Journal: ESSD

Title: AIMERG: a new Asian precipitation dataset (0.1°/half-hourly, 2000-2015) by calibrating GPM IMERG at daily scale using APHRODITE

Author(s): Ziqiang Ma et al.

MS No.: **essd-2019-250**

MS Type: Data description paper

General Comments:

The study focuses on the application of a new calibration approach, Daily Spatio-Temporal Disaggregation Calibration Algorithm (DSTDCA), to daily scale on the retrospective IMERG data using APHRODITE product during 2000 to 2015. The quality of the calibrated AIMERG precipitation is analyzed against observation data. The study contains useful and novel information, and is generally well written. I recommend it for publication after some minor revisions. I have only a few minor comments listed below. More specific comments:

Authors Response: All the coauthors greatly appreciate you for your final recommendation with "Publication after some minor revisions (review by editor)". Though minor revision is needed, the first author and the co-authors have paid great attentions on each bullet pointed out by you, which greatly improved the quality of this manuscript. Based on the comments from you and the other reviewer, we have made careful modifications on the original manuscript.

Below, the original comments are in **black**, our responses are in **blue**, and our changes in manuscript are in **red**.

Point 1:

Referee Comments: Section 1 (Introduction): It would be good to add a review section on the calibration approaches that has been used in previous studies in this section.

Authors Response: A very constructive suggestion. Driven by this point, we have carefully reviewed the development of the calibration approaches, and then added a new review section the third paragraph in the **Introduction**.

Author's changes in manuscript: we have added a review paragraph in the third paragraph in the **Introduction** in lines from 94 to 121. The content of the review paragraph is as follows: "Satellite-based precipitation products have significant advantages in detecting the variations of precipitation at fine spatio-temporal resolutions, especially over the poorly gauged regions. However, as the indirect estimates of precipitation, satellite-based precipitation products are inherently containing regional, seasonal, and diurnal systematic biases and random errors (Ebert et al., 2007), which could be effectively alleviated by anchoring the satellite-only precipitation products using gauge-based observations (Huffman et al., 2007). Therefore, great efforts have been taken on exploring the calibrations on the satellite-only precipitation estimates using gauge analysis. Historically, GPCP has provided the lion's share of the early efforts in the process of developing calibration algorithms for the satellite-only precipitation estimates in generating Satellite-Gauge products (2.5°/monthly). For instance, to correct the bias of the multi-satellite only estimates (mainly based on PMW and IR data) on a regional scale, the multi-satellite estimate was firstly multiplied by the ratio of the large-scale (with moving window size 5 × 5) average gauge analysis to the large-scale average of the multi-satellite estimate, and then the satellite-gauge (SG) estimate was finally derived by combining the gauge-adjusted multi-satellite estimate and the gauge analysis with inverse-error-variance weighting (Huffman et al., 1997; Adler et 2003; Adler 2018). Recently, a two-step strategy was proposed to remove the bias inherent in the multi-satellite only precipitation estimates using the probability density function (PDF) matching method and to combine the bias-corrected estimates with the gauge analysis using the optimal interpolation (OI) algorithm (Xie and Xiong, 2011; Shen et al., 2014). And a similar improved PDF algorithm was applied to generate the GSMaP data, which was adjusted at the daily scale by the gauge analysis (0.5°/daily) from the climate prediction center (CPC) (Mega et al., 2014). While GPM IMERG adjusted the multi-satellite precipitation estimates (0.1°/half hourly) at the monthly scale using the ratios between the original monthly multi-satellite only and the monthly SG data, in the combination with the original monthly multi-satellite only and GPCC (1.0°), in the month (Huffman et al., 2019). There is still much room for exploring the improved algorithms for calibrating the multi-satellite-only precipitation estimates at finer spatiotemporal scales, e. g, 0.25°/daily, which is also one of the next vital focuses by the GPM (Huffman et al., 2019)."

Point 2:

Referee Comments: Line 215: de-capitalize "Final"

Authors Response: Good idea, we have revised this error and checked such errors throughout the manuscript.

Author's changes in manuscript: we have de-capitalize "Final" as “final” in line 246.

Point 3:

Referee Comments: Table 1: add horizontal lines to avoid confusion

Authors Response: Good idea, adding the horizontal lines is greatly helpful in making it more clearly.

Author's changes in manuscript: we have added the horizontal lines to in Table 1, in lines from 200 to 202. The revised Table 1 shown as follows:

Table 1. List of satellite-based, gauge-based, and satellite-gauge combination precipitation products used in this study.

