Bias adjustment of satellite-based precipitation estimation using gauge observations: A case study in Chile

Zhongwen Yang1,2, Kuolin Hsu2, Soroosh Sorooshian2, Xinyi Xu1, Dan Braithwaite2, and Koen M. J. Verbist1,4

1College of Water Sciences, Key Laboratory of Water and Sediment Sciences of Ministry of Education, Beijing Normal University, Beijing, China. 2Department of Civil and Environmental Engineering, University of California, Irvine, California, USA. 3UNESCO-IHP, Hydrological Systems and Global Change Section, Santiago, Chile. 4International Centre for Eremology, Department of Soil Management, Ghent University, Ghent, Belgium

Abstract

Satellite-based precipitation estimates (SPEs) are promising alternative precipitation data for climatic and hydrological applications, especially for regions where ground-based observations are limited. However, existing satellite-based rainfall estimations are subject to systematic biases. This study aims to adjust the biases in the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks–Cloud Classification System (PERSIANN-CCS) rainfall data over Chile, using gauge observations as reference. A novel bias adjustment framework, termed QM-GW, is proposed based on the nonparametric quantile mapping approach and a Gaussian weighting interpolation scheme. The PERSIANN-CCS precipitation estimates (daily, 0.04°×0.04°) over Chile are adjusted for the period of 2009–2014. The historical data (satellite and gauge) for 2009–2013 are used to calibrate the methodology; nonparametric cumulative distribution functions of satellite and gauge observations are estimated at every 1°×1° box region. One year (2014) of gauge data was used for validation. The results show that the biases of the PERSIANN-CCS precipitation data are effectively reduced. The spatial patterns of adjusted satellite rainfall show high consistency to the gauge observations, with reduced root-mean-square errors and mean biases. The systematic biases of the PERSIANN-CCS precipitation time series, at both monthly and daily scales, are removed. The extended validation also verifies that the proposed approach can be applied to adjust SPEs into the future, without further need for ground-based measurements. This study serves as a valuable reference for the bias adjustment of existing SPEs using gauge observations worldwide.

1. Introduction

Precipitation is one key input variable for hydrological process modeling and climatic studies of extreme events, such as floods and droughts. The quality of precipitation estimates can largely influence the inferred outcomes of these applications. It is widely recognized that the ground-based gauge can provide reliable precipitation measurements at gauge points. However, uncertainty from gauges increases when the precipitation measurement is extended from the point scale to a spatial coverage [Huff, 1970]. This limitation of gauge observation can be much worse in semiarid or mountainous regions where the gauge network is generally sparse. Satellite estimates provide an alternative for precipitation measurements for those regions where ground observations are limited or not available. Satellite-based retrieval algorithms take advantage of the visible and infrared spectral information observed from Geostationary Earth Orbiting (GEO) satellites or the passive microwave (PMW) images from Low Earth Orbiting (LEO) satellites to generate rainfall estimates, offering much better coverage than gauge observations [Boushaki et al., 2009]. A range of satellite-based precipitation products have been developed during the past two decades [Adler et al., 2003; Ashouri et al., 2015; Hong et al., 2004; Hsu et al., 1997; Hsu et al., 1999; Huffman et al., 2009; Huffman et al., 2010; Huffman et al., 2007; Huffman et al., 2014; Joyce et al., 2004; Kubota et al., 2007; Kuligowski, 2002; Mitchell et al., 2004; Sorooshian et al., 2000; Xie and Arkin, 1997].

However, without direct reference to ground-based measurements, satellite-based precipitation estimates (SPEs) are subject to systematic bias. The bias of SPEs may come from different sources, including sampling error, sensor limitations, and during estimation from retrieval algorithms. SPEs are interpreted from snapshot images from GEO and LEO satellites, which are different from “true” values. Samples from GEO satellites are within the spectral range of visible and infrared wavelengths providing cloud albedo or cloud top thermal temperatures which are only indirectly related to the rainfall below the clouds. More direct sensing of rainy
clouds can be obtained from low-frequency PMW spectrums; however, samples from PMW sensors are less frequent and limited in spatial coverage compared to those from GEO-based sensors.

