Error Analysis on PERSIANN Precipitation Estimations: Case Study of Urmia Lake Basin, Iran

N. Ghajarnia1; P. Daneshkar Arasteh2; M. Liaghat3; and S. Araghinejad4

Abstract: In-depth evaluation and analysis of the error properties associated with satellite-based precipitation estimation algorithms can play an important role in the future development and improvements of these products. This study evaluates the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) daily data set from 2000 to 2011 in 69 pixels over a semiarid basin in northwest Iran and compares it with the data set of the existing rain-gauge network. Different analytical approaches and measures are used to examine PERSIANN performance seasonally and categorically. The residuals are also decomposed into true positive (hit), false negative (miss), and false alarm (FA) estimate biases in addition to systematic and random error components. The results show seasonal variability of PERSIANN precision in rainfall detection with substantial errors during winter and summer that are associated with high rates of FA ratio (more than 60%). The value of miss and FA biases (124 and −77,000 mm, respectively, within the total data set) are considerably larger than hit and total bias (27 and 74,000 mm, respectively) because these components contribute conversely and compensate each other by their opposite signs. Moreover, PERSIANN detects heavy rainfalls well with a probability of detection (POD) over 80%, but with serious biases. Generally, although the detection ability of PERSIANN improves as the rate of rainfall increases, its systematic error in simulation of the rainfall process also increases (from 5% systematic error to 90% in heavier rainfalls), leading to a low level of accuracy in the estimation of precipitation rate. DOI: 10.1061/(ASCE)HE.1943-5584.0001643. © 2018 American Society of Civil Engineers.

Author keywords: PERSIANN; Error components; Systematic/random error; Satellite-based precipitation estimation; Urmia Lake Basin.

Introduction

Recent years have witnessed an increasing tendency to incorporate many global hydrometeorological data sets into hydrological studies. Among these global data sets are the precipitation estimation products which are mostly developed based on global satellite information. The concept behind most of these algorithms is to benefit from both infrared (IR) and passive microwave (PMW) data as inputs to detect occurrence and estimate magnitude of precipitation. Infrared data are available from geostationary (GEO) satellites and have high spatial and temporal resolutions but do not measure precipitation directly. In fact, IR sensors measure cloud albedo and top temperature that can be associated with precipitation rate using an indirect relationship (Nasrollahi et al. 2013). On the other hand, the PMW sensors mounted on low Earth orbiting (LEO) satellites that measure the thermal emission and scattering of raindrops are recognized as more-reliable sources of precipitation remote sensing (Alder et al. 2001; Ebert et al. 1996). However, because LEO satellites have a low temporal resolution of only one or two times per day for a specific location (Marzano et al. 2004), many satellite-derived precipitation products take advantage of multiple remote-sensing devices.

In general, by using different technical approaches, higher temporal resolution of IR data and the more accurate information from PMW data are combined in various precipitation estimation models such as Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA) (Huffman et al. 2007), National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC) Morphing Technique (CMORPH) (Joyce et al. 2004), and Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Network (PERSIANN; Hsu et al. 1997; Sorooshian et al. 2000). Because of the potential for wide application of these global satellite-based precipitation estimation products, efforts have been dedicated to evaluating their quality and analyzing their associated estimation accuracy. Therefore many researchers have focused on evaluation and comparison of high-resolution precipitation estimation products (HRPEP) in different regions (e.g., Alder et al. 2001; Ebert et al. 2007; Tian et al. 2007, 2009; Turk et al. 2008; Habib et al. 2009; Javanmard et al. 2010; AghaKouchak et al. 2011; Tang and Hossain 2012; Jamandre and Narisma 2013; Katariea et al. 2013; Mozamini et al. 2013; Nastos et al. 2013; Turk and Xian 2013; Lo-Conti et al. 2014; Porcu et al. 2014; Ghajarnia et al. 2015; Liu 2015). Such investigations and studies can give better guidance to users in selecting a product for their particular purpose and can help them assess the impact of those errors on the applications (Tian et al. 2009). The pioneers of these kinds of studies are the Algorithm Inter-Comparison Program (AIP) (Arkin and Xie 1994; Ebert et al. 1996), the Precipitation Inter-Comparison Project (PIP) (Smith et al. 1998; Adler et al. 2001), and the more recent Program to Evaluate High-Resolution Precipitation Products (PEHRPP) (Arkin and Turk 2006).

