

Research Article

Bias Correction in Monthly Records of Satellite Soil Moisture Using Nonuniform CDFs

Shan Wang ¹, Huiling Shan,¹ Chi Zhang,¹ Yuexing Wang,¹ and Chunxiang Shi²

¹School of Information Engineering, East China Jiaotong University, Nanchang 330013, China

²National Meteorological Information Center, Beijing 100081, China

Correspondence should be addressed to Shan Wang; patrick_shan@163.com

Received 12 February 2018; Revised 24 April 2018; Accepted 16 May 2018; Published 16 July 2018

Academic Editor: Jifu Yin

Copyright © 2018 Shan Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

It is important to eliminate systematic biases in the field of soil moisture data assimilation. One simple method for bias removal is to match cumulative distribution functions (CDFs) of modeled soil moisture data to satellite soil moisture data. Traditional methods approximate numerical CDFs using 12 or 20 uniformly spaced samples. In this paper, we applied the Douglas–Peucker curve approximation algorithm to approximate the CDFs and found that three nonuniformly spaced samples can achieve the same reduction in standard deviation. Meanwhile, the matching results are always closely related to the temporal and spatial availability of soil moisture observed by automatic soil moisture station (ASM). We also applied the new nonuniformly spaced sampling method to a shorter time series. Instead of processing a whole year of data at once, we divided it into 12 datasets and used three nonuniformly spaced samples to approximate the model data's CDF for each month. The matching results demonstrate that NU-CDF3 reduced the SD, improved R , and reduced the RMSD in over 70% of the stations, when compared with U-CDF12. Additionally, the SD and RMSD have been reduced by over 4% with R improved by more than 9%.

1. Introduction

Estimates of satellite soil moisture can be improved using statistical correction or scaling approaches, which can be particularly valuable prior to using a satellite soil moisture product in an assimilation system [1]. The cumulative distribution function (CDF) matching approach is a statistical correction method that has been used to adjust microwave satellite observations using China Land Data Assimilation System- (CLDAS-) simulated soil moisture. This approach was used in a number of previous studies. Lee and Anagnostou [2] presented a retrieval scheme for near-surface soil moisture, which is based on a combined passive/active microwave remote-sensing algorithm. Through experiments we found that the accuracy of the retrieval was shown to depend on the overlying vegetation and soil wetness conditions. In moderate vegetation cover, the retrieved values seem to reproduce the trend well in soil moisture dry down. Atlas et al. [3] found that the radar-retrieved CDFs of rain rate then replicate the CDF of gage-measured rates nicely;

a real change in the mean rain rate is manifested by the change in the probability distribution of reflectivities. Liu et al. [4] produced a merged dataset covering from late 1978 through 2006 using the CDF matching technique and confirmed the strong impact of ENSO on soil moisture and vegetation condition across Australia. Yuan et al. [5] introduced data matching between four remotely sensed soil moisture products (ASCAT, WindSat, FY3B, and SMOS) and ASM stations using CDF method. And then, results show that all these four products performed well in the Northwest China. Furthermore, the same satellite soil moisture product showed great spatial differences in different regions. Reichle and Koster [6] found that a simple method of bias removal was to match the cumulative distribution functions (CDF) of the satellite and model data. And then, by using spatial sampling with a 2-degree moving window, they can obtain local statistics based on a one-year satellite record that are a good approximation to those that would be derived from a much longer time series.

Over the years, several different multifrequency passive microwave sensors have been used to estimate surface soil

TABLE 1: A brief introduction of three satellites.

Satellite	Nation	Resolution	Function	Characteristics
FY3B [7]	China (2010)	25 km	FY3B is the second satellite of the FY3 series, and it is the first in the series to carry a microwave imager (MWRI) for observing the Earth's surface in 10 microwave channels.	The FY3B satellite has 5 frequencies (10.65, 18.7, 23.8, 36.5, and 89 GHz) and dual polarization (vertical and horizontal polarization) mode
SMOS [8, 9]	Europe (2009)	25 km	SMOS is the only satellite in the world that can observe changes in soil moisture and sea salinity.	The SMOS satellite which emits the microwave signals at 1.4 GHz frequency (L band) can penetrate the soil to a depth of 5 cm.
WindSat [10]	America (2003)	25 km	Besides providing the sea surface wind vector measurement for the Navy, WindSat also observed other sea surface environmental parameters and was also used to measure soil moisture and sea ice.	The WindSat discrete frequency band is 6.8, 10.7, 18.7, 23.8, and 37 GHz, of which 10.7, 18.7, and 37 GHz are all polarization mode and 6.8 and 23.8 GHz are dual polarization mode

moisture. The soil moisture data from WindSat, SMOS, and FY3B passive microwave remote-sensing soil moisture products are introduced in this paper. The satellite sensors are introduced, respectively, in Table 1.

A simple method of bias removal is to match the CDFs of the satellite and model data. However, accurate CDF estimation typically requires a long record of satellite data. When the data sample is not enough, the matching effect will be unstable. This paper will analyze the matching effect based on the annual scale and the monthly scale.

The objective of this study is to use modeled soil moisture to improve the dynamic range of the temporal variability of the surface soil moisture from the three satellites described above (FY3B, SMOS, and WindSat). We applied the CDF matching technique to adjust the limited temporal variability of the satellite data using the common land model (CLM). To investigate the impact of the bias correction, we used three statistical indicators as the evaluation criteria: the standard deviation (SD), the correlation coefficient (R), and the centered root mean square difference (RMSD). Traditionally, the CDF curve is represented by 12 uniformly spaced samples. The resulting set of piecewise linear equations approximate the original CDF [11].

2. Data and Methods

2.1. Satellites and AWS Data Description. The automatic soil moisture station (ASM) was established in 2009. And to October 1, 2013, a total of 2111 stations has been established, of which 1555 were put into operation and distributed throughout the country. The daily average soil moisture volume data (unit m^3/m^3) of 376 stations after 0–10 cm level quality control from 2011 to 2013 was used in this study.

All the data used in this paper were from January 1, 2011 to December 31, 2013, including the ASM 0–10 cm level daily average station (after quality control) observation data (text format), WindSat global land surface soil moisture day product (binary format), SMOS global land surface soil moisture 3 day product (Buffer format), and FY3B global surface soil moisture level 2 daily products (HDF5 format, EASE-GRID projection).