Short name	Full name	Spatial and temporal sampling	Time period	References
IMERG	Integrated Multi-satellitE Retrievals for Global Precipitation Measurement	0.1°/half-hourly	2000-present	Huffman et al. (2019) https://pmm.nasa.gov/data-access/downloads/gpm (last access: 17 January 2020)
APHRODITE	Asian Precipitation Highly Resolved Observational Data Integration Towards Evaluation of Water Resources	0.25°/daily	1951-2015	Yatagai et al. (2012) http://aphrodite.st.hirosaki-u.ac.jp/download/ (last access: 17 January 2020)
CMPA	China Merged Precipitation Analysis	0.1°/hourly	2008-present	Shen et al. (2014) http://data.cma.cn (last access: 17 January 2020)
	Point-based rain gauge data	hourly	2010-present	Shen et al. (2010) http://data.cma.cn (last access: 17 January 2020)

Point 4:

Referee Comments: Table 2: add a column of value ranges for the metrics considered

Authors Response: Good idea, we have added a column of value ranges for the metrics considered, and also the horizontal lines suggested by **Point 3**.

Author's changes in manuscript: we have added a column of value ranges for the metrics considered, and also the horizontal lines in lines from 261 to 265. The revised Table 2 is shown as follows:

Table 2 Formulas and perfect values of the evaluation metrics used in this study[a].

Statistic metrics	Equation	Perfect value	Value ranges
Correlation Coefficient (CC)	$CC = \frac{\frac{1}{N}\sum_{n=1}^{N}(S_n - \bar{s})(G_n - \bar{G})}{\sigma_S \sigma_G}$	1	[-1, 1]
Mean Error (ME)	$ME = \sum_{n=1}^{N}(S_n - G_n)$	0	$(-\infty, +\infty)$
Relative Bias (BIAS)	$BIAS = \frac{\sum_{n=1}^{N}(S_n - G_n)}{\sum_{i=1}^{n} G_n} \times 100\%$	0	$(-\infty, +\infty)$
Root Mean Square Error (RMSE)	$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(S_n - G_n)^2}$	0	$[0, +\infty)$
Probability of Detection (POD)	$POD = \frac{n_{11}}{n_{11} + n_{01}}$	1	[0, 1]
False Alarm Ratio (FAR)	$FAR = \frac{n_{10}}{n_{11} + n_{10}}$	0	[0, 1]
Critical Success Index (CSI)	$CSI = \frac{n_{11}}{n_{11} + n_{10} + n_{01}}$	1	[0, 1]

Point 5:

Referee Comments: Figure 4: Use the exact boundary of each subregion if the green boxes are not the exact boundary. How the subregion boundary defined and by what criteria?

Authors Response: Actually, we have used the exact boundary of each sub-region. To make it clearer, we have added the exact boundary information in this revised manuscript. For deciding the sub-regions, we have mainly considered three aspects: the representative climatic zones in China, the local spatial distributions of the gauge stations, and the complexity of the topography. For instances, Sub-Region 1 represents the high latitude plain in the most north-eastern region of China under a cold climate (left top: 115.0° E, 54.0°N; right bottom: 135.0° E, 47.0°N); Sub-Region 2 represents the south-eastern coastal area of China influenced greatly by the Asian Monsoons (left top: 115.0° E, 26.0°N; left bottom: 119.0° E, 24.0°N; right bottom: 124.0° E, 31.0°N; right top: 120.0° E, 34.0°N); Sub-Region 3 represents the most southern region including the island Hainan in the tropical zone (left top: 105.0° E, 24.0°N; right bottom: 115.0° E, 18.0°N); Sub-Region 4 represents the inner area of China covering the Yunnan-Kweichow Plateau and Sichuan Basin,

under a humid inland climate (left top: 100.0° E, 33.0°N; right bottom: 107.0° E, 27.0°N); Sub-Region 5 represents the most southern Tibet Plateau along the Himalayas with complex terrains and high elevations above ~ 4000.0 meters (left top: 80.0° E, 33.0°N; right bottom: 95.0° E, 27.0°N); Sub-Region 6 represents the central Asia with complex terrains covering the entire Tianshan Mountains in China under an arid inland climate (left top: 80.0° E, 45.0°N; right bottom: 92.0° E, 40.0°N).

Author's changes in manuscript: we have added the reasons and the criteria for selecting the sub-regions by providing the accurate boundary information in the Section 4.2 in lines from 303 to 316, as well as in the caption of Figure 4 in lines from 320 to 325.