During the past 15 years, numerous studies have been carried out to compare and evaluate these products in different regions.
This paper considered the Urmia Lake Basin in northwestern Iran (Hughes 2006; Dinku et al. 2008; Hirpa et al. 2010); the South American region, Ecuador, Chile, and the United States in the Americas (de Gonçalves et al. 2006; Ward et al. 2011; Ebert et al. 2007; Hong et al. 2007; Sapiano and Arkin 2009; Tian et al. 2009; Habib et al. 2009; AghaKouchak et al. 2011); Australia, China, South Korea, Japan, and Indonesia in the Asia Pacific region (Ebert et al. 2007; Zhou et al. 2008; Sohn et al. 2010; Yatagai et al. 2012; Vernimmen et al. 2012); the Middle East region, including Iran (Javamard et al. 2010; Moazami et al. 2013; Kaitiraei et al. 2013; Moazami et al. 2014; Ghajarnia et al. 2015), and northern Europe, northwestern Europe, and Italy in the European region (Ebert et al. 2007; Kidd et al. 2012; Lo-Conti et al. 2014).

Altogether, the results of these studies show that (1) none of the HRPEPs can completely outperform the others and the performances of these products vary in different regions and countries; (2) generally, the ability of these products to detect rainfall occurrences takes their accuracy in estimation of precipitation magnitude; and (3) these models mainly tend to underestimate heavy rainfall and overestimate light events.

Because of the complexity and wide extent of the concepts incorporated in the field of satellite-based precipitation estimation, Sorooshian et al. (2011), after The Advanced Concepts Workshop on Remote Sensing of Precipitation at Multiple Scales, stated that “future improvements in satellite-based precipitation retrieval algorithms will rely on more in-depth research on error properties in different climate regions, storm regimes, surface conditions, seasons, and altitudes. Given such information, precipitation algorithms for retrieval, downscaling, and data fusion can be optimized for different situations” (Sorooshian et al. 2011). Therefore, although the necessity for evaluation studies originates from the spatial heterogeneity of different products’ performances, linked to various factors such as climate, geographical location, local rain-gauge networks, and so on, the usefulness of new research in this field must be evaluated considering the real insights provided by the study, in particular those that can contribute to the further improvement of such products.

Therefore this study focuses on error decomposition and in-depth analysis of PERSIANN product over the Urmia Lake Basin through seasonal statistics and tries to introduce a collection of statistical and conceptual methods and measures that can potentially provide all essential information for a comprehensive evaluation of HRPEPs. This set of evaluation metrics can be calculated for different products and in different regions to evaluate almost every aspect of HRPEPs’ performances.

Precipitation Products and Case Study

Case Study

This paper considered the Urmia Lake Basin in northwestern Iran (44–47° latitude and 35–38° longitude) as a case study. This basin contains a famous salty lake, Urmia Lake (Fig. 1) (Ghajarnia et al. 2015) which has recently suffered greatly from consequences of climate change and unsustainable developments in the upstream area. The total area of the Urmia basin (51,862 km²) includes of mountainous area (34,100 km²), foothills (13,123 km²), and flat areas (4,639 km²) in addition to Urmia Lake itself.

The climate of the Urmia basin is influenced by cold winters and modest summers. The overall precipitation received in the basin from 1951 to 2012 has decreased in two separate stages, a stationary stage (1951–1994), with 385 mm/year, and a decreasing stage (1994–2012) with 317 mm/year (Urmia Lake Rescue Program 2013).