In order to facilitate comparative evaluation, WindSat, SMOS, and FY3B soil moisture products are unified into

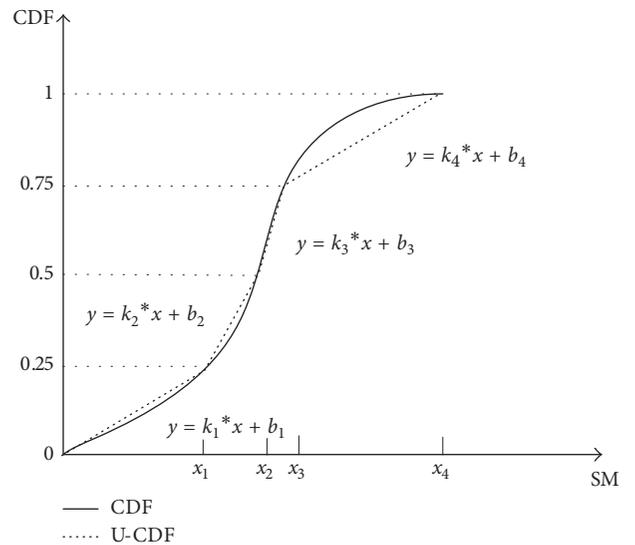


FIGURE 1: Approximating CDF with uniform samples.

binary format storage, the corresponding projection mode of latitude and longitude, region range (0° – 60° N, 70° – 150° E), daily product, and 25 km spatial resolution. The latitude and longitude information of ASM stations are known, microwave remote sensing soil moisture products have been projected to the corresponding latitude and longitude, and region range has been cut. Therefore, the corresponding product rank numbers can be calculated according to the latitude and longitude information of the station, and WindSat, SMOS, and FY3B soil moisture gridding data are interpolated to ASM stations, respectively. Then the spatial matching of the data is completed.

2.2. Uniform and Nonuniform CDF Matching Methods

2.2.1. CDF Matching. The principle behind CDF matching is straightforward. Let X and Y denote the soil moisture of the original and scaled satellite data, respectively. The original satellite data and the scaled satellite data have probability density functions. The CDF of random variable X is

$$F_X(x) = \int_0^x f_X(x) dx. \quad (1)$$

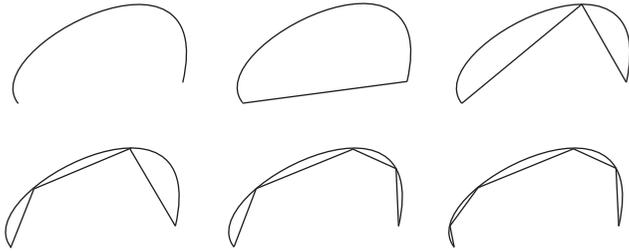


FIGURE 2: The Douglas–Peucker algorithm computes a polyline approximation of a smooth curve.

The CDF of random variable Y is

$$G_Y(y) = \int_0^y g_Y(y) dy. \quad (2)$$

When given a value x of X , y can be found from the following equation:

$$F_X(x) = G_Y(y). \quad (3)$$

2.2.2. Uniform CDF. For simple computation, we usually do not take the whole CDF curve into the matching calculation. Instead, we use several straight lines to approximate the CDF curve. Computer-stored straight lines only need to store its slope and intercept, which can greatly improve the efficiency of calculation. Uniform sampling of CDF curves to get piecewise straight lines is a conventional method. The concept is very simple, taking 4 straight lines as an example (Figure 1). The value of CDF is between 0 and 1, and it is equally divided into 4 segments, that is, the sampling values are 0, 0.25, 0.5, 0.75, and 1. Then the corresponding soil moisture values $x_1, x_2, x_3,$ and x_4 were obtained from the CDF curve. Connecting these sampling points, we can get 4 straight lines, which can be expressed as $y = k_i x + b_i$ ($i = 1, 2, 3, 4$). In the next CDF matching calculation, these 4 lines were used in data calibration instead of CDF curve. In practical applications, the empirical value of the number of sampling segments is 12 segments, so that a good approximate CDF curve can be obtained.

2.2.3. Nonuniform CDF. The polygonal curve approximation algorithm can be used to compress a densely sampled CDF by representing it with a reduced set of nonuniform samples. When compared with a uniform sampling method, it provides a more compact yet accurate approximation of the original function. Polygonal approximation algorithms take as input a curve represented by an N -segment polyline and produce an M -segment polyline with vertices that minimize the difference between the two (typically $M < N$). Although there are algorithms that output the optimal solution [12–14], we have used the Douglas–Peucker [15, 16] algorithm because it is simple and fast. It has also been shown that these greedy algorithms typically produce results within 80% accuracy of the optimal solution [17].

The Douglas–Peucker curve approximation algorithm looks for the next sample that is furthest from the current polyline (Figure 2). Initially, only the end points of the curve

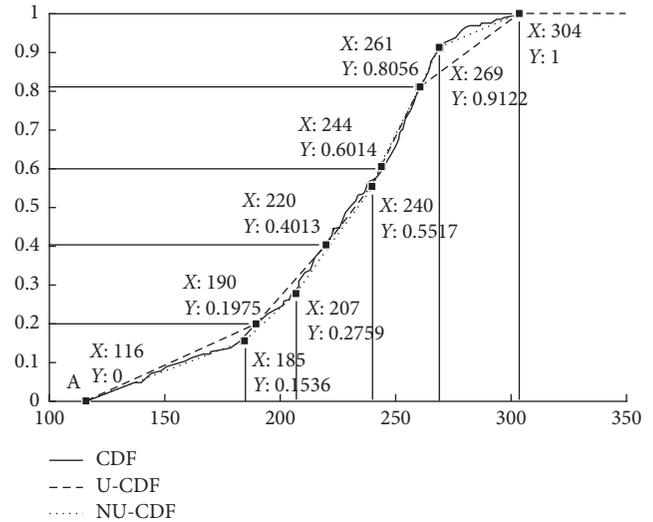


FIGURE 3: Approximate CDF curve with five uniformly and nonuniformly spaced samples.

are selected and the algorithm iteratively inserts the vertex as the approximation, until reaching an error threshold or maximum number of vertices [18]. This algorithm can produce a near-optimal approximation to a numerical CDF using a small number of samples.

We must first sample the CDF of the CLDAS model data and then apply piecewise CDF matching. A key problem is how to determine the number of samples that should be used to achieve the best matching results for bias correction. We first used a 5-segment sampling to demonstrate that the NU-CDF matching method is superior to the U-CDF matching method and then chose the optimal sampling segment for both methods. Figure 3 shows the CDF curve for the CLDAS model data (black line) and six uniformly and nonuniformly spaced samples. The concept behind uniform sampling is easy to understand. It divides the vertical axis into five parts with equal lengths, so six equally divided points are found in the interval $[0, 1]$; that is, (0, 0.2, 0.4, 0.6, 0.8, and 1). The corresponding six samples in the CDF curve can be easily located. The CDF curve typically starts with a series of 0s and ends with a series of 1s (as can be seen in the figure). We select the last 0 (marked with “A”) and the first 1 (marked with “B”) as the two initial samples. Meanwhile, because the CDF curve is numerical, the CDF curve does not often contain points that are exactly 0.2, 0.4, 0.6, and 0.8. We instead choose samples that are closest to the division points (0.1975, 0.4013, 0.6014, and 0.8056, in this plot). Finally, we connect these six samples to produce a polyline that approximates the CDF curve. This uniform sampling method is simple, but it does not take into account the shape of the CDF curve. NU-CDF can produce a more compact approximation by using the characteristics of the CDF curve. It takes A as the start point and B the end point, connects them, and finds a third point that is furthest away from the current polyline approximation. This point is 185, 0.1536. The fourth sample (269, 0.9122) is obtained by connecting the third sample and B to a polyline and finding the point