(1) The added content of the reasons and the criteria for selecting the sub-regions and the accurate boundary information, in lines from 303 to 316, are shown as follows: "For deciding the sub-regions (Fig. 4 d), we have mainly considered three aspects: the representative climatic zones in China, the local spatial distributions of the gauge stations, and the complexity of the topography. For instances, Sub-Region 1 represents the high latitude plain in the most north-eastern region of China under a cold climate (left top: 115.0° E, 54.0°N; right bottom: 135.0° E, 47.0°N); Sub-Region 2 represents the south-eastern coastal area of China influenced greatly by the Asian Monsoons (left top: 115.0° E, 26.0°N; left bottom: 119.0° E, 24.0°N; right bottom: 124.0° E, 31.0°N; right top: 120.0° E, 34.0°N); Sub-Region 3 represents the most southern region including the island Hainan in the tropical zone (left top: 105.0° E, 24.0°N; right bottom: 115.0° E, 18.0°N); Sub-Region 4 represents the inner area of China covering the Yunnan-Kweichow Plateau and Sichuan Basin, under a humid inland climate (left top: 100.0° E, 33.0°N; right bottom: 107.0° E, 27.0°N); Sub-Region 5 represents the most southern Tibetan Plateau along the Himalayas with complex terrains and high elevations above ~ 4000.0 meters (left top: 80.0° E, 33.0°N; right bottom: 95.0° E, 27.0°N); Sub-Region 6 represents the central Asia with complex terrains covering the entire Tianshan Mountains in China under an arid inland climate (left top: 80.0° E, 45.0°N; right bottom: 92.0° E, 40.0°N)."

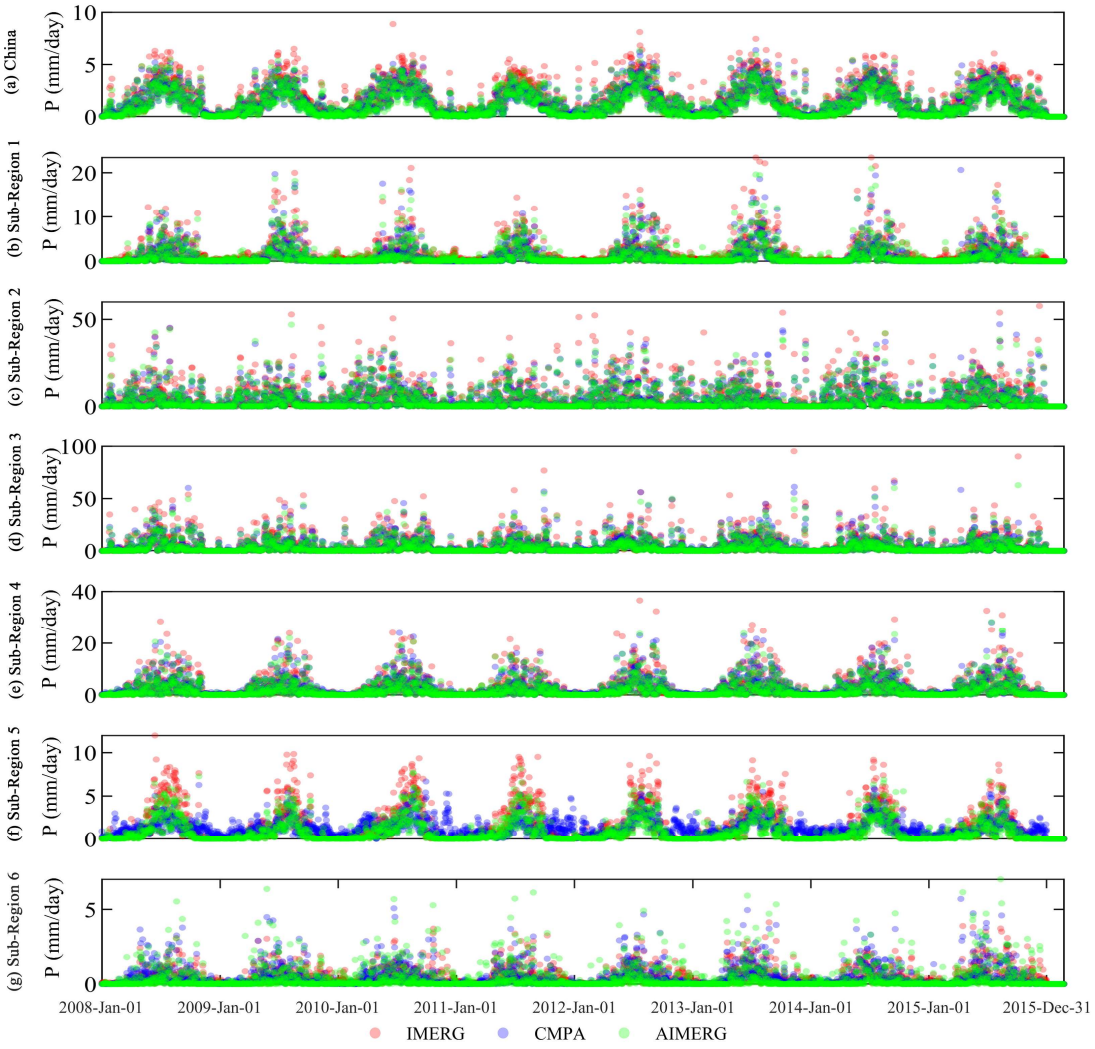
(2) The added content of the accurate boundary information in the caption of Figure 4, in lines from 320 to 325, are shown as follows: "Figure 4. The spatial patterns of (a) CMPA, (b) IMERG, and (c) AIMERG over China Mainland From 2008~2015, and (d) the spatial distributions of the ~ 50, 000 automatic meteorological stations in China Main land. The accurate boundary information