Data Sets

Reference Rain-Gauge Network

The reference data set used in this study consisted of the rain-gauge network of the Islamic Republic of Iran Meteorological Organization (IRIMO) and the Iran Water Management Research Institute (TAMAB). Overall, 249 rain gauges measure daily precipitation reaching the surface. This study obtained daily rainfall measurements between 2000 and 2011. By applying the typical homogeneity and outlier tests, 218 of these sets were verified to be reliable; Fig. 1 shows the rain-gauge distribution. The errors and uncertainties of the reference data set were analyzed, and based on the results, approximately 85% of the data passed the quality-control stage. The inherent existing uncertainties within the reference data set after aggregation were also calculated and judged to be less than an acceptable level of error, which makes the rest of the study reliable; Ghajarnia et al. (2015) presented more details. After the quality-control stage, the gauge observations were aggregated in space to match the scale of PERSIANN estimates by averaging rain rates from the gauges within each PERSIANN pixel by using a weighted interpolation scheme, inverse distance weighting (IDW).

PERSIANN

The satellite-based PERSIANN model globally estimates precipitation at 0.25 × 0.25° pixels (Hsu et al. 1997; Sorooshian et al. 2000). It uses the global infrared information from geosynchronous satellites (GOES-8, GOES-10, GMS-5, Metsat-6, and Metsat-7) provided by CPC, NOAA to generate rain rates every half hour, and aggregates these estimations into 6-h rainfall. To increase the accuracy of the estimations, rainfall estimates of PMW sensors mounted on low-orbit satellites, including TRMM, NOAA-15, NOAA-16, NOAA-17, DMSP F13, DMSP F14, and DMSP F15, are used for regular periods. In other words, the PERSIANN model uses a combination of IR-PMW information to estimate precipitation globally. The IR–rainfall relationships of the PERSIANN model are produced using a classification method called a self-organizing feature map (SOFM) (Hsu et al. 1997; Sorooshian et al. 2000). This study used daily PERSIANN estimations of precipitation from 2000 to 2011.

Methodology

The PERSIANN error analysis was performed using the daily rain-gauge data set within the Urmia Lake Basin for 12 years of historical data (4,749 days) from 2000 to 2011. Altogether, 69 pixels of size 0.25 × 0.25° exist within the Urmia Lake Basin containing both PERSIANN and daily rain-gauge data in the time span of this study. An average of 3.6 gauges per pixel are spatially distributed over the Urmia Lake Basin, ranging from a minimum of one to maximum of seven gauges per pixel. The authors acknowledge that the reference data set is not error free and is subject to different uncertainties such as observation errors or even more important uncertainties associated with the lack of areal representativeness. However, Ghajarnia et al. (2015) demonstrated that the inherent uncertainty associated with the reference data set within each pixel in the Urmia Lake Basin is still less than the PERSIANN (and some other HRPEPs) estimation error levels.

To analyze error properties of PERSIANN estimations, this study used several statistical and categorical metrics. All these error analysis techniques have been used in previous studies; however, they were not structured and summarized in a single framework to be used in similar research. The novelty of this paper is its introduction of a well-designed package of error analysis that can help to understand different aspects and performances of a precipitation...
product. By performing the different types of error characterizations that are used in this paper, it is possible to evaluate many different aspects of a precipitation estimation product: (1) overall statistical evaluations; (2) categorical analysis in different rainfall ranges, to evaluate the performance of model around extreme values; (3) rainfall detection capabilities, to evaluate the accuracy of the model in detecting rainfall/no-rainfall events; (4) seasonal evaluations, to determine the effects of different climates and rainfall regimes; (5) error decomposition, to evaluate the portion of false alarm (FA), hit, and miss event biases that normally compensate each other by their opposite signs; and (6) the systematic and random error decomposition, to determine whether improvement of the estimation model should focus on the model algorithm or uncertainty of the precipitation phenomena.

Fig. 2 is a flowchart of the different steps of the error analysis to make clear the methodology of this study. The first step is to obtain the local observations and PERSIANN estimations. Next, it is necessary to prepare the time series of these two data sets and perform some initial setup, such as quality control for the gauge stations, averaging the time series of gauges that share the same pixels, and decoding/reading the PERSIANN estimations from the downloaded files. After this stage, the main error analysis and calculations starts in different categories.