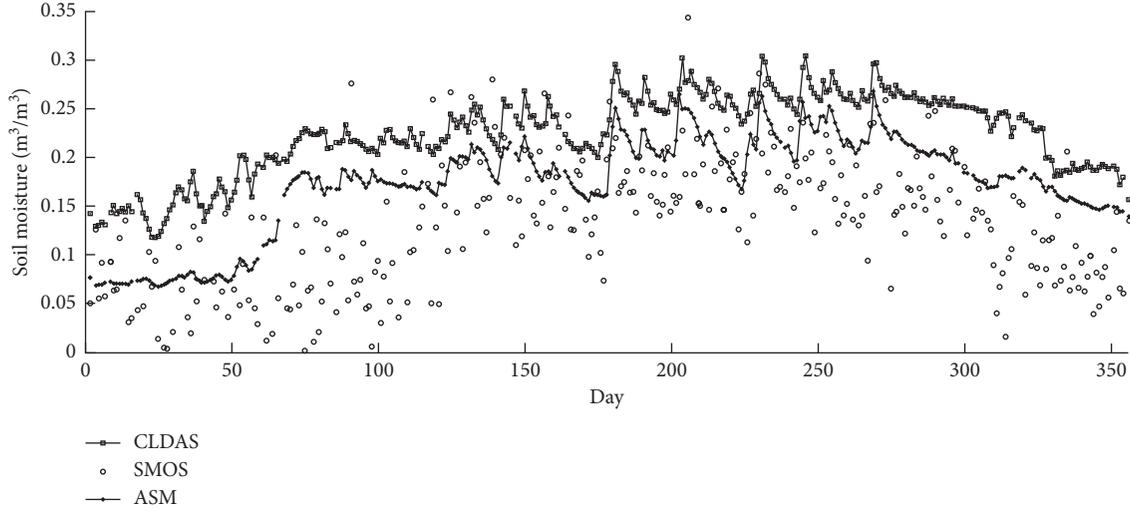


FIGURE 4: Time series of soil moisture estimates from the CLDAS model, SMOS data, and ASM data for Gansu Province.

that is furthest away. The fifth and sixth samples are determined using the same technique. As the plot shows, the black dashed line that represents NU-CDF is closer to the original CDF curve than the gray dashed line that represents U-CDF. The difference between the two is particularly obvious in the last segment. Because U-CDF only depends on the vertical axis and does not consider the shape of the CDF, “information” is lost during the approximation process.

2.3. Statistics. To evaluate the bias correction, we used three indicators: the standard deviation (SD), the correlation coefficient (R), and the centered root mean square difference (RMSD). For a set of soil moisture values, x_1, x_2, \dots, x_n ,

$$SD = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}, \quad (4)$$

where

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i. \quad (5)$$

The correlation between the first time series and a second one (ASM) with observations y_1, y_2, \dots, y_n is

$$R = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{N \sqrt{(1/N) \sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{(1/N) \sum_{i=1}^N (y_i - \bar{y})^2}}, \quad (6)$$

and

$$RMSD = \sqrt{\frac{\sum_{i=1}^N [(x_i - \bar{x}) - (y_i - \bar{y})]^2}{N}}. \quad (7)$$

In general, SD and RMSD are smaller, or the correlation coefficient is closer to 1. We think the calibration results are better.

3. Results and Discussions

The experimental results of this study are divided into three parts. First, the data deviation correction results of a single satellite in a specific province are given. Then we divided one year of data into 12 parts according to each month, and separately implemented CDF matching. Finally, we extend the research area to the whole region of China and consider the results of multiple satellites at the same time.

3.1. Case Study in Gansu Province. To illustrate the CDF matching process, we analyzed SMOS soil moisture retrievals for Gansu Province in 2012. Figure 4 shows the soil moisture value of the satellite retrievals, the corresponding CLDAS model dataset, and the ASM dataset. The horizontal axis represents the time in days, and the vertical axis represents the soil moisture value in m^3/m^3 . The satellite data are in dispersed circle, the model data are in black line with square, and the ASM are in gray line with dot. The soil moisture values from SMOS tend to be drier, and so we must scale the satellite data using the model data to reduce the bias. Note that some days do not have soil moisture data, but this does not affect the matching process.

As discussed in CDF matching, the scaled SMOS soil moisture data can be easily computed by both U-CDF and NU-CDF matching. Figure 5 shows the ASM dataset and two piecewise CDF-scaled SMOS soil moisture datasets. Both methods can convert the satellite data value range to the ASM data, but the difference between the two is not obvious. To clearly highlight the differences, we considered the data from Day 200 to Day 230, as shown in Figure 6. The data from NU-CDF are closer to the ASM on many days (e.g., Days 207, 212, 214, and 219). A Taylor diagram is an intuitive and convenient way to represent these three parameters. It can be used to summarize the relative merits of a collection of different models or to track changes in the performance of a model as it is modified [18]. Figure 7 contains a Taylor diagram. Four points are plotted on the polar style graph, with ASM representing the reference data,

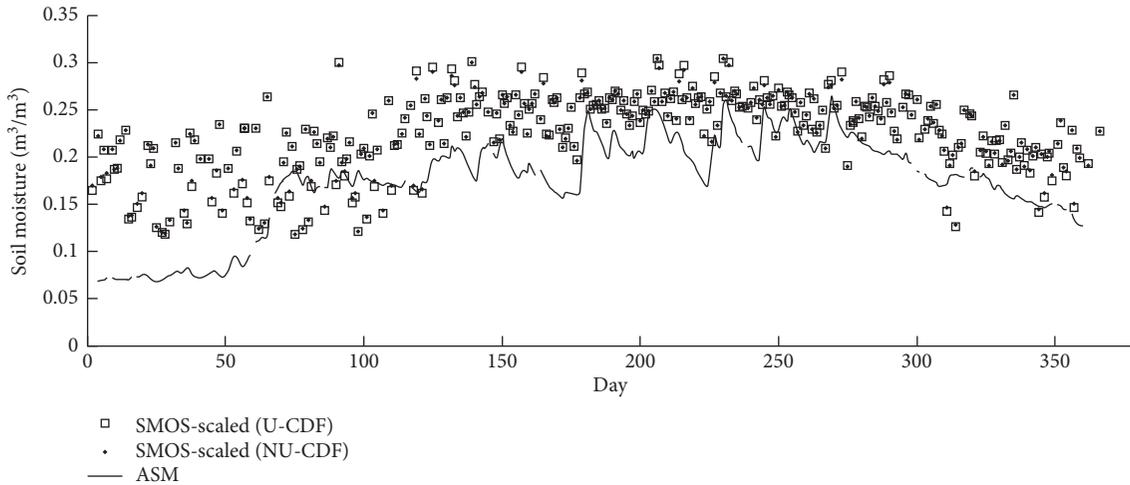


FIGURE 5: Time series of ASM data and two scaled SMOS soil moisture datasets using the U-CDF and NU-CDF matching methods.