To better understand the presence of bias, satellite-based rainfall products have been extensively evaluated in recent years. Li et al. [2013] conducted a multiscale evaluation of four mainstream products over the Yangtze River and found that these products exhibited considerable biases in different temporal scales. Hirpa et al. [2010] and Romilly and Gebremichael [2011] indicated that the PMW-based rainfall products outperformed the infrared-based counterparts and the biases in SPEs depended on different rainfall regimes. Over mountainous regions, studies showed that the SPEs examined could not fully capture the dependence of mean precipitation on elevation [Gao and Liu, 2013; Krakauer et al., 2013]. For extreme rainfall, it was pointed out that no single rainfall product could be considered ideal for detecting extreme events, and the SPEs tended to miss a significant volume of rainfall [Aghakouchak et al., 2011; Miao et al., 2015]. Under these situations, some uncertainty analysis models were also proposed to examine the biases in SPEs [Sarachi et al., 2015; Tian et al., 2009]. Other assessments focus on the utility of SPEs from the perspective of hydrological modeling. Studies found that the biases in SPEs led to systemic overestimation/underestimation in hydrological simulations [Behrangi et al., 2014; Behrangi et al., 2011; Bitew and Gebremichael, 2011; Gebregiorgis et al., 2012; Thiemig et al., 2013]. Those studies concluded that considerable biases exist in SPEs and that bias adjustment is an essential step prior to the hydrological applications using SPEs.

Different bias adjustment approaches were used to improve the data quality of SPEs recently. Among these approaches, the simple scaling method was commonly used [Boushaki et al., 2009; Lin and Wang, 2011; Tesfagiorgis et al., 2011; Vila et al., 2009]. This method calculates the additive or multiplicative bias factors (at daily or monthly) of SPEs against reference data (i.e., gauge observations). The original SPEs are then rescaled according to the bias factors, reducing the biases. But, there are two limitations existing in the simple scaling method. This method cannot work without simultaneous reference data, and it fails to take advantage of historical data that may still contain useful information about the spatial/temporal patterns of rainfall. In contrast, the quantile mapping (QM) approach can successfully avoid these limitations in the bias correction of modeled precipitation data. It is a distribution-based approach, which uses historical data for bias adjustment, and simultaneous reference data are therefore not needed. Specifically, the QM approach is designed to transform the cumulative distribution function (CDF) of modeled data into the CDF of observed climatology at the nearest station [Bennett et al., 2014; Gudmundsson et al., 2012; Themessl et al., 2012]. Some comparison studies indicated that the QM method had the best skill in reducing the systematic bias of regional climate model precipitation estimates [Jie et al., 2013; Themessl et al., 2011]. In particular, the nonparametric QM approach is highly valued for bias adjustment, because it does not rely on any predetermined function and as such provides more flexibility. In addition, from a hydrological modeling perspective, it was found that the sophisticated QM method resulted in better hydrological performance than that of the more simple scaling method [Piani et al., 2010; Thiemig et al., 2013]. Recently, Zhang and Tang [2015] employed the QM approach to adjust satellite precipitation using gauge observation in China and produced reliable precipitation data for hydrological monitoring. In addition, some other researchers have used a similar distribution-based approach (probability density function matching) to adjust SPEs [Sheffield et al., 2014; Shen et al., 2014; Xie and Xiong, 2011]. Based on these, examining the degree of effectiveness of QM approach in adjusting the bias of SPEs is a logical next step.

In this study, the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks–Cloud Classification System (PERSIANN-CCS) [Hong et al., 2004] precipitation product over Chile is adjusted. The objective is to reduce the systematic biases existing in the daily PERSIANN-CCS data set over Chile. A bias-adjusting framework is developed that incorporates the QM approach and Gaussian weighting (GW) interpolation. It is tested using PERSIANN-CCS estimates and gauge observations, as reference data, over Chile. We expect that the results will serve as a good test and reference for the application of the proposed approach for bias correction of satellite-based precipitation products in other regions of the world. This paper is organized as follows: section 2 presents the study area, precipitation data sources, and the construction of methodology; section 3 exhibits the evaluation of the bias adjustment results; section 4 presents the discussion and conclusions; and the future outlook for the method application is given in section 5.
2. Study Area, Data Sources, and Methodology

2.1. Study Area and Data Sources

2.1.1. Study Area

Figure 1 shows the location of the study area, covering the whole territory of Chile, extending from 17°S to 56°S, bounded by the South Pacific Ocean on the west and by the Andes mountain range on the east. The long territory and diverse topography jointly result in quite different climate zones in Chile. Northern Chile has arid/semiarid conditions with extremely scarce precipitation, while in the south, abundant precipitation is observed reaching amounts of up to 6000 mm/yr [Nunez et al., 2011; Smith and Evans, 2007; Verbist et al., 2010]. Temporally, there are four seasons in Chile: fall (March-April-May (MAM)), winter (June-July-August (JJA)), spring (September-October-November (SON)), and summer (December-January-February (DJF)), and most precipitation falls in winter time. Unusually severe droughts have affected the country in the last decade, both because of their intensity and multiannual duration [Garreaud, 2015]. This is consistent with climate change projections for that region that indicate central-southern Chile as a global hot spot for increased drought frequency, causing water security issues in these regions [Prudhomme et al., 2014].
In the high-altitude regions, the impact of floods due to fast rainfall-runoff response and high runoff coefficients results in significant socioeconomic costs [Blume et al., 2008; Favier et al., 2009]. Thus, reliable SPEs with high spatial/temporal resolution are extremely useful for drought analysis and flood prediction over Chile.