of the Sub-Regions: Sub-Region 1 (left top: 115.0° E, 54.0°N; right bottom: 135.0° E, 47.0°N); Sub-Region 2 (left top: 115.0° E, 26.0°N; left bottom: 119.0° E, 24.0°N; right bottom: 124.0° E, 31.0°N; right top: 120.0° E, 34.0°N); Sub-Region 3 (left top: 105.0° E, 24.0°N; right bottom: 115.0° E, 18.0°N); Sub-Region 4 (left top: 100.0° E, 33.0°N; right bottom: 107.0° E, 27.0°N); Sub-Region 5 (left top: 80.0° E, 33.0°N; right bottom: 95.0° E, 27.0°N); and Sub-Region 6 (left top: 80.0° E, 45.0°N; right bottom: 92.0° E, 40.0°N). The background map used in this study was provided by Esri, USGS and NOAA (http://goto.arcgisonline.com/maps/World_Terrain_Base, last access: 17 January 2020)."

Point 6:

Referee Comments: Figure 6: The presentation of the points are somewhat difficult to see. especially around the zeros.

Authors Response: A very constructive point for improving the quality of the Figure 6. We have paid great efforts on improving this figure, mainly considering the two aspects: (1) according to the value patterns of each dataset, CMPA, IMERG, and AIMERG, according to the Figure 5 at monthly scale, the sequence of the layers were adjusted from the sequence of CMPA, IMERG, and AIMERG to that of IMERG, CMPA, and AIMERG; and (2) the transparency of the plot in red, blue, and green, were all set as 50%, which made the figure greater clear to view the differences among the IMERG, CMPA, and AIMEG. The revised Figure 6 is shown as follows:



Revised Figure 6. The temporal patterns of mean areal precipitation of the IMERG, CMPA, and AIMERG, over China Mainland and sub-regions from 2008 to 2015, at daily scale.

Author's changes in manuscript: we have substituted the revised Figure 6, mentioned-above, for the original Figure 6 in line 355.

Point 7:

Referee Comments: Section 5 (Discussion): It would be good to discuss and quantitatively compare horizontally to other high resolution precipitation products that exists over the same area, using the same metrics evaluated.

Authors Response: A very constructive suggestion. To quantitatively and horizontally compare AIMEG with other high resolution precipitation products is greatly necessary to give the readers an overall view and understanding on the quality of the AIMERG. Recently, Tang et al (2020, *Remote*

Sensing of Environment) has conducted a comprehensive comparison of GPM IMERG with other nine state-of-the-art high resolution precipitation products, six satellite-based precipitation products (TRMM 3B42, 0.25°/3 hour; CMORPH, 0.25°/3 hour; PERSIAN-CDR, 0.25°/1 day; GSMaP 0.1°/1 hour; CHIRPS, 0.05°/1 day; SM2RAIN, 0.25°/1 day) and three reanalysis datasets (ERA5, ~0.25°/1 hour; ERA-Interim, ~0.75°/3 hour; MERRA2~0.5° × 0.625°/1 hour), from 2000 to 2018, and found that the IMERG product generally outperformed other datasets, except the Global Satellite Mapping of Precipitation (GSMaP), which was adjusted at the daily scale by the gauge analysis (0.5°/daily) from the CPC (Mega et al., 2014). Therefore, we have compared the AIMERG with GSMaP, in case of the typhoon, which was coordinated with those in Figure 11. As shown in Figure S1, though the spatial patterns of the GSMaP are similar with those of the AIMERG, the AIMERG provides much more details than GSMaP, especially over the northeastern Zhejiang Province. Meanwhile, AIMERG significantly overwhelms GSMaP in terms of both bias and random errors. For instance, GSMaP underestimates the precipitation (bias ~ -31%) twice as large as AIMERG (bias ~ -15%), and the random errors of GSMaP (MAE ~ 1.97 mm/hour, RMSE ~ 3.26 mm/hour) are also significantly larger than those of AIMERG (MAE ~ 1.44 mm/hour, RMSE ~ 2.50 mm/hour). Compared with the original IMERG, though the random errors of GSMaP are relatively larger, the bias of GSMaP (~ -31%) is significantly smaller than that of the original IMERG (~ -50%), which owes to the calibrations on the GSMaP at the daily scale. In future, we also encourage researchers to comprehensively evaluate and compare the AIMERG with other high resolution precipitation products at various spatio-temporal scales.