### Statistical Metrics

The statistical metrics used to quantify the differences between PERSIANN and the reference data set were correlation coefficient (CC), mean bias error (MBE), root-mean-square error (RMSE) and coefficient of variation of RMSE (CV-RMSE):

\[
CC = \frac{\text{cov}(\text{Pest}, \text{Pobs})}{\sigma(\text{Pest}) \sigma(\text{Pobs})}
\]  

(1)
The contingency table indexes used in this study were probability of detection (POD), false alarm ratio (FAR), bias, and critical success index (CSI), which verified the accuracy of PERSIANN in estimating the occurrence or nonoccurrence of precipitation. The POD represents the ratio of correct rainfall identification to the total number of observed rainfall events, whereas FAR denotes the ratio of false rainfall reports to the total number of rainfall estimates by PERSIANN. Bias shows the ratio of the number of rainfall alarms by the model to the total number of actual rainfall occurrences, and the CSI index measures the fraction of observed events that were correctly estimated when correct negatives (CN) are removed from consideration. Ebert et al. (1996) provided detailed definitions of the contingency table indexes.

**Categorical Evaluations**

The daily rainfall in the historical time series data of this study varied from 0 to 92 mm/day; the annual average rainfall in the whole country and in the Urmia basin are approximately 250 and 350 mm/year, respectively. Thus, although the error values and properties of PERSIANN estimations are expected to vary in this wide range of rainfall magnitude, it might be meaningful to analyze the behavior of PERSIANN during heavy and light rainfalls to verify probable differences. Likewise, it is also important to determine how PERSIANN estimations would change as the daily rainfall rate of the Urmia Lake Basin increases or decreases. Therefore, considering the frequency distribution of the rainfall data set, five categories were defined: [0-1), [1-5), [5, 10), [10, 20), [20-50), and ≥50 mm/day. Different analyses and measurements were performed and calculated for each category. A simple categorical measure is also introduced, the reliability of PERSIANN estimations, which is formulated as

\[
Rel_j = \frac{N_{\text{Pest} - \text{obs}}}{N_{\text{Pest} - \text{est}}}, \quad j = 1, 2, 3, \ldots, N_j
\]

where \(Rel_j\) = reliability of PERSIANN estimates in category \(j\); \(N_{\text{Pest} - \text{obs}}\) = number of PERSIANN estimates in category \(j\); \(N_{\text{Pest} - \text{est}}\) = number of PERSIANN estimates in category \(j\) for which the corresponding observations were also reported in the same category; and \(N_j\) = total number of defined categories. This index calculates the ratio of PERSIANN success in each category, ranging from 0 to 1, with perfect score equal to 1.

**Error Decomposition**

**Categorical Error Decomposition**

It is possible to break down the total bias or absolute error of the PERSIANN data set into three components, the absolute error during successful detections (hit events), errors due to rainfall misses (miss events), and errors due to false detections (false alarm events) (Habib et al. 2009).

\[
\text{Hit Bias} = \sum (P_{\text{obs}} - P_{\text{est}}) \quad \forall \text{Hit events}
\]

\[
\text{Miss Bias} = \sum P_{\text{obs}} \quad \forall \text{Miss events}
\]

\[
\text{FA Bias} = \sum P_{\text{est}} \quad \forall \text{FA events}
\]

If estimation error is calculated over the total data set, it does not provide information on the source of differences. However, when these are decomposed, it is possible to distinguish among the three possible error sources. Each of these three error sources is caused by different situations and causes in the main estimation model.
Therefore different mechanisms are also needed to deal with these errors, which would lead to different refinement, calibration, or bias reduction methods.

**Error Decomposition into Systematic and Random Components**

The difference between observed and estimated rainfall rates can be described by the mean square difference (MSD), which provides a measure of the average difference (Habib et al. 2009). According to Willmott (1981), MSD can be decomposed into a systematic component (MSD_s) and a random component (MSD_r):

\[
MSD = MSD_s + MSD_r
\]

\[
\sum_{i=1}^{n} (P_{est}^i - P_{obs}^i)^2 = \frac{\sum_{i=1}^{n} (\hat{P}_{est}^i - P_{obs}^i)^2}{n} + \frac{\sum_{i=1}^{n} (P_{est}^i - \hat{P}_{est}^i)^2}{n}
\]

where \(\hat{P}_{est}\) is defined by the least-square linear regression relationship \(\hat{P}_{est} = a + bP_{obs}\), where \(a\) and \(b\) are the intercept and slope, respectively. By performing this decomposition, the proportion of MSD that is attributed to systematic errors can be described by \((MSD_s/MSD)\) and the random contribution can be described by \((MSD_r/MSD)\) or \((1 - MSD_s/MSD)\).

Fig. 3 illustrates the concept of systematic and random errors. The system of an estimation model (or a prediction model) is correct only when the trend line of estimated points is located on the perfect agreement line (estimates = observations). Otherwise, the difference between the existing trend line and the perfect agreement line demonstrates the value of the systematic error. In contrast, random error is defined by the distance between each point’s location and its equivalent value on the existing trend line of estimation system.

### Results and Discussions

#### Analysis of Statistical Metrics

Fig. 4 shows the seasonal values of all statistical measurements in the basin. These indexes were calculated seasonally for all basin
pixels, and their average and standard deviations were calculated (Fig. 4). These indexes provide an overview of the accuracy of PERSIANN precipitation estimation of this region. This figure shows that generally the correlation coefficient of PERSIANN estimations with observed rainfalls is quite low (less than 35%), which is not enough to make this data set a reliable precipitation product for use in hydrologic studies. This correlation values show the necessity of PERSIANN model adjustment or postcalibration in this region. The seasonal correlation coefficient varied from 0.14 in winter to 0.34 in fall, with the lowest correlation in the cold season when the surface of the basin was filled with snow cover. The MBE index indicated underestimation in fall, winter, and summer and overestimation in spring. The MBE values in each season may not necessarily be attributed to less or more rainfall estimation magnitude than the real observations on rainy days. In other words, the bias in many cases is likely to be caused by false detection or missing events, which can be a great source of bias error in the estimation data set.

The RMSE measure also shows more estimation error during spring and the least seasonal error in summer. However, abundant sunny days (no rainy days) in summer might have led to a lower level of RMSE for the estimation data set. Therefore the CV-RMSE index is a better measure for seasonal comparisons. The relative RMSE (CV-RMSE) of PERSIANN in summer in the Urmia Lake Basin is greater than in other seasons, which mainly show little fluctuation during fall, winter, and spring (Fig. 4). To better analyze causes and reasons associated with these error levels, the following sections perform more evaluations.

**Analysis of Contingency Table Measures**

Contingency table indexes measure the capability of the PERSIANN model to accurately estimate rainfall occurrence (Fig. 5).

A higher level of POD was observed in spring and fall, whereas POD values in summer and winter were much less. This was probably caused by the existence of nonprecipitating cold clouds at high altitudes in winter and orographic precipitating warm clouds in summer in the Urmia Lake Basin. On the other hand, the overall value of PERSIANN FAR in this area was above 50% in winter, slightly less than 70%. This indicates that the PERSIANN algorithm had many detection errors in winter seasons over the Urmia Lake Basin. The land surface of the mountainous areas in the Urmia Lake Basin is mainly covered by snow, which can be a source of missed rainfalls and lower levels of POD values (Tian et al. 2009; Nasrollahi et al. 2013). Nasrollahi et al. (2013) also observed very high levels of FAR during the winter season in the United States, which was associated with the presence of nonprecipitating high cirrus clouds and the presence of snow and ice cover on the ground.

The bias index indicates the ratio of rainfall estimates (hit + FA) to the rainfall observations (hit + miss). Fig. 5 shows underestimation of the number of rainy days in fall, winter, and summer. The highest level of underestimation was observed in summers, when the occurrence of sunny days (CN events) is dominant in the Urmia Lake Basin and mainly convective rainfall patterns may cause precipitation. The slight overestimation of PERSIANN in the spring was also accompanied with a FAR of nearly 60%. This denotes that although the total number of rainfall detections was close to the observed events, more than half were detected on days when no actual rainfall occurred. The CSI index generally demonstrated better performance of the PERSIANN algorithm in spring and fall rather than in summer and winter, which can be associated with high-altitude cold clouds and snow cover in winter and a convective pattern of warmer clouds in summer.

In summary, the overall behavior and error characteristics of PERSIANN estimations in each season varied. Therefore the

---

**Fig. 5.** Seasonal contingency table indexes calculated for daily data set of PERSIANN model compared with reference data set; 1 mm/day was considered the daily rainfall threshold for these calculations.
approach and method of PERSIANN modifications or bias reductions need to be specified separately for each season, considering the seasonal differences and behaviors.

**Categorical Evaluations**

Fig. 6 presents the categorical performance of PERSIANN compared with the reference data set. This graph was generated by pooling data pairs from all pixels in the total historical data set into one plot (daily data between 2000 and 2011), which were categorized in each actual (observed) rainfall category. The analysis was separately done in six categories (Fig. 6). On the left and right of each category, bar charts present the average values of reference and PERSIANN data, and each category plots all relevant rainfall estimates of the PERSIANN product, e.g., the plotted points in the (0–1) category represent the PERSIANN rainfall estimations between 0 and 1 mm/day. The dark rectangles in the graph show areas between the minimum and maximum values of each category in which the PERSIANN estimations are expected to be reported.

It is evident from the left (average gauge data) and right (average PERSIANN estimations) bar charts for each category of Fig. 6 that as the value of observed precipitation increased, the value of PERSIANN estimations also increased, which shows similar behavior in both data sets. However, although PERSIANN captured the increasing order of observations in gauge precipitation, this increase was not large enough to reach the allowable zone (dark rectangles). On the other hand, the overall values of PERSIANN estimations in all categories had relatively similar ranges regardless of the value of observed precipitations, i.e., the ranges of plotted points in all categories were the same. For instance, even when all rainfall events were observed in the ranges 0–1 or 1–5 mm/day, some PERSIANN estimations were still reported at a rate of more than 40 mm/day. This finding indicates that not all the aspects and nature of rainfall processes in the Urmia Lake Basin were reasonably modeled by the PERSIANN algorithm. This might be related to the low density of the rain-gauge network in this area or to the number of stations that report data to international meteorological organizations. However, it seems that an error modeling approach can locally be useful for modeling some of the remaining relations and dependencies between PERSIANN estimations and observed data which are not primarily modeled by PERSIANN algorithm.

Given these points, it is necessary to see how much it is possible to rely on any single PERSIANN estimation in the future. This was determined by considering the concept of the reliability of PERSIANN estimations [Eq. (5)]. This index was categorically calculated (Fig. 7). The first category of this figure is the zero estimation or no-rain condition. Other categories are the same as in Fig. 6, although, because PERSIANN did not report any rainfall rate more than 50 mm/day, Fig. 7 does not present the last category (>50). Because the information in Fig. 7 was obtained from calculations over the historical data from 2000 to 2011 in the Urmia Lake Basin, it is possible to trust the PERSIANN estimations of a no-rain situation with a confidence level of 86.3%. In other words, if PERSIANN estimates a no-rain condition, this estimation can be accepted with 86.3% confidence, which can be evaluated as a good result regardless of the 14% incorrect estimates. Nevertheless, the values of the reliability index of PERSIANN estimations in other categories were all less than 20%, which is not a desirable result.

![Fig. 6. Comparison of PERSIANN estimations and rainfall observations in different categories; although PERSIANN is in general able to follow the increasing manner of observations, the magnitude of its increase is much less than it should be.](image1)

![Fig. 7. Reliability of PERSIANN estimations index calculated for different categories; first category is allocated to estimation of value 0, which represents no-rain condition](image2)
For example, if PERSIANN estimates the precipitation rate higher than 20 mm/day, it can almost be confidently assumed that the gauge stations observed rainfall in other categories. Because the actual rainfall mostly is not observed in the same category as PERSIANN estimations, the true categories must be determined (Figs. 8 and 9). The graphs of Fig. 9 have two horizontal axes and one vertical axis; the horizontal axes are the PERSIANN and gauge categories, and the vertical axis represents the frequency percentage or probability of joint occurrences.

Figs. 8 and 9 present the seasonal posterior probabilities of true rainfall observations versus PERSIANN estimations in different categories. In other words, this information shows the percentage of gauge observation frequencies in different categories for different ranges of PERSIANN estimations in each season separately. For instance, in the third category of the fall season (1–5) and when PERSIANN estimations were not accurate, 51.8% of the gauge station observations were 0 (no-rainfall, i.e., FA); in 8.2% of the cases, gauges observed rainfall of 0–1 mm/day; and so on.

Most noticeable in Figs. 8 and 9 is that probabilities related to the zero categories (no-rain conditions) were much higher than were other gauge categories. This indicates that the majority of PERSIANN mistakes, when its estimations were not confirmed by gauge observations, were FA errors, and this condition was more dominant in the winter season. Moreover, with increasing rainfall rate in higher categories, the portion of FA errors decreased in every season except winter. This means that when PERSIANN estimated higher magnitudes of rainfall, the probability of FA error decreased, although this was not the case in winter.

To better investigate the detection capability of the PERSIANN algorithm in higher-rainfall categories, another categorical analysis of POD and missed data percentages of PERSIANN estimates (Fig. 10) was conducted. As the rainfall increased in higher categories, the portion of missed events decreased (Fig. 10). In fact, during heavy rainfalls, PERSIANN’s ability to detect rainfall events increased; the missed percentage decreased and POD values increased. This general rule was observed in every season, including winter. The only exception, with a reverse trend, was observed in the category of >50 mm/day in the fall season; this may be related to the small number of events in this category, which may have influenced the statistical results.

### Error Decomposition

#### Results of Categorical Error Decomposition

The total and seasonal bias of PERSIANN compared with the reference data set were decomposed into three components, hit bias, FA rainfall bias, and missed rain bias (Fig. 11). These volumes are combined totals from the individual pixels analyzed over the total daily data set from 2000 to 2011. The miss and FA bias portions were more dominant than the hit bias but contributed conversely and canceled each other because of their opposite sign. Seasonal comparisons show that although the FAR of spring was less than that of winter (in Fig. 5), the value of FA bias for spring was the greatest (Fig. 11). In fact, spring is the only season for which the value of FA bias was less than the missed rainfall bias. The value of missed rainfalls during winter and fall was considerable, which can be due to the effect of snow and ice cover over the Urmia Lake Basin lands during these two seasons.

In addition to the bias, the absolute error was also decomposed among these three components categorically (Fig. 12). The results of analysis show that as the value of rainfall rates increased in different categories, the portion of hit MAE also became more

![Table](https://example.com/table.png)

**Fig. 8.** Seasonal posterior probabilities of true rainfall observations versus PERSIANN estimations in different categories; shaded cells in tables represent seasonal values of reliability index of PERSIANN estimations.
dominant and the portion of missed event MAE decreased. This means that in the case of heavier rainfalls, the PERSIANN algorithm mainly needs to increase its accuracy in estimation of rainfall magnitude rather than its precision in detection of rainfall occurrence. Fig. 12 also indicates that in the first category (0–1), the portion of FA errors was highly dominant, approximately 83% of the total MAE by three components. This is another confirmation of the importance of FA error removal in PERSIANN estimations to improve its accuracy.

Error Decomposition into Systematic and Random Components

The mean square of differences and its two components as defined in Eqs. (9) and (10), systematic and random errors, were also assessed. Obviously, according to Eqs. (9) and (10), these calculations were performed using only hit events, for which both the reference and estimation data sets had reported rainfall values. In general, according to the average value of all pixels in the Urmia Lake Basin, approximately 60% of the PERSIANN total MSD value was systematic error and the remaining 40% was related to random errors. Therefore it can be concluded that the PERSIANN algorithm in the Urmia Lake Basin mainly encountered problems modeling the system and pattern of rainfalls rather than overcoming uncertainties associated with this phenomenon.

Fig. 13 shows PERSIANN’s total MSD values and the proportion of its components in different seasons over the Urmia Lake Basin. The portions of systematic error in summer and winter were greater than those in spring and fall. This means that although PERSIANN had shortcomings capturing the general system of precipitation in this area, it was more successful in simulating spring and fall precipitation than that of winter and summer. This is further proof that high, cold, nonprecipitating clouds and snow cover in winter and low, warm clouds in summer may cause many problems for satellite-based precipitation estimation models to capture the system of precipitation. On the other hand, it seems that PERSIANN is more capable of modeling the convective nature of rainfalls in spring because its systematic error during this season was less than in the other seasons.

Finally, Fig. 14 presents the changing systematic and random errors as the value of rainfall varied in different categories. As the value of precipitation increased in higher categories, the proportion of systematic error dominated the random error. This means that for light precipitation, the detection of rainfall occurrence by the PERSIANN model plays an important role and seems to be enough in terms of systematic modeling. Therefore, in this condition, precipitation uncertainties and random errors are dominant and significant. On the other hand, although PERSIANN detects the occurrence of heavier rainfalls more accurately (Fig. 10), this is absolutely insufficient to reduce the systematic error, and much more accurate modeling of the rainfall rate is needed during these conditions.

Summary and Conclusions

This paper investigated different characteristics and error components of the PERSIANN precipitation estimation product over the Urmia Lake Basin, Iran. Different approaches and indexes were used to evaluate PERSIANN precision and accuracy in detection and estimation of rainfall events, including statistical, categorical, and seasonal analysis. Total residuals were decomposed into hit/miss and FA biases, and the systematic and random error
proportions were analyzed seasonally and categorically. This study yielded new insights into the nature of the errors in PERSIANN estimates over the Urmia Lake Basin as a semiarid region in the Middle East, including the following:

1. The analysis showed poor rainfall-detection ability for PERSIANN during winter and summer seasons. The analyzed contingency table indexes indicated better detection precision during spring and fall over the Urmia Lake Basin and showed high error values during winter and summer. This can be caused by high-altitude cold clouds and ground snow cover in winter, and the convective pattern of warmer clouds in summer.

2. A relatively constant level of error was generally observed in PERSIANN estimations among different categories (Fig. 6). This study showed that the range of precipitation estimates in different categories from 0–1 to more than 50 mm/day of rainfall was nearly invariant, which is a sign of the existence of systematic error in the PERSIANN precipitation estimation algorithm.

3. Further investigations showed a low reliability level of PERSIANN estimations in different categories, which were

![Fig. 10. Seasonal categorical POD values and missed data percentages; index of missed data percentage is calculated by dividing total number of missed events by total number of rainfall events (missed events plus hits)](image)

![Fig. 11. Decomposition of total bias of PERSIANN product from reference data set into bias during successful hits, bias due to FA events, and bias due to missed rains)](image)
mostly associated with a high level of FA error in this area. Error decomposition of total bias into hit, miss, and FA portions also indicated considerable biases which were caused by miss and FA errors in PERSIANN estimations. However, as the rate of precipitation increased in higher category levels, the ability and precision of PERSIANN in rainfall detection also increased.

4. Finally, the systematic and random error decomposition indicated that PERSIANN in the Urmia Lake Basin was less successful in modeling the system and pattern of rainfalls rather than in overcoming uncertainties associated with precipitation. The level of systematic error considerably increased in heavier rainfalls, and PERSIANN showed poor performance in accurate estimation of rainfall rate.

Fig. 12. Comparison of portion of hit, miss, and FA mean absolute error values in different categories; as rainfall increases, the portion of hit MAE becomes more dominant.

Fig. 13. Proportion of systematic and random errors in different seasons over Urmia Lake Basin; the portion of PERSIANN systematic error is mainly dominant in this area.
Fig. 14. Categorical evaluation of changing manner of systematic and random error proportions by increase of rainfall rate in different categories

Furthermore, PERSIANN error characteristics in each season varied due to the condition and rainfall patterns of each season, which shows the necessity of seasonally different approaches for the calibration of this product. Overall, the authors believe that these different types of error analysis can substantially help any further local studies for postcalibration and bias reduction of PERSIANN estimations. In other words, without performing such preliminary studies over the error characteristics of a satellite-based precipitation estimation product, pursuing fundamental studies for bias reduction of such products are hardly possible and successful.

References