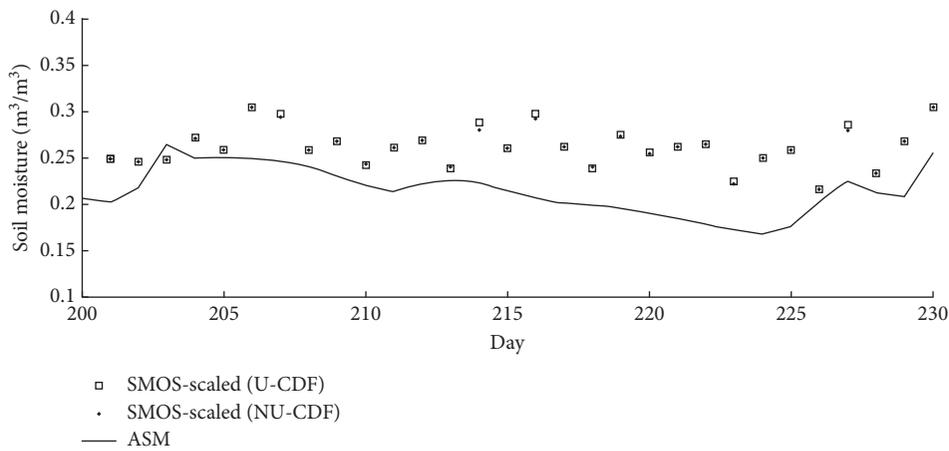
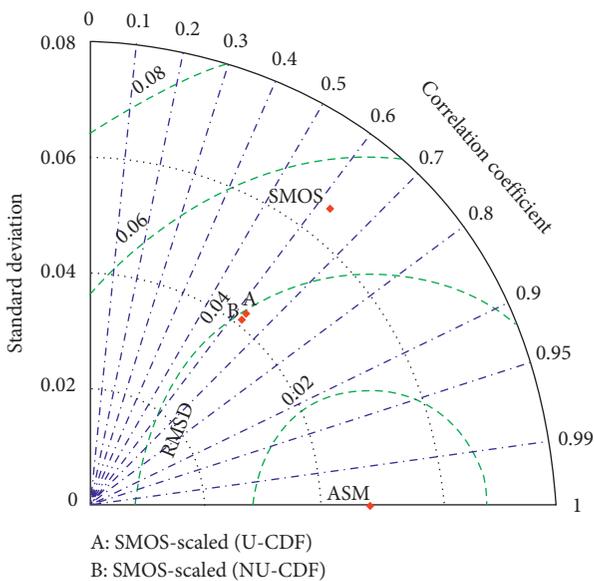


FIGURE 6: Soil moisture data for Days 200–230 in Figure 3. In most cases, NU-CDF is more accurate.



A: SMOS-scaled (U-CDF)
 B: SMOS-scaled (NU-CDF)

FIGURE 7: Taylor diagram displaying three parameters: standard deviation (SD), correlation coefficient (R), and centered root mean square difference (RMSD).

SMOS representing the original satellite data, “A” representing the U-CDF-scaled SMOS data, and “B” representing the NU-CDF-scaled SMOS data. The radial distances from the origin to the points are proportional to the pattern SD, and the azimuthal positions represent the value of R between the two fields. The RMSD between the scaled satellite data and the ASM data is proportional to the distance between them (in the same units as SD). Both U-CDF and NU-CDF reduced the SD of the satellite data. The SD of the original SMOS data was 0.0658, and it decreased to 0.0453 using U-CDF and 0.04134 using NU-CDF. The SD of the model data was 0.04014, and, as expected, the NU-CDF matching method effectively reduced the bias. However, both methods failed to improve the correlation or to reduce the RMSD of the satellite data. We will discuss this problem in Section 3.2.

It is obvious that more or less samples can be inserted into the CDF curve, so we must decide on the optimal number of segments in a polyline approximation for bias correction. We used the above data and approximated the CDF curve using different numbers of segments, using both U-CDF and NU-CDF matching. We used the SD to evaluate the bias. Figure 8 plots the relationship between the SD and

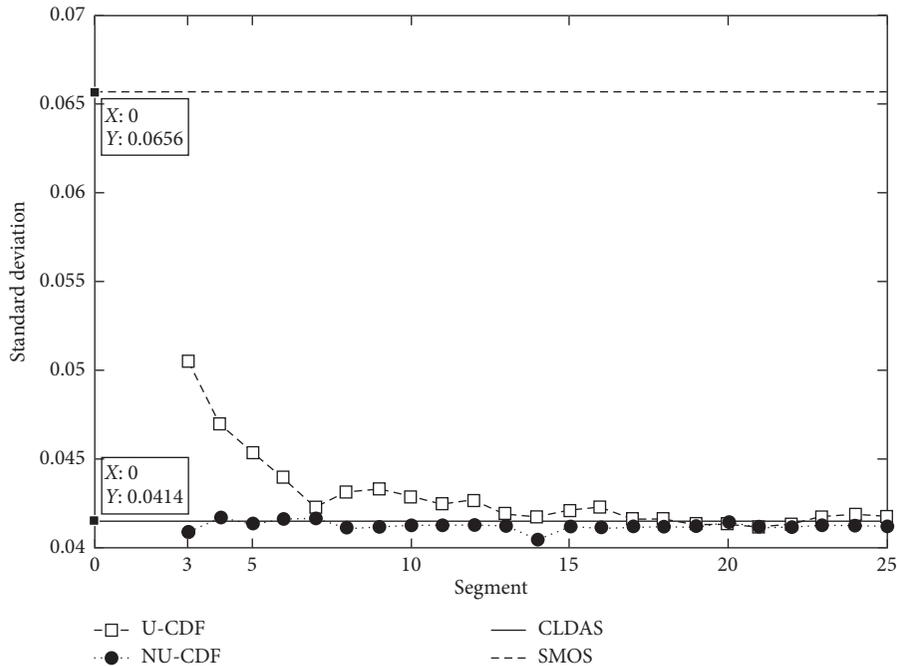


FIGURE 8: Relationship between SD and different sampling segments for both U-CDF and NU-CDF.

different samplings from U-CDF and NU-CDF. The horizontal axis represents the number of segments in the polyline approximation, and the vertical axis represents the SD. As shown in the Taylor diagram, the SMOS data had an SD of 0.0658, and CLDAS had an SD of 0.0414. Consider U-CDF (the hollow square). Initially, the SD quickly decreased as the number of segments increased. When there were 12 or more segments, the SD tended to be stable; the values fluctuated around the CLDAS SD value (0.0414) and had a maximum of 0.04259 (when using 12 segments). Therefore, we used 12 segments to approximate the CDF curve using the traditional method. Now, consider NU-CDF (the points). Regardless of the number of segments, the SD was always stable around the CLDAS value. The minimum value was 0.04029 (using 14 segments), and the maximum was 0.0417 (using four segments). Therefore, NU-CDF requires fewer samples to reduce the SD, when compared with U-CDF. Note that we have only used SMOS soil moisture data for Gansu Province in 2012 to reach this conclusion. Our more extensive experiments described in Section 4 demonstrated that, in most cases, this conclusion still holds when the study area is expanded to the entire China region and considers different satellite data (SMOS, FY3B, and WindSat).