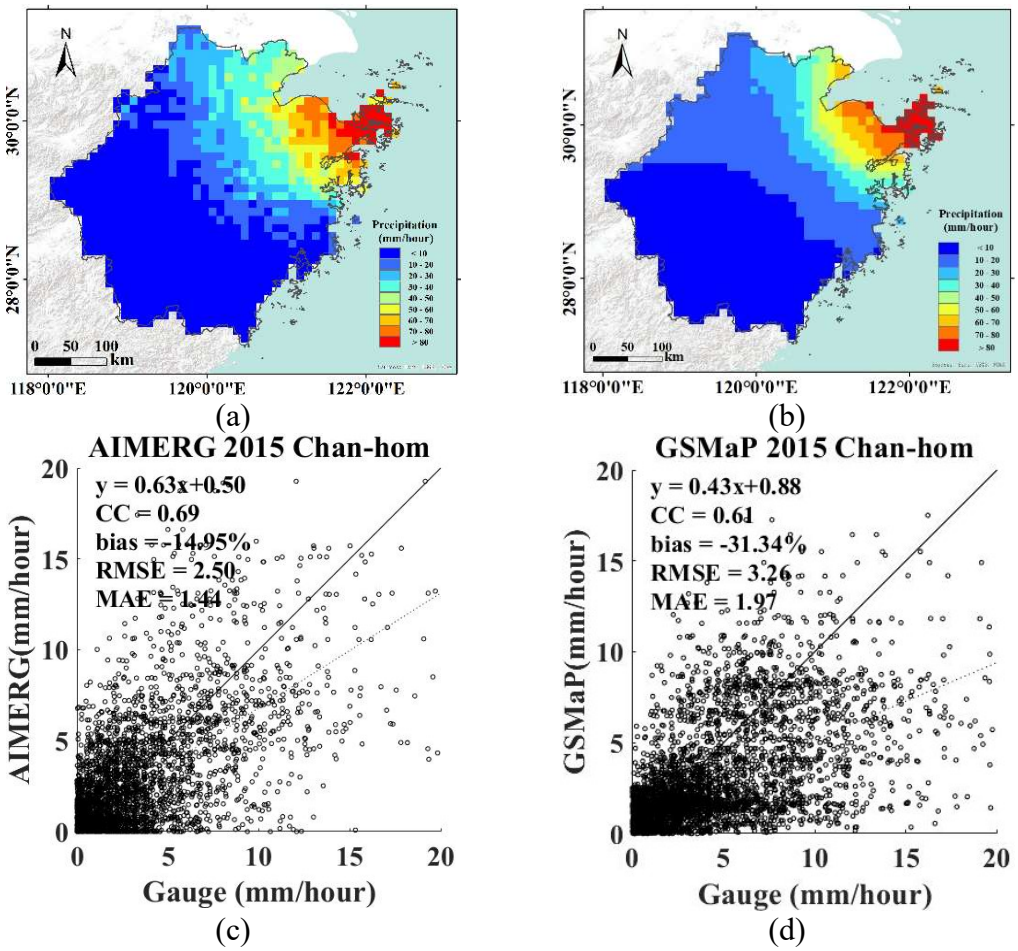


Figure S1. The typhoon, Chan-hom, is selected as an example for assessing the quality of the AIMERG and Gauge adjusted GSMaP, occurred in the typical period 0 a.m., – 11 a.m., June, 11, 2015, in Zhejiang Province.

Author's changes in manuscript: We have added a section to quantitatively and horizontally compare AIMEG with other high resolution precipitation products, GSMaP, in the **Discussion** section in lines from 530 to 554. And the content is shown as follows:

"**5.4. Quantitatively and horizontally comparisons with other high resolution precipitation product**

Recently, Tang et al (2020) has conducted a comprehensive comparison of GPM IMERