. Average errors associated with filling a large (~2 yr) gap in the records at 36 rainfall stations. REL MAE is the average relative absolute mean error, REL MED is the average relative absolute median error, D-PoP is the average difference in the probability of rainfall between the observed and the gap-filled data at 36 sample stations, and the other metrics are defined in section 2f. The lowest errors for each metric are shown in boldface font. The methods are defined in section 1 and described in section 2b.

Method	MBE (mm)		MAE (mm)		REL MAE (%)		RMSE (mm)		MED (mm)		REL MED (%)		R^2		D-PoP (%)	
	Mean	Std dev	Mean	Std dev	Mean	Std dev	Mean	Std dev	Mean	Std dev	Mean	Std dev	Mean	Std dev	Mean	Std dev
	NR	-0.31	±0.7	3.2	±2.1	63.4	±24.3	7.8	±4.4	0.9	±1.0	12.5	±8.8	0.61	±0.25	-0.5
LR	-0.19	±1.0	3.3	±2.1	62.2	±20.0	7.8	±4.0	1.2	±1.2	16.5	±12.8	0.61	±0.21	-1.7	±10.3
ID	-0.15	±2.3	3.3	±2.1	65.4	±23.7	8.0	±4.3	1.0	±1.0	14.5	±10.9	0.62	±0.22	-1.4	±16.0
QM	0.21	±1.1	3.4	±2.3	66.3	±24.3	8.5	±4.7	1.0	±1.1	13.2	±10.1	0.61	±0.20	-2.8	±8.9
BE	-0.34	±2.3	3.6	±2.3	71.9	±34.4	8.6	±4.6	1.0	±1.0	14.1	±10.7	0.60	±0.21	-1.3	±10.3

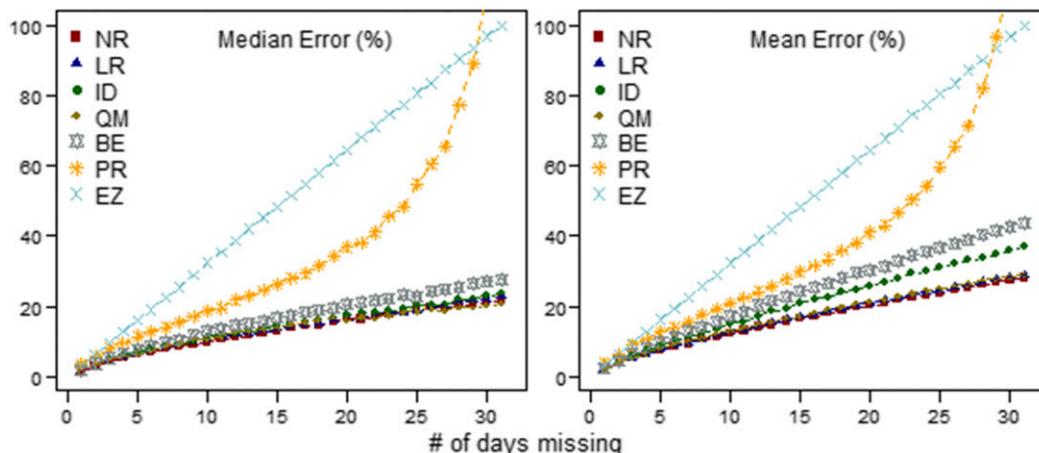


FIG. 6. (left) MED and (right) MAE between observed monthly total rainfall and monthly rainfall, when daily gaps are filled, for between 1 and 31 days missing and for seven gap-filling methods: NR, LR, ID, QM, BE, PR, and EZ.

Another essential optimization approach is the inclusion of a rain/no-rain bias correction. For all of the methods tested, using either the COR or DIST correction significantly reduced the instances where spurious rainfall was added to the gap-filled dataset. The inclusion

of this correction also resulted in a small negative bias in the overall estimates.

For longer (2yr) gaps, all five of the gap-filling approaches produced similar results, and no statistical differences were identified between the error distributions.

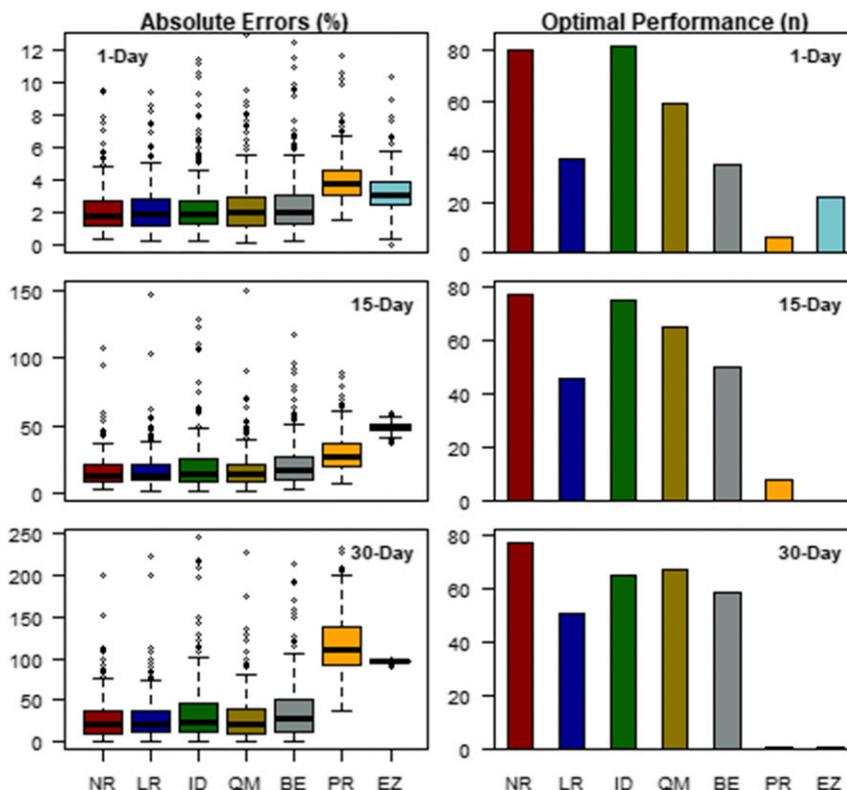


FIG. 7. (left) Absolute errors between observed monthly rainfall sums and monthly sums with daily gaps filled and (right) the frequency in which a given method was optimal for a given month for (top) 1 day of missing data, (middle) 15 days of missing data, and (bottom) 30 days of missing data for seven infilling methods: NR, LR, ID, QM, BE, PR, and EZ. Boxplot data falling outside the Q1–Q3 range are plotted as outliers of the data, and those outliers exceeding the y-axis bounds on the graph are not shown.

TABLE 2. Absolute mean and median (MED) deviations between gap-filled and observed monthly rainfall for seven infilling methods and for three missing data scenarios, along with RMSE. All metrics are percentages. The lowest errors for each metric are shown in boldface font.

	1-day gap			15-day gap			30-day gap		
	MED	Mean	RMSE	MED	Mean	RMSE	MED	Mean	RMSE
NR	2.2	2.8	1.7	17.1	23.3	4.8	25.0	37.5	6.1
LR	2.3	2.8	1.7	18.5	24.7	5.0	28.9	41.7	6.5
ID	2.4	3.6	1.9	20.1	33.9	5.8	29.2	59.2	7.7
QM	2.3	3.3	1.8	18.7	30.4	5.5	27.0	51.0	7.1
BE	2.5	3.7	1.9	21.1	37.7	6.1	31.5	65.7	8.1
PR	4.1	4.3	2.1	30.5	34.0	5.8	121.9	130.4	11.4
EZ	3.1	3.2	1.8	48.6	48.6	7.0	96.9	96.8	9.8

For smaller gaps, no statistical differences were found between these same five methods. However, all of these methods produced errors significantly lower than the PR and EZ methods. These results show that under no circumstances should either of these methods be employed (regardless of the gap size). For our analysis of smaller gaps, the QM method produced the lowest median error when gaps were greater than 15 days. However, mean errors were consistently lowest when using the NR method. Given that the NR approach utilizes multiple predictor stations, this has a smoothing effect on the estimates and can reduce errors associated with large outliers. Overall, the NR method was the most consistent at reproducing monthly rainfall and was consistently the most optimal method overall. The results in this study were similar to those of [Tardivo and Berti \(2013\)](#), who also found that the NR method (using up to three predictor stations) was optimal among several other methods for gap filling daily rainfall in Italy.

One of the most important findings of this study is the added value that gap filling a daily rainfall time series has prior to spatial interpolation of these data. In diverse topographical areas such as Hawaii Island, gap filling critical stations can help to capture complex rainfall gradients that exist there.

The method described in this paper can be readily adapted for use with other datasets at a daily temporal resolution. The most appropriate gap-filling approach for any unique station may vary in both space and time, especially considering that all the methods presented here are dependent on information from predictor stations that may or may not be available on a given day. Also, because many of these methods require established statistical relationships between target and predictor stations, their ability to fill all gaps within a dataset may be limited. Thus, filling a dataset to completion may require the use of multiple approaches.

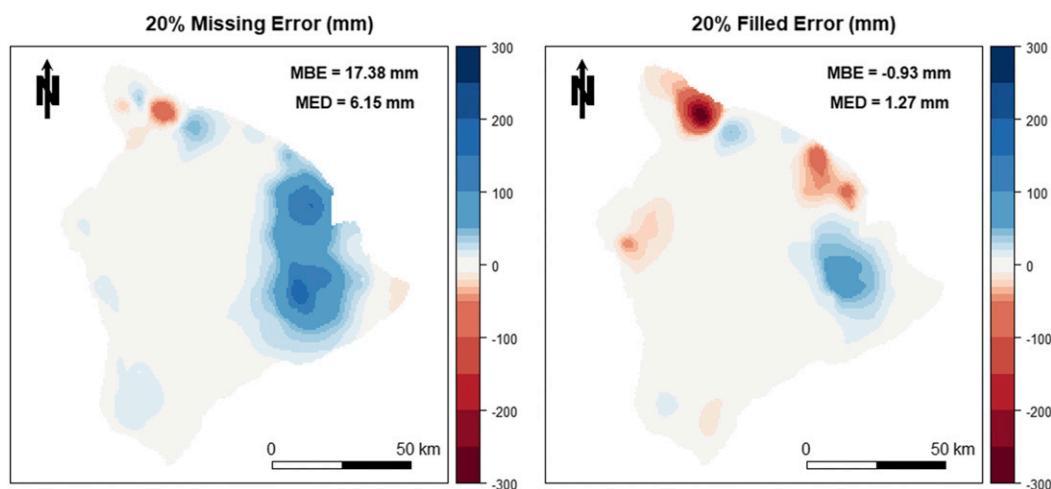


FIG. 8. Fourteen-month composite of deviations between a gridded surface created with monthly observations and a gridded surface created with (a) 20% of the stations removed and (b) 20% of the stations filled, and their associated MBE and MED.

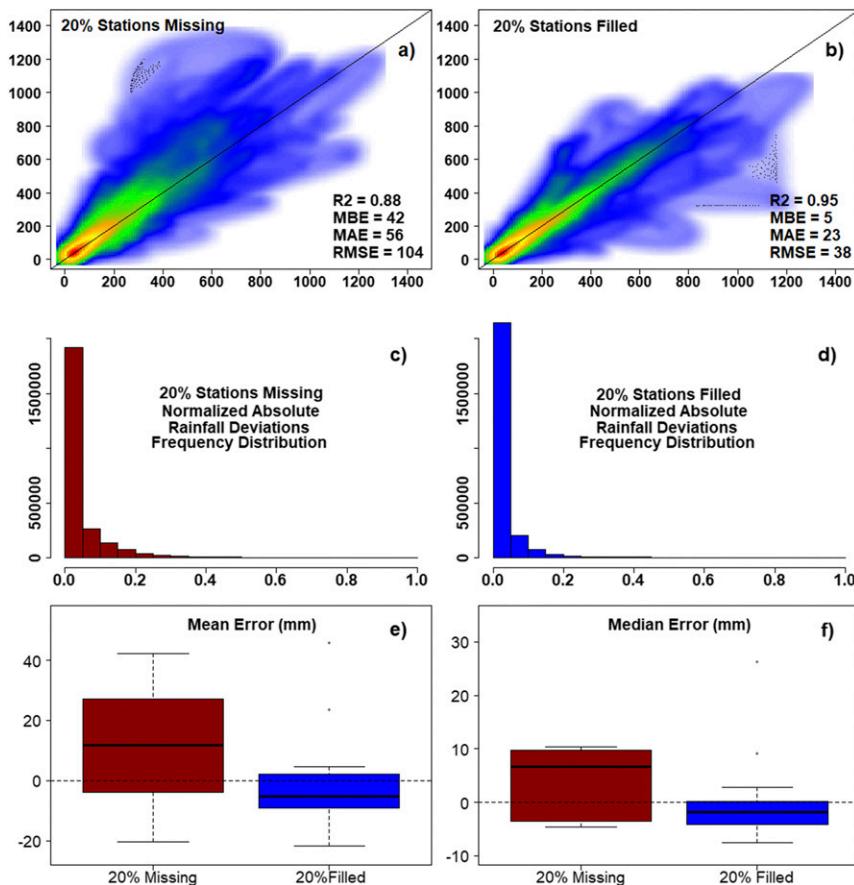


FIG. 9. Fourteen-month composite of gridded estimates of observed rainfall (x axis) compared with gridded estimates made with (a) 20% of the observations missing (RF data are in millimeters) and (b) 20% of the observations filled; frequency distributions of normalized absolute rainfall for gridded estimates made with (c) 20% of the observations missing and (d) 20% of the observations filled; (e) mean and (f) median deviations in rainfall between gridded estimates made with observed rainfall data, and gridded estimates made with 20% of the data missing and 20% of the data filled. In (e) and (f), data falling outside the Q1–Q3 range are plotted as outliers of the data, and those outliers exceeding the y -axis bounds on the graph are not shown.

The results of this study will be used to support an ongoing research effort to produce a serially complete daily rainfall dataset for the State of Hawaii (from 1990 to present day), which is part of a bigger project to develop a near-real-time data acquisition and rainfall mapping system for the State of Hawaii.

Acknowledgments. We thank Abby Frazier, East-West Center, and Mike Nullet, Department of Geography and Environment, University of Hawai‘i at Mānoa. Funding for the Hawaii EPSCoR Program, which supported this work, is provided by the National Science Foundation’s Research Infrastructure Improvement (RII) Track-1: ‘Ike Wai: Securing Hawaii’s Water Future Award OIA-1557349. Andrew Newman was funded by the U.S. Army Corps of Engineers Climate Preparedness and Resilience Program. We also thank the Pacific Island

Climate Adaptation Science Center for continued support. The National Center for Atmospheric Research is a major facility sponsored by the National Science Foundation under Cooperative Agreement 1852977.

REFERENCES

Adhikary, S. K., N. Muttill, and A. G. Yilmaz, 2016: Genetic programming-based ordinary kriging for spatial interpolation of rainfall. *J. Hydrol. Eng.*, **21**, 04015062, [https://doi.org/10.1061/\(asce\)he.1943-5584.0001300](https://doi.org/10.1061/(asce)he.1943-5584.0001300).
 Caldera, H. P. G. M., V. R. P. C. Piyathisse, and K. D. W. Nandalal, 2016: A comparison of methods of estimating missing daily rainfall data. *Eng. J. Inst. Eng. Sri Lanka*, **49**, 1–8, <https://doi.org/10.4038/engineer.v49i4.7232>.
 Eischeid, J. K., C. B. Baker, T. R. Karl, and H. F. Diaz, 1995: The quality control of long-term climatological data using objective data analysis. *J. Appl. Meteor.*, **34**, 2787–2795, [https://doi.org/10.1175/1520-0450\(1995\)034<2787:TQCOLT>2.0.CO;2](https://doi.org/10.1175/1520-0450(1995)034<2787:TQCOLT>2.0.CO;2).

- , P. A. Pasteris, H. F. Diaz, M. S. Plantico, and N. J. Lott, 2000: Creating a serially complete, national daily time series of temperature and precipitation for the western United States. *J. Appl. Meteor.*, **39**, 1580–1591, [https://doi.org/10.1175/1520-0450\(2000\)039<1580:CASCND>2.0.CO;2](https://doi.org/10.1175/1520-0450(2000)039<1580:CASCND>2.0.CO;2).
- Frazier, A. G., T. W. Giambelluca, H. F. Diaz, and H. L. Needham, 2016: Comparison of geostatistical approaches to spatially interpolate month-year rainfall for the Hawaiian Islands. *Int. J. Climatol.*, **36**, 1459–1470, <https://doi.org/10.1002/joc.4437>.
- Gao, Y., C. Merz, G. Lischeid, and M. Schneider, 2018: A review on missing hydrological data processing. *Environ. Earth Sci.*, **77**, 47, <https://doi.org/10.1007/s12665-018-7228-6>.
- Giambelluca, T. W., Q. Chen, A. G. Frazier, J. P. Price, Y. L. Chen, P. S. Chu, J. K. Eischeid, and D. M. Delparte, 2013: Online rainfall atlas of Hawai'i. *Bull. Amer. Meteor. Soc.*, **94**, 313–316, <https://doi.org/10.1175/BAMS-D-11-00228.1>.
- Hasan, M. M., and B. F. W. Croke, 2013: Filling gaps in daily rainfall data: A statistical approach. *20th Int. Congress on Modelling and Simulation*, Adelaide, SA, Australia, Modelling and Simulation Society of Australia and New Zealand, 380–386, <https://www.mssanz.org.au/modsim2013/A9/hasan.pdf>.
- Kagawa-Viviani, A. K., and T. W. Giambelluca, 2020: Spatial patterns and trends in surface air temperatures and implied changes in atmospheric moisture across the Hawaiian Islands, 1905–2017. *J. Geophys. Res. Atmos.*, **125**, e2019JD031571, <https://doi.org/10.1029/2019jd031571>.
- Kim, J., and J. H. Ryu, 2016: A heuristic gap filling method for daily precipitation series. *Water Resour. Manage.*, **30**, 2275–2294, <https://doi.org/10.1007/s11269-016-1284-z>.
- Li, J., and A. D. Heap, 2008: A Review of Spatial Interpolation Methods for Environmental Scientists. Geoscience Australia Record 2008/23, 137 pp., https://d28rz98at9flks.cloudfront.net/68229/Rec2008_023.pdf.
- Longman, R. J., H. F. Diaz, and T. W. Giambelluca, 2015a: Sustained increases in lower-tropospheric subsidence over the central tropical North Pacific drive a decline in high-elevation rainfall in Hawaii. *J. Climate*, **28**, 8743–8759, <https://doi.org/10.1175/JCLI-D-15-0006.1>.
- , T. W. Giambelluca, M. A. Nullet, and L. L. Loope, 2015b: Climatology of Haleakalā. University of Hawai'i at Mānoa Pacific Cooperative Studies Unit Tech. Rep., 126 pp.
- , and Coauthors, 2018: Compilation of climate data from heterogeneous networks across the Hawaiian Islands. *Sci. Data*, **5**, 180012, <https://doi.org/10.1038/sdata.2018.12>.
- , and Coauthors, 2019: High-resolution gridded daily rainfall and temperature for the Hawaiian Islands (1990–2014). *J. Hydrometeorol.*, **20**, 489–508, <https://doi.org/10.1175/JHM-D-18-0112.1>.
- Lu, G. Y., and D. W. Wong, 2008: An adaptive inverse-distance weighting spatial interpolation technique. *Comput. Geosci.*, **34**, 1044–1055, <https://doi.org/10.1016/j.cageo.2007.07.010>.
- McCuen, R. H., 2004: *Hydrologic Analysis and Design*. 3rd ed. Prentice Hall, 888 pp.
- Newman, A. J., and Coauthors, 2015: Gridded ensemble precipitation and temperature estimates for the contiguous United States. *J. Hydrometeorol.*, **16**, 2481–2500, <https://doi.org/10.1175/JHM-D-15-0026.1>.
- Panofsky, H. A., and G. W. Brier, 1968: *Some Applications of Statistics to Meteorology*. The Pennsylvania State University, 224 pp.
- Paulhus, J. L., and M. A. Kohler, 1952: Interpolation of missing precipitation records. *Mon. Wea. Rev.*, **80**, 129–133, [https://doi.org/10.1175/1520-0493\(1952\)080<0129:IOMPR>2.0.CO;2](https://doi.org/10.1175/1520-0493(1952)080<0129:IOMPR>2.0.CO;2).
- Salas, J. D., 1993: Analysis and modeling of hydrological time series. *Handbook of Hydrology*, D. R. Maidment, Ed., Vol. 199, McGraw-Hill, 19.1–19.72.
- Shepard, D., 1964: A two-dimensional interpolation for irregularly-spaced data function. *Proc. ACM National Conf.*, Las Vegas, NV, ACM, 517–524, <https://doi.org/10.1145/800186.810616>.
- Simolo, C., M. Brunetti, M. Maugeri, and T. Nanni, 2010: Improving estimation of missing values in daily precipitation series by a probability density function-preserving approach. *Int. J. Climatol.*, **30**, 1564–1576, <https://doi.org/10.1002/joc.1992>.
- Tannenbaum, C., 2009: The empirical nature and statistical treatment of missing data. Ph.D. dissertation, University of Pennsylvania, <https://repository.upenn.edu/dissertations/AAI3381876>.
- Tardivo, G., 2015: Spatial and time correlation of thermometers and pluviometers in a weather network database. *Theor. Appl. Climatol.*, **120**, 19–28, <https://doi.org/10.1007/s00704-014-1148-5>.
- , and A. Berti, 2013: Comparison of four methods to fill the gaps in daily precipitation data collected by a dense weather network. *Sci. J. Environ. Eng. Res.*, **2013**, 1–9, <https://doi.org/10.7237/sjeer/265>.
- Teegavarapu, R. S. V., 2014: Statistical corrections of spatially interpolated missing precipitation data estimates. *Hydrol. Processes*, **28**, 3789–3808, <https://doi.org/10.1002/hyp.9906>.
- , and V. Chandramouli, 2005: Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records. *J. Hydrol.*, **312**, 191–206, <https://doi.org/10.1016/j.jhydrol.2005.02.015>.
- , M. Tufail, and L. Ormsbee, 2009: Optimal functional forms for estimation of missing precipitation data. *J. Hydrol.*, **374**, 106–115, <https://doi.org/10.1016/j.jhydrol.2009.06.014>.
- , A. Aly, C. S. Pathak, J. Ahlquist, H. Fuelberg, and J. Hood, 2018: Infilling missing precipitation records using variants of spatial interpolation and data-driven methods: Use of optimal weighting parameters and nearest neighbour-based corrections. *Int. J. Climatol.*, **38**, 776–793, <https://doi.org/10.1002/joc.5209>.
- Vicente-Serrano, S. M., S. Beguería, J. I. López-Moreno, M. A. García-Vera, and P. Stepanek, 2009: A complete daily precipitation database for northeast Spain: Reconstruction, quality control, and homogeneity. *Int. J. Climatol.*, **30**, 1146–1163, <https://doi.org/10.1002/joc.1850>.
- Webster, R., and M. A. Oliver, 2001: *Geostatistics for Environmental Scientists*. Statistics in Practice, 2nd ed. John Wiley and Sons, 300 pp.