3.2. CDF Matching Using a Month Data. Accurate CDF estimation typically requires a long record of satellite data. To correct the biases, the temporal statistical moments of both the simulated soil moisture and the satellite-derived soil moisture must be well established. Without further assumptions, this would require many years of data. However, we can still use a short record of satellite data under the constraint that we do not have global estimates of the data's

temporal statistical moments [6]. In the following discussion, we divided one year of data into 12 parts according to each month, and then separately implemented CDF matching.

Because the efficient number of satellite soil moisture data in a month is 31 or less, the 12-segment U-CDF (U-CDF12) method is obviously no longer feasible. However, the NU-CDF method still works in this situation. We directly selected the three-segment polyline approximation of the CDF curve computed using 1 month of CLDAS data. Figure 9 displays two scaled SMOS soil moisture datasets from the monthly NU-CDF matching (points) and the traditional U-CDF matching (squares) methods for a 1-year period. Note that the point sequence is composed of 12 independent NU-CDF matching results. In this figure, the squares are dispersed relatively far away from the ASM data, but the points are in the vicinity of the ASM data. The soil moisture values for May are given in Figure 10, to highlight the details. This magnified view shows that soil moisture values from the three-segment NU-CDF (NU-CDF3) are almost always closer to the ASM data. As previously mentioned, the discontinuity is due to unavailable satellite data.

We computed the SD, R , and RMSD to evaluate the bias correction. They are plotted in Figure 11. It is obvious that NU-CDF3 significantly improved R and reduced the RMSD, although the SD did not benefit from the monthly matching technique when compared with U-CDF12 for a 1-year period.

3.3. The Entire China Region. In the previous section, we analyzed soil moisture data from SMOS for Gansu Province in 2012 and found that NU-CDF3 can achieve almost the

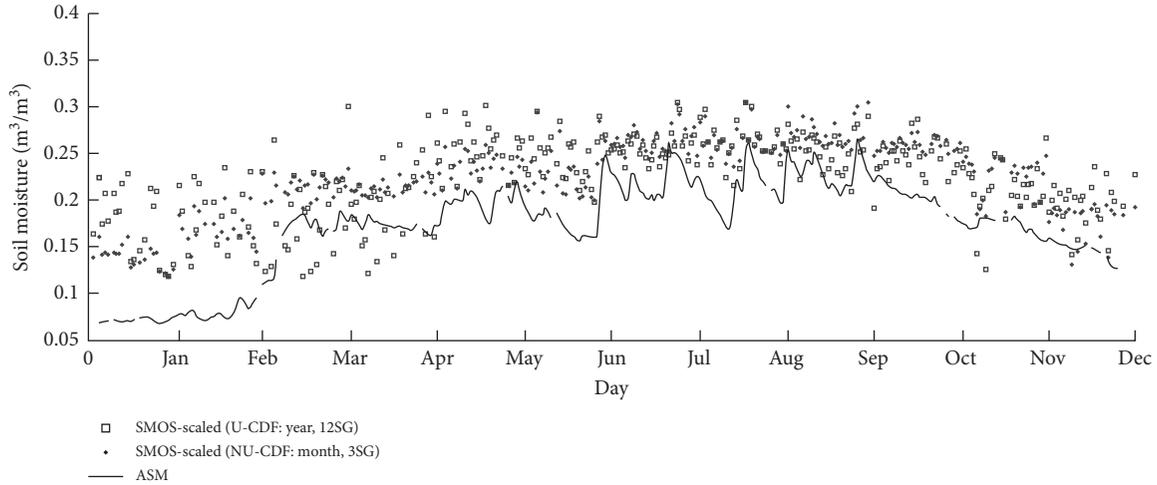


FIGURE 9: Time series of the ASM soil moisture data.

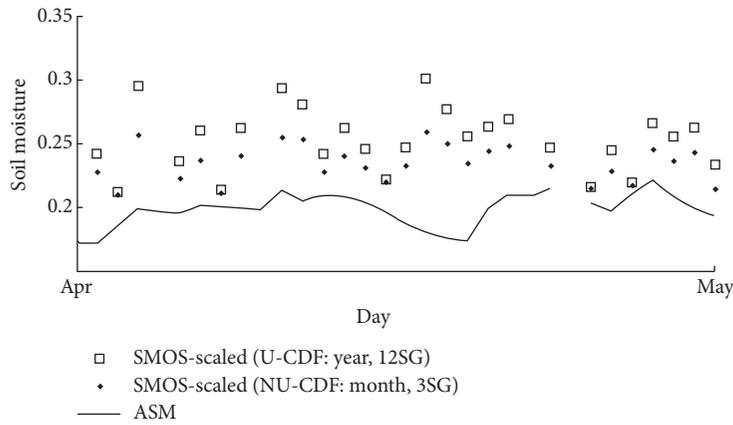
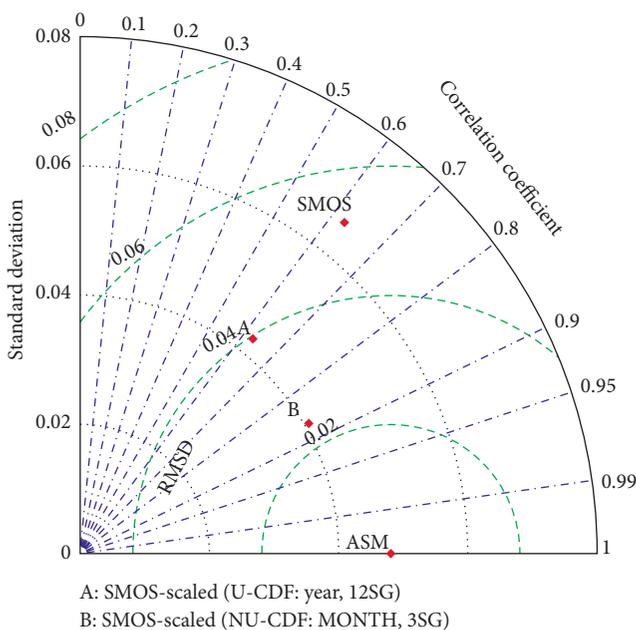


FIGURE 10: Time series of the soil moisture data from May.



A: SMOS-scaled (U-CDF: year, 12SG)
 B: SMOS-scaled (NU-CDF: MONTH, 3SG)

FIGURE 11: The SD, R, and RMSD of four datasets.

same SD reduction, can improve R , and can reduce the RMSD when compared with U-CDF12. In this section, we analyze the data for the whole China area from 376 automatic soil moisture stations using FY3B, SMOS, and WindSat. The results agree with our previous conclusions.

Figure 12 shows the automatic soil moisture station distribution in China. Each point in the map represents a station, and there are 376 in total. For each station, we converted the FY3B, SMOS, and WindSat soil moisture content data for 1 year so that it was consistent with the CLDAS model data, using U-CDF12- or NU-CDF3 matching methods. We then investigated the impact of the bias correction using the SD, R , and RMSD. First, consider the SD. Figure 13 shows the relationship between the SD and each station. The three subplots from top to bottom show the matching results using original soil moisture data from FY3B, SMOS, and WindSat, respectively. The horizontal axes represent the station numbers and the vertical axes represent the SDs. The SDs of the original satellite data are represented by bold gray lines, and the NU-CDF3-scaled satellite data are represented by black lines. NU-CDF3 reduced the SD in most stations. Consider Figure 13(a) as an example. It contains a plot of the SDs of the actual FY3B

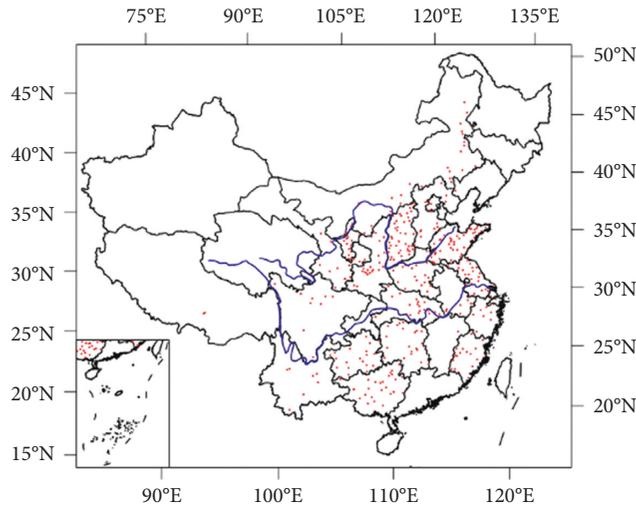


FIGURE 12: Automatic soil moisture stations in China.

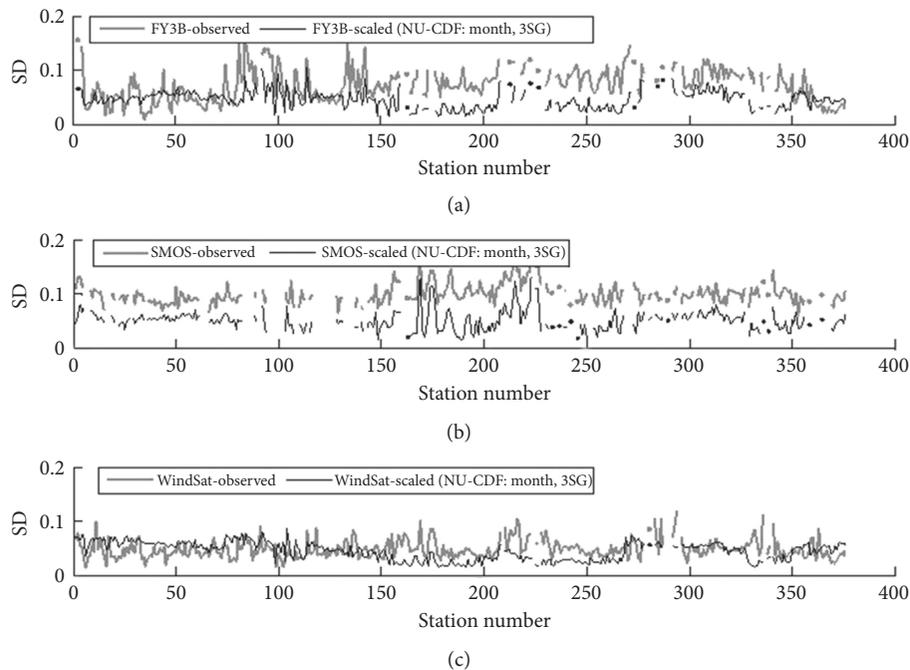


FIGURE 13: SD of the original satellite soil moisture data and the NU-CDF3-scaled data from each station.

soil moisture data and the NU-CDF3 scaled FY3B soil moisture data. For a better illustration, we should plot two other sets of data: the ASM data and the U-CDF12 scaled data. However, it is difficult to show the infinitesimal differences between these four sets of data in one plot, so Figure 14(a) shows the two scaled FY3B datasets for a time period of 51 days. Squares represent the U-CDF12 matching results, and black dots represent the NU-CDF3 matching results. These three sets of SD data are very similar. Similarly, Figures 14(b) and 14(c) show the SDs from the SMOS and WindSat retrievals, which have the same characteristics. We conclude that NU-CDF3 using a monthly matching method has the same ability to reduce SD as the U-CDF12 using a year of data in the matching method.

Figure 15 displays the relationship between R and each station. The vertical axes represent R , the correlation coefficient, between one set of data and the ASM data. Note that R of the model data itself is always 1 (the maximum of the vertical axis). The black lines represent R for NU-CDF3, and the gray lines represent U-CDF12. The three different satellite datasets have a common characteristic: the R value of the NU-CDF3-scaled soil moisture is much closer to 1 than U-CDF12, which represents an improvement.

Figure 16 displays the relationship between the RMSD and each station. The situation is quite similar to our previous analysis of the correlation. In most stations, the RMSD of the NU-CDF3-scaled satellite data (gray line) is less than the U-CDF12-scaled satellite data (black line). It is obvious

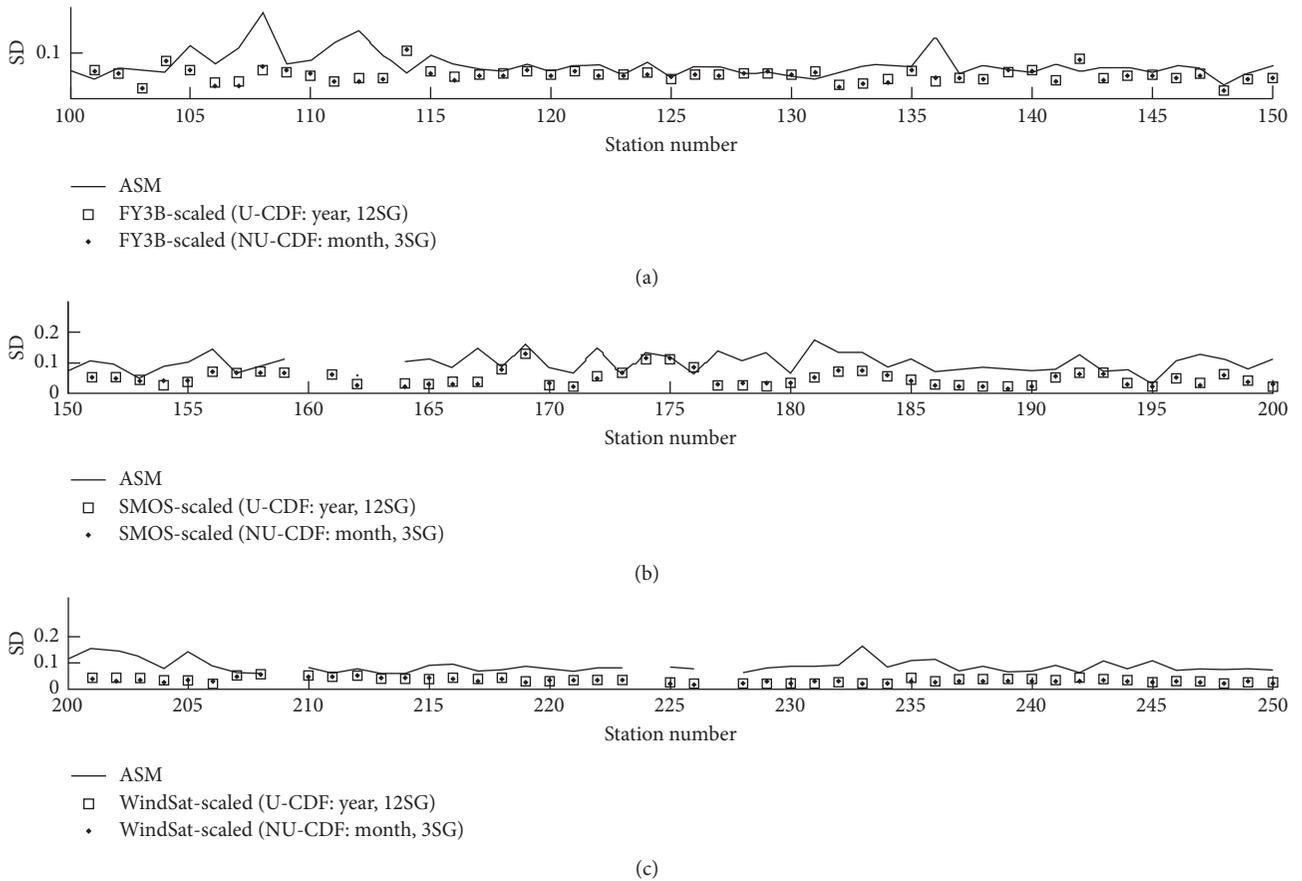


FIGURE 14: SDs of ASM, U-CDF12-scaled data, and NU-CDF3-scaled data from 51 automatic soil moisture stations.

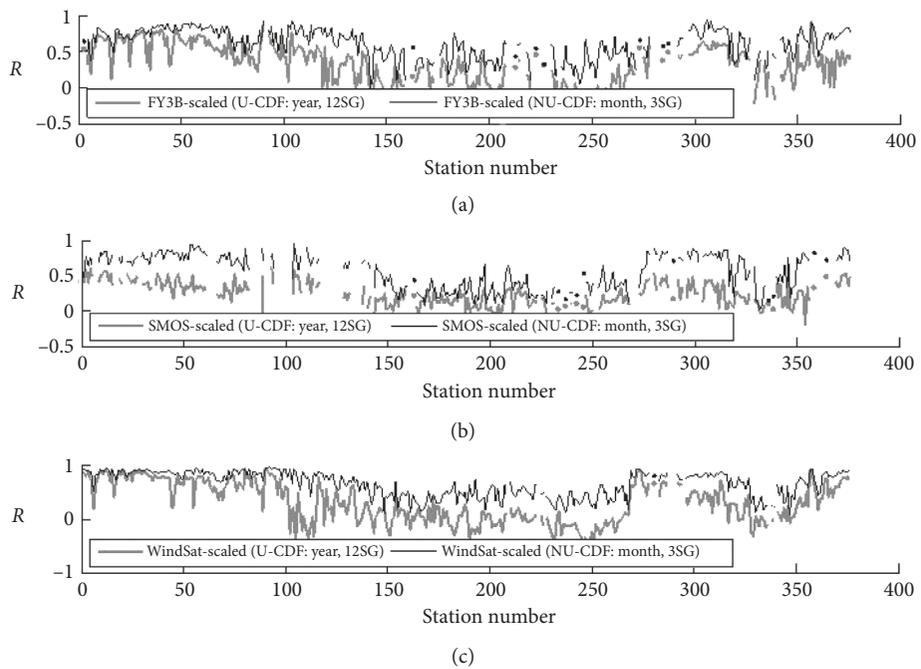


FIGURE 15: Correlation (R) between the scaled satellite data and the ASM data at each station.

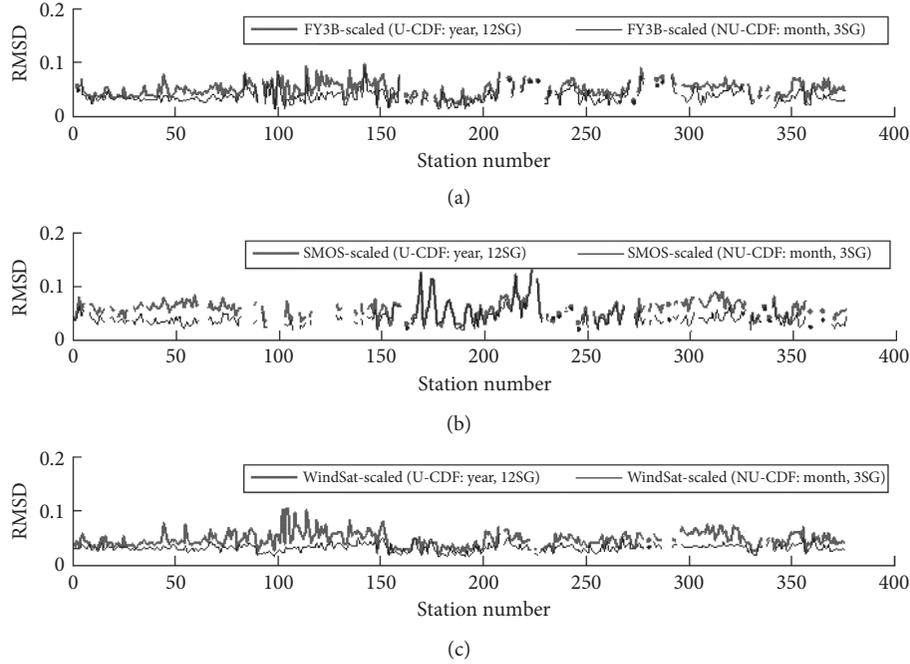


FIGURE 16: RMSD of the scaled satellite data when compared with the ASM.

TABLE 2: Statistical properties of the datasets. The improvement ratio was computed according to (7)–(9).

Satellite	SD			R			RMSD		
	FY3B	SMOS	WindSat	FY3B	SMOS	WindSat	FY3B	SMOS	WindSat
Valid stations	333	288	363	333	288	363	333	288	363
NU-CDF3 superior	266	231	274	235	220	275	245	229	280
Ratio (%)	79.88	80.21	75.48	70.57	76.39	75.76	73.57	79.51	77.13
Improvement ratio (%)	6.05	4.89	4.37	9.66	14.48	14.05	4.69	7.13	6.19

that the RMSD of the original satellite soil moisture data has been reduced by the NU-CDF3 matching method.

To quantitatively analyze the “superiority” of NU-CDF3 for bias reduction, we calculated the number of automatic soil moisture stations where the satellite and CLDAS model all have valid estimates and the number of stations where the NU-CDF3 was better than U-CDF12 in terms of the three indicators. Table 2 displays these results. Row four shows that NU-CDF3 performed better for 70% to 80% of the stations in terms of SD, R , and RMSD. We use A to represent the SD, the R , or the RMSD improvement from U-CDF12-scaled satellite data when compared with ASM and B to represent the improvement from the NU-CDF3-scaled satellite data. Then A and B are defined as

$$A = \sum_{i=1}^N \text{abs}[P_{\text{U-CDF}}(i) - P_{\text{ASM}}(i)], \quad (8)$$

$$B = \sum_{i=1}^N \text{abs}[P_{\text{NU-CDF}}(i) - P_{\text{ASM}}(i)], \quad (9)$$

where P represents one of the three indicators. Take SD as an example. Then, P_{ASM} represents the SD of the ASM data at each station, $P_{\text{U-CDF}}$ is the SD of the U-CDF12-scaled

satellite data at each station, and $P_{\text{NU-CDF}}$ is the SD of the NU-CDF3-scaled satellite data at each station. N is the total number of automatic soil moisture stations.

We further define the improvement ratio for NU-CDF as

$$\eta = \frac{B - A}{B}, \quad (10)$$

where A and B denote that, regardless of the indicator being considered, the improvement always represents the distance between the scaled data indicator and the ASM indicator.

Table 2 displays the improvement ratio calculated according to the above equations. NU-CDF3 improved the SD of the FY3B data by 6.05%, the SD of the SMOS data by 4.89%, and the SD of the WindSat data by 4.37%. This implies that NU-CDF3 was slightly more effective than U-CDF12 at reducing the SD. The improvement ratios for R and RMSD are also shown in Table 2. NU-CDF3 matching improved R for the three satellite retrievals by over 9% and reduced the RMSD by over 4%.

4. Conclusions

We compared the effectiveness of two different CDF-sampling methods for bias correction: uniformly spaced

sampling and nonuniformly spaced sampling. When the CDF was computed from a year of satellite soil moisture data, three nonuniformly spaced samples reduced the standard deviation to the same extent as 12 uniformly spaced samples. We made use of the high temporal and spatial availability of ASM datasets by separately implementing CDF matching for each month of satellite data. The correlation has been significantly improved using this monthly three-segment NU-CDF matching method. Finally, we expanded the study area to cover all of China and analyzed the soil moisture data from FY3B, SMOS, and WindSat at 376 automatic soil moisture stations in 2012. In our results, NU-CDF3 reduced the SD, improved R , and reduced the RMSD in over 70% of the stations, when compared with U-CDF12. Additionally, the SD and RMSD have been reduced by over 4% with R improved by more than 9%.

Data Availability

The data used to support the findings of this study are included in the supplementary information files.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

The authors not only wish to acknowledge the National Meteorological Center to provide three satellite observation data but also thank the teacher Shi for her guidance and support. This work was supported in part by the Major Program of National Natural Science Foundation of China (no. 91437220), Natural Science Foundation of Jiangxi Province for Youths (no. 20171ACB21038), and Jiangxi Municipal Science and Technology Project.

Supplementary Materials

The supplementary materials include the three satellite observation data (FY3B, SMOS, and WindSat) and 376 automatic soil moisture stations (ASM, 0–10 cm) daily data in 2012, and all data are arranged according to the stations. (*Supplementary Materials*)

References

- [1] M. Choi and J. Jacobs, “Temporal variability corrections for advanced microwave scanning radiometer E (AMSR-E) surface soil moisture: case study in Little River Region, Georgia, U.S.,” *Sensors*, vol. 8, no. 4, pp. 2617–2627, 2008.
- [2] K. H. Lee and E. N. Anagnostou, “A combined passive active microwave remote sensing approach for surface variable retrieval using tropical rainfall measuring mission observations,” *Remote Sensing of Environment*, vol. 92, no. 1, pp. 112–125, 2004.
- [3] D. Atlas, D. Rosenfeld, and D. B. Wolff, “Climatologically tuned reflectivity rain rate relations and links to area time integrals,” *Journal of Applied Meteorology*, vol. 29, no. 11, pp. 1120–1135, 1990.
- [4] Y. Y. Liu, A. I. J. M. van Dijk, R. A. M. de Jeu, and T. R. H. Holmes, “An analysis of spatiotemporal variations of soil and vegetation moisture from a 29-year satellite-derived data set over mainland Australia,” *Water Resources Research*, vol. 45, no. 7, article W07405, 2009.
- [5] Z. Yuan, S. Chunxiang, S. Runping et al., “Quality evaluation of multi-microwave remote sensing soil products over China,” *Journal of the Meteorological Science*, vol. 35, no. 3, pp. 289–296, 2015.
- [6] R. H. Reichle and R. D. Koster, “Bias reduction in short records of satellite soil moisture,” *Geophysical Research Letters*, vol. 31, no. 19, article L19501, 2004.
- [7] R. Parinussa, G. Wang, and R. de Jeu, “Soil moisture from the recent AMSR2 and FY3B multi-frequency passive microwave sensors,” in *Proceedings of the Geophysical Research Abstracts, EGU2014–3335*, vol. 16, EGU General Assembly, Vienna, Austria, April–May 2014.
- [8] Y. Kerr, P. Waldteufel, J. P. Wigneron et al., “The SMOS mission: new tool for monitoring key elements of the global water cycle,” *Proceedings of the IEEE*, vol. 98, no. 5, pp. 666–687, 2010.
- [9] D. J. Leroux, Y. H. Kerr, E. F. Wood, A. K. Sahoo, R. Bindlish, and T. J. Jackson, “An approach to constructing a homogeneous time series of soil moisture using SMOS,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 393–405, 2012.
- [10] R. M. Parinussa, T. R. H. Holmes, and R. A. M. de Jeu, “Soil moisture retrievals from the Windsat spaceborne polarimetric microwave radiometer,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 7, pp. 2683–2694, 2012.
- [11] Y. Y. Liu, R. M. Parinussa, W. A. Dorigo et al., “Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals,” *Hydrology and Earth System Sciences*, vol. 15, no. 2, pp. 425–436, 2011.
- [12] M. T. Goodrich, “Efficient piecewise-linear function approximation using the uniform metric: (preliminary version),” in *Proceedings of the Tenth Annual Symposium on Computational Geometry*, pp. 322–331, ACM Press, Stony Brook, NY, USA, June 1994.
- [13] W. Chan and F. Chin, “On approximation of polygonal curves with minimum number of line segments or minimum error,” *International Journal of Computational Geometry and Applications*, vol. 6, no. 1, pp. 59–77, 1996.
- [14] D. Z. Chen and O. Daescu, “Space-efficient algorithms for approximating polygonal curves in two dimensional space,” *International Journal of Computational Geometry and Applications*, vol. 13, no. 2, pp. 95–111, 2003.
- [15] D. H. Douglas and T. K. Peucker, “Algorithms for the reduction of the number of points required to represent a digitized line or its caricature,” *Cartographica: The International Journal for Geographic Information and Geovisualization*, vol. 10, no. 2, pp. 112–122, 1973.
- [16] J. Hersherberger and J. Snoeyink, “Speeding up the Douglas-Peucker line-simplification algorithm,” in *Proceedings of the 5th International Symposium on Spatial Data Handling*, pp. 134–143, Charleston, SC, USA, August 1992.
- [17] P. L. Rosin, “Techniques for assessing polygonal approximations of curves,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 6, pp. 659–666, 1997.
- [18] K. E. Taylor, “Summarizing multiple aspects of model performance in a single diagram,” *Journal of Geophysical Research*, vol. 106, no. 7, pp. 7183–7192, 2001.



Hindawi

Submit your manuscripts at
www.hindawi.com