2.1.2. Data Sources
The precipitation data used in this study include the PERSIANN-CCS precipitation estimates and the historical gauge observations. The PERSIANN-CCS data are provided from the Center for Hydrometeorology and Remote Sensing, University of California, Irvine, and the gauge observations are retrieved from the Chilean Climate Data Library (website: http://www.climatedatalibrary.cl/SOURCES/Chile/DGA/meteorological/.Precipitation/.Historical/.Daily/). The two data sources are further described as follows:

1. PERSIANN-CCS precipitation estimates
PERSIANN-CCS [Hong et al., 2004] is an infrared-based satellite estimation system. It employs computer image processing and pattern recognition techniques to develop patch-based cloud classification. Based on the cloud classification and artificial neural network model, pixel precipitation intensity is estimated globally. PERSIANN-CCS can produce hourly SPEs at a spatial resolution of 0.04°x0.04°, which is a desirable resolution for hydrological applications at local scale. In this study, the daily original PERSIANN-CCS (Ori-CCS) precipitation data, accumulated from hourly Ori-CCS estimates, are adjusted. Because insufficient samples are available from GEO sensors over part of southern Chile (46°S–56°S), the Ori-CCS estimates are not available every hour for that area. Thus, the study area was limited to the latitude from 17°S to 46°S. To ensure temporal consistency between the data sets, the UTC time of Ori-CCS is adjusted to match Chilean local time that is used for the gauged rainfall data.

2. Historical gauge observations
The gauge observations are used as the reference data in testing the bias adjustment approach. These data were produced by the Chilean National Water Authority, which imposes a strict quality control on the data measurements. Missing and unrealistic, nonphysical data are flagged and checked through a double-mass curve analysis, and data are eliminated from the data series if apparent errors are identified. Over this area (17.5°S to 46°S), there are a total of 456 gauges selected, for which daily precipitation observations have been archived over the 2009–2014 period (Figure 2a). Over this period, on average there is only a 2% missing data rate. Sufficiency of historical ground data is important for constructing reliable transformations for bias adjustment using the QM approach. For the CDF construction of the QM approach the Ori-CCS precipitation data must overlap the observed precipitation records for the same time period.

2.2. Methodology
In this section, the bias adjustment framework, called QM-GW, is presented. It is based on the QM approach and a GW interpolation scheme. The QM approach is used to primarily correct the Ori-CCS precipitation estimates using CDFs calculated from gauge and satellite estimation at every 1°x1° box area. The GW interpolation aims to grid the adjusted PERSIANN-CCS (Adj-CCS) rainfall series at satellite pixels (0.04°x0.04°), based on the bias-corrected estimates.

2.2.1. CDFs of Gauge and Satellite Estimation at 1°x1° Grid Boxes
The nonparametric QM approach is used for mapping the Ori-CCS rainfall estimates to the gauge measurements. The nonparametric QM is constructed with the nonparametric CDFs that are calculated from the Ori-CCS precipitation at gauged satellite pixels and the corresponding gauge observations. In previous related studies, CDFs were estimated individually at each gauged pixel or calculated for every satellite pixel with gridded gauge observations [De Vera and Terra, 2012; Themessl et al., 2011; Thiemig et al., 2013]. However, in this study, the CDFs are calculated for every 1°x1° box region rather than at each gauge point. It is assumed that those gauges within each 1°x1° box share the same CDF. As shown in Figure 2b, the whole study area is divided into 98 1°x1° boxes, of which 31 boxes do not include any gauge station and are marked with heavy border/edge grid boxes.

The CDFs at each box are calculated based on the included gauged pixel Ori-CCS estimates and the concurrent ground observations, which are mixed together in advance, respectively. The method for calculating the nonparametric CDFs follows Wilks [1995]. This method was employed by Themessl et al. [2011] as well. Notably, we calculate four pairs of CDFs at each box using the historical daily rainfall samples from four seasons (MAM, JJA, SON, and DJF), separately (see Figure 2c). This means that two nonparametric CDFs are
calculated for each season at a given CDF box. This is different from previous studies which considered the entire year of data as a whole, such as De Vera and Terra [2012]. These seasonally divided CDFs enable specific transformations in mapping SPEs, providing more reliable adjustment results.

As shown in Figure 2b, for those 1°×1° grid boxes where gauges are not available, the gauge observations and satellite estimations from neighboring 1°×1° grid boxes are collected to estimate the CDF of that 1°×1° grid box. Similarly, if no gauge in the 1°×1° neighborhood is available, the collection of concurrent gauge and satellite measurements is further extended to the closest nearby 1°×1° grid boxes where gauge observations are available.

### 2.2.2. Bias Adjustment of Satellite Estimation

Based on the seasonally calculated nonparametric CDFs over Chile, the Adj-CCS daily rainfall (0.04°×0.04°), \( R_i(t) \), from daily satellite estimation is calculated from daily Ori-CCS rainfall, \( r(t) \), at a given satellite pixel \( i \) at 0.04°×0.04° resolution and time \( t \) within season \( s \) as follows:

\[
R_i(t) = \sum_{j \in \Omega} w_j \cdot CDF_{GCCS}^{-1}(CDFCCS_{j}(r_i(t))) = \sum_{j \in \Omega} w_j \cdot r'_j(t)
\]

(1)

where \( CDF_{GCCS}^{-1} \) is the inverse CDF of gauge observations for season \( s \) at box \( j \); \( CDFCCS_{j} \) denotes the CDF of Ori-CCS precipitation for season \( s \) at the box \( j \). In this study, CDFs from the four 1°×1° boxes near pixel \( i \) (\( j \in \Omega \)) are used to obtain bias-corrected estimates \( r'_j(t) \), based on the nonparametric QM approach. The nearby CDF boxes are identified with the distances from pixel \( i \) to the centers of those CDF boxes (see Figure 2c).
weighting factor \( w_{ij} \) of each \( r_{ij}'(t) \) is estimated based on the distance of satellite pixel \( i \) at 0.04° resolution to the center of box \( j \). A Gaussian function (see Figure 2d) is proposed to generate interpolation weights:

\[
    w_{ij} = \exp\left(-\frac{(\bar{d}_{ij}/C)^2}{2}\right) \sum_{j=1}^{4} \exp\left(-\frac{(\bar{d}_{ij}/C)^2}{2}\right)
\]

\[
    \bar{d}_{ij} = d_{ij}/D
\]

where \( w_{ij} \) is the weighting factor assigned for \( r_{ij}'(t) \); \( \bar{d}_{ij} \) represents the relative distance from pixel \( i \) to the center of nearby 1°×1° box \( j \); \( d_{ij} \) is the distance from pixel \( i \) to the center of nearby CDF box \( j \); \( D \) is a constant, which equals the distance between the two 1°×1° grid centers that are farthest apart; and \( C \) is the shape parameter of the Gaussian function. \( C \) is assigned a value of 0.33 in this study (see Figure 2d).

The flowchart of the bias adjustment is listed in Figure 3. The proposed bias adjustment methodology is evaluated over the time period of 2009–2014. From that data, 5 years of data (2009–2013) are used for the estimation of seasonal CDFs at the 98 boxes, with 1 year data (2014) used to verify the bias-adjusted CCS estimation using the CDFs estimated from the 5 year calibration period.

**3. Evaluation**

The results from the calibration and validation are evaluated using gauge observations as reference. In detail, the performance of the Ori-CCS and Adj-CCS data sets are assessed in two aspects, including the spatial pattern and temporal distribution. Due to the long latitudinal extent of Chile’s territory, the annual precipitation shows large variation from north to south (Figure 4). The spatial pattern evaluation aims to indicate spatial improvement of the satellite-based precipitation estimates over the adjusting area, while the temporal scale evaluation presents the ability of the bias adjustment of satellite-based precipitation time series at specified regions that are relevant to the potential hydrological application of SPEs at local scale.

Figure 4 shows spatial distribution of average annual precipitation from gauge observations for the period of 2009 to 2014 over the study area. It shows low-precipitation rates apparent over northern Chile, with high precipitation observed over southern Chile. In our evaluation, three evaluation zones, representing low-, medium-, and high-precipitation regions, are selected for the evaluation. Each zone is an area covering a 2° latitude extent. According to Figure 4, the mean areal precipitation for annual average over zone nos. 1, 2, and 3 are calculated to be 173, 1056, and 2027 mm/yr, respectively. To illustrate the systematic errors of Ori-CCS to be adjusted, seasonal CDFs from Ori-CCS and gauge areal mean daily estimates over
the zones are derived and shown in Figure 5. It is presented that the CDFs of the two data sources for zone no. 2 are relatively close to each other during the seasons, indicating less systematic bias in Ori-CCS estimation over zone no. 2. However, Ori-CCS trends to systematically overestimate and underestimate daily precipitation over zone nos. 1 and 3, respectively, according to the corresponding seasonal CDFs.

The performance of the Ori-CCS and Adj-CCS precipitation estimates are evaluated based on three statistics: the correlation coefficient (CORR), the root-mean-square error (RMSE), and the mean bias (BIAS). They are calculated in the spatial and temporal evaluations with equations 4–6:

\[
\text{CORR} = \frac{\sum_{k=1}^{n} (G_k - \bar{G})(S_k - \bar{S})}{\sqrt{\sum_{k=1}^{n} (G_k - \bar{G})^2} \sqrt{\sum_{k=1}^{n} (S_k - \bar{S})^2}}
\]

\[
\text{RMSE} = \sqrt{\frac{1}{n} \sum_{k=1}^{n} (S_k - G_k)^2}
\]

\[
\text{BIAS} = \frac{1}{n} \sum_{k=1}^{n} (S_k - G_k)
\]

where \(G_k\) are the gauge observations and \(\bar{G}\) is the average of the gauge observations; \(S_k\) and \(\bar{S}\) are the satellite estimates (i.e., Ori-CCS or Adj-CCS) and their average, respectively. The performance improvement of the Adj-CCS estimations can be indicated by increased CORR and by reduced RMSE and absolute BIAS.

### 3.1. Spatial Pattern Evaluation

Figure 6 shows the annual average precipitation for the gauge observations, Ori-CCS, and Adj-CCS, during the calibration years (2009–2013) and validation year (2014). Compared to the gauge observations, Ori-CCS overestimated precipitation over northern Chile and underestimated over southern Chile. On the other hand, the Adj-CCS has consistently improved upon Ori-CCS for both calibration and validation periods (Figure 6). Locally, all of the evaluation zones have presented consistent improvement as well. As shown in Figure 7, the scatterplots of the observed annual precipitation at gauged pixels with respect to the Ori-CCS and Adj-CCS are derived for zone nos. 1, 2, and 3, respectively. Statistically, the spatial pattern improvement over the zones is indicated by the increased CORRs and the great reduction of RMSEs (mm/yr) and absolute BIASs (mm/yr), respectively (see Table 1).
Figure 5. Seasonal CDFs of areal mean daily gauge and Ori-CCS estimations for the evaluation zones.

Figure 6. Annual precipitation of gauge observation, Ori-CC, and Adj-CCS in (left) calibration and (right) validation years. The evaluation zone nos. 1, 2, and 3 are highlighted with rectangles from top to bottom.
In addition, the monthly spatial patterns of the Ori-CCS and Adj-CCS precipitation are assessed. Figures 8 and 9 show the evaluation results from January, April, July, and October. The monthly precipitation averages of the four months during calibration years exhibit similar spatial pattern changes to those of the annual results. Over the evaluation zones, the monthly data of Ori-CCS overestimated monthly average precipitation in zone no. 1 and underestimated in zone nos. 2 and 3. As shown in Table 1, zone no. 3 has averagely the lowest CORR (0.3) and highest RMSE (94.2 mm/month) and absolute BIAS (89.9 mm/month) with respect to the Ori-CCS monthly average. After bias adjustment, the spatial disagreements have been significantly reduced.

The Adj-CCS shows advanced spatial patterns of monthly average precipitation as compared to that of the gauge observations (Figures 8 and 9). From Table 1, almost all of the CORRs of Adj-CCS for the zones are increased, and the RMSEs are reduced on average by 90%, 45%, and 64% for zone nos. 1, 2, and 3, respectively. The absolute BIASs for the months decrease greatly to lower than 1.0 mm/month after bias adjustment, except for April over zone no. 3 (decrease to 28.82 mm/month) and July over zone no. 2 (decrease to 15.86 mm/month).

For the validation year (2014), the monthly Adj-CCS estimates show improved precipitation spatial patterns for the four evaluation months, as well. From Figures 8 and 9, the monthly Ori-CCS precipitation shows overestimation in zone no. 1 and underestimation in zone nos. 2 and 3 for those four months. The CORRs of the Ori-CCS monthly precipitation in 2014 are low, with some negative CORRs for zone nos. 2 and 3 (see Table 1). The RMSEs and absolute BIASs for zone no. 3 are the highest with the averages of 136.7 mm/month and 123.6 mm/month, respectively. After bias adjustment, each of the four months of Adj-CCS precipitation in 2014 shows distinct spatial pattern improvement, except for zone no. 2 which presents limited changes for January, April, and October (Figures 8 and 9). Compared to Ori-CCS, Adj-CCS shows reduced RMSEs by 85%, 29%, and 48% on average for zone nos. 1, 2, and 3, respectively (see Table 1). The absolute BIASs for

Table 1. Statistics for Spatial Pattern Evaluation

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Zone No. 1</td>
<td>Zone No. 2</td>
<td>Zone No. 3</td>
</tr>
<tr>
<td>Correlation</td>
<td>Ori-CCS</td>
<td>Adj-CCS</td>
<td>Ori-CCS</td>
</tr>
<tr>
<td>Annual</td>
<td>0.91</td>
<td>0.93</td>
<td>0.82</td>
</tr>
<tr>
<td>RMSE</td>
<td>676.91</td>
<td>36.83</td>
<td>393.17</td>
</tr>
<tr>
<td>BIAS</td>
<td>618.10</td>
<td>0.31</td>
<td>365.92</td>
</tr>
<tr>
<td>January</td>
<td>0.91</td>
<td>0.94</td>
<td>0.60</td>
</tr>
<tr>
<td>Correlation</td>
<td>123.29</td>
<td>12.69</td>
<td>11.39</td>
</tr>
<tr>
<td>RMSE</td>
<td>133.29</td>
<td>12.69</td>
<td>11.39</td>
</tr>
<tr>
<td>BIAS</td>
<td>121.28</td>
<td>7.96</td>
<td>10.11</td>
</tr>
<tr>
<td>April</td>
<td>0.83</td>
<td>0.72</td>
<td>0.62</td>
</tr>
<tr>
<td>Correlation</td>
<td>21.70</td>
<td>26.1</td>
<td>17.72</td>
</tr>
<tr>
<td>RMSE</td>
<td>14.25</td>
<td>3.17</td>
<td>84.60</td>
</tr>
<tr>
<td>BIAS</td>
<td>17.94</td>
<td>1.50</td>
<td>14.59</td>
</tr>
<tr>
<td>July</td>
<td>0.11</td>
<td>0.24</td>
<td>0.74</td>
</tr>
<tr>
<td>Correlation</td>
<td>31.11</td>
<td>1.00</td>
<td>17.32</td>
</tr>
<tr>
<td>RMSE</td>
<td>31.34</td>
<td>0.48</td>
<td>9.94</td>
</tr>
</tbody>
</table>

Figure 7. Scatterplot of the annual precipitation of Ori-CC and Adj-CCS at gauged pixels over the evaluation zones for (left three plots) calibration and (right three plots) validation years.
Figure 8. (top to bottom) The same as Figure 6 but for monthly precipitation of January, April, July, and October, respectively.
the months in 2014 decrease on average by 95%, 65%, and 81%, correspondingly, after bias adjustment. However, these changes are not significant as those from the average results during the calibration years (Figure 9 and Table 1).

3.2. Time Series Evaluation

3.2.1. Monthly Time Series

In the monthly precipitation time series evaluation, plots of gauge observation (red), Ori-CCS (green), and Adj-CCS (blue) over zone nos. 1, 2, and 3 for the calibration and validation years are shown in Figure 10. Compared to the reference gauge observations, the monthly area mean precipitation of Ori-CCS shows systematic biases over the three zones. For zone no. 1, consistent overestimation of monthly precipitation by Ori-CCS is observed during 2009–2014; the RMSEs and BIASs are over 38.5 and 23.9 mm/month, respectively. As for zone nos. 2 and 3, Ori-CCS underestimates the areal mean monthly precipitation; the RMSEs are over 71.0 mm/month with significant negative BIASs of −125.2 mm/month. The CORRs at these zones, however, are very high (close to or exceeding 0.90) during the calibration and validation periods. These high CORRs indicate that the temporal structures of the monthly Ori-CCS estimates are highly consistent with that of gauge observations.

After the adjustment, the Adj-CCS precipitation series present great improvement in data quality. During 2009–2014, the Adj-CCS monthly area mean of precipitation over zone no. 1 is reduced, and over zone nos. 2 and 3 are increased (Figure 10). These changes lead to good agreement of Adj-CCS monthly precipitation estimates to gauge observations. As a result, the RMSEs and absolute BIASs of monthly precipitation during the calibration period decrease on average by 47% and 86%, respectively, and for the validation period the RMSEs and absolute BIASs are reduced by 69% and 90%, respectively. The CORRs of the monthly area mean precipitation stay almost the same during both periods.
### 3.2.2. Daily Time Series

The overall evaluation is also extended to the daily scale. The CORRs, RMSEs, and BIASs of Ori-CCS and Adj-CCS areal mean daily estimates over each zone are calculated with respect to calibration and validation periods. In order to evaluate the hits and misses of satellite estimation, we further estimate the probability of detection (POD), false alarm ratio (FAR), and Heidke skill score (HSS) for these daily series. The estimations of POD, FAR, and HSS follow Su et al. [2011] and Hyvarinen [2014]. All of the statistics are shown as Table 2.

According to Table 2, the RMSEs of Ori-CCS daily areal mean estimates over the zones range from 2.31 to 9.71 mm/d, and the absolute BIASs range from 4.76 to 1.13 mm/d, with the CORRs of about 0.70. For PODs, they are high (about 0.90) over zone no. 1 for the calibration and validation years and are lower (0.71–0.81) over zone nos. 2 and 3. The FARs and HSSs tend to decrease over zone nos. 1 to 3, with higher HSSs (reaching 0.42) for zone no. 1 and low FARs (to 0.06) for zone no. 3. After bias adjustment, biases are reduced significantly in the daily precipitation series for 2009–2014. It shows that the absolute BIASs of the Adj-CCS estimates over the three zones decrease greatly to less than 0.60 mm/d. Zone no. 1 presents largely reduced RMSEs for the evaluating periods (by over 65%) due to bias adjustment, while zone nos. 2 and 3 show limited improvement in RMSEs. Furthermore, the CORRs, PODs, FARs, and HSSs remain almost unchanged after the correction, except for zone no. 1 which presents slightly decreased PODs and FARs and increased HSSs (Table 2).

Figure 11 shows the year-by-year accumulated daily precipitation time series from the gauge observations (red), original satellite estimations (green), and bias-adjusted estimations (blue) for the period of 2009–2014 for the three zones. Ori-CCS overestimates over zone no. 1 and underestimates over zone nos. 2 and 3. After the adjustment, the cumulative precipitation time series of Adj-CCS are consistent with those of gauge observations. This indicates that the proposed QM approach is effective for adjusting the systematic bias of SPEs for both calibration and validation years over different regions.

Additionally, the number of rainy days (NRDs), RMSE, and BIAS with respect to Ori-CCS and Adj-CCS daily estimations for different precipitation amount ranges are analyzed over the zones. As shown in Figure 12, zone no. 1 has a high NRD from Ori-CCS (green bar) for the nonrainy days (indicating false alarms), and the bias...
adjustment has reduced the NRD as well as that from Ori-CCS for the rainy days with observed rainfall rate of less than 1 mm/d. This is the main reason why the FARs and PODs for zone no. 1 decrease (Table 2). For zone nos. 2 and 3, the top NRDs of Ori-CCS are estimated for the wet days with rainfall amount of less than 10 mm/d, although the Ori-CCS missed a large number of rainfall events compared to the NRDs estimated from gauge observations (red bars). However, all of the NRDs from Ori-CCS stay almost unchanged for the two zones after bias adjustment, which helps to explain the stable PODs, FARs, and HSSs for zone nos. 2

<table>
<thead>
<tr>
<th>Evaluation Zone</th>
<th>Statistic</th>
<th>Ori-CCS</th>
<th>Adj-CCS</th>
<th>Ori-CCS</th>
<th>Adj-CCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zone no. 1</td>
<td>CORR</td>
<td>0.72</td>
<td>0.67</td>
<td>0.78</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>3.14</td>
<td>0.93</td>
<td>2.31</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>BIAS</td>
<td>1.13</td>
<td>−0.15</td>
<td>0.85</td>
<td>−0.06</td>
</tr>
<tr>
<td></td>
<td>POD</td>
<td>0.90</td>
<td>0.84</td>
<td>0.86</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>POD</td>
<td>0.45</td>
<td>0.36</td>
<td>0.48</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>HSS</td>
<td>0.42</td>
<td>0.53</td>
<td>0.31</td>
<td>0.38</td>
</tr>
<tr>
<td>Zone no. 2</td>
<td>CORR</td>
<td>0.67</td>
<td>0.68</td>
<td>0.66</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>5.67</td>
<td>5.87</td>
<td>8.19</td>
<td>8.01</td>
</tr>
<tr>
<td></td>
<td>BIAS</td>
<td>−0.96</td>
<td>0.15</td>
<td>−2.23</td>
<td>−0.60</td>
</tr>
<tr>
<td></td>
<td>POD</td>
<td>0.72</td>
<td>0.71</td>
<td>0.71</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>POD</td>
<td>0.26</td>
<td>0.25</td>
<td>0.21</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>HSS</td>
<td>0.30</td>
<td>0.30</td>
<td>0.20</td>
<td>0.19</td>
</tr>
<tr>
<td>Zone no. 3</td>
<td>CORR</td>
<td>0.65</td>
<td>0.70</td>
<td>0.75</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>7.77</td>
<td>6.97</td>
<td>9.71</td>
<td>7.95</td>
</tr>
<tr>
<td></td>
<td>BIAS</td>
<td>−3.35</td>
<td>−0.23</td>
<td>−4.76</td>
<td>−0.54</td>
</tr>
<tr>
<td></td>
<td>POD</td>
<td>0.74</td>
<td>0.74</td>
<td>0.81</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>POD</td>
<td>0.06</td>
<td>0.06</td>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>HSS</td>
<td>0.20</td>
<td>0.20</td>
<td>0.24</td>
<td>0.24</td>
</tr>
</tbody>
</table>

Figure 11. Cumulative precipitation of gauge observation (red), Ori-CCS (green), and Adj-CCS (blue) over the evaluation zones for 2009–2014.
and 3. Zone no. 1 shows consistent reductions in the RMSEs and absolute BIASs for Adj-CCS as the observed rainfall amount increases. In comparison, the bias adjustment has decreased RMSEs of satellite estimation particularly for the heavy precipitation (rainfall rate $>30$ mm/d) for both zone nos. 2 and 3. Similarly, the absolute BIASs have been largely reduced as well.

4. Discussion and Conclusions

The QM-GW method is proposed to adjust the bias of PERSIANN-CCS using localized gauge observation over Chile. Our experiments show that the spatial patterns of gauge-adjusted estimations (Adj-CCS) over Chile present good agreement with the gauge observations over both calibration and validation periods, especially at annual scale (Figures 6–9 and Table 1). The adjusted precipitation time series have shown considerably reduced biases (Figures 10–12 and Table 2). In the dry north (as zone no. 1), the NRD gets improved for the no rain periods, and significant decreases in RMSEs and BIASs are shown. In the rainy south (as zone nos. 2 and 3), the biases in heavy (extreme) precipitation have been reduced to a large extent. These improvements can support the potential application of PERSIANN-CCS precipitation estimates in flood and drought analysis as well as water resource management in Chile. This is especially the case for those areas with limited or lacking gauge coverage (Figure 2). For these regions the bias-adjusted PERSIANN-CCS can provide reliable precipitation estimates.

From the results of the validation study, one can conclude that the QM-GW bias adjustment approach is capable of removing systematic bias in future SPEs. This is demonstrated in Figures 10 and 11; the proposed
Acknowledgments
This study is sponsored by the National Technology Support Program Funds for the Key Technology for Digital Basin (project 2013BAB05804) and by the China Scholarship Council (201406040177). The support of U.S. Army corps of Engineering’s ICIWaRM program and Will Logan’s feedback in this study are greatly acknowledged. The constructive comments and suggestions by the journal referees are gratefully acknowledged. The gauge and PERSIANN-CCS precipitation data used in this study are available at http://www.climatedatalibrary.cl/SOURCES/Chile/, respectively.

5. Future Prospect
According to our findings, the QM-GW approach has resulted in effective reduction of the systematic biases in SPEs by taking advantage of the distribution properties of historical satellite and gauge data. It can be projected that the framework is able to be applied in other regions to improve satellite-based precipitation products. There are some requirements that need to be considered in future applications.

First, a relatively dense gauge network is required. According to the proposed framework, the study area has been divided into 1°x1° grid boxes for CDF calculation. If the gauge network is sparse, most of the boxes may not include even one rain gauge for reference. This will highly reduce the reliability of the calibrated model and bias-adjusted data. Thus, evenly distributed rain gauges with relatively high network density are recommended. Second, enough high-quality historical precipitation data are needed for model calibration. Specifically, the framework works based on the precipitation CDFs, calculated using satellite and gauge data for each season. In future applications, the use of monthly CDFs will further improve the bias adjustment. While the total amount and data accuracy of considered rainfall samples can impact the uncertainty of CDFs, a larger number of high-quality rainfall data will result in more reliable CDFs for bias adjustment. Last but not least, due to the stationary assumption of the CDFs, model recalibration using the precipitation series is effective in correcting the consistent biases in SPEs without simultaneous gauge observations. This property of the bias adjustment framework is related to the distribution-based QM method. The QM method maps adjusted SPEs based on the transformations from the CDFs of historical SPEs and gauge observations. Under stationary assumption for the CDFs from the historical events being usable for future events, the same CDFs are used for future (validation) events. However, care should be taken in applying the same QM farther into the future, as CDFs may be altered by the impacts of climate variability and change. The nonstationary process of future climate poses a great challenge as well as an opportunity for follow-up studies.

However, some random errors (including misses and false alarms) still exist in the Adj-CCS precipitation series at the local time steps of 2009–2014. These random errors cause very limited improvement of the CORRs, RMSEs, PODs, FARs, HSSs, and NRDs in daily precipitation time series, especially at zone nos. 2 and 3 (Table 2 and Figure 12). It indicates that the QM-GW framework cannot effectively remove the day-by-day random errors in PERSIANN-CCS estimates. This issue can be explained by the probability distribution property of the QM approach. According to the QM approach, the original satellite rainfall rate is generally corrected to a historical rainfall rate from gauge observations, in which the two rainfall rates share the same quantile from the CDFs at a neighboring CDF box. It means that the QM correction takes only the probability distribution of precipitation into consideration but neglects day-by-day adjustment with simultaneous observations. Therefore, the QM approach can help to correct the biases in the overall magnitude of precipitation rather than to reduce the local time random errors. Particularly, the misses and false alarms for satellite estimation can be rarely corrected using the model. The random errors might change with the year-to-year fluctuation in climatology. These can also help to explain why the model tends to lose effectiveness at time scales less than or equal to a month shown especially in the monthly spatial pattern evaluation (Figure 9 and Table 1). The limitation of QM method was also acknowledged previously [Jie et al., 2013; Mueller and Thompson, 2013]. However, we believe that if the simultaneous ground observations are used as reference in further adjustment, the random errors can be largely reduced. In our future work, efforts will focus on further improvement of the Adj-CCS precipitation data quality by correcting the random errors.

References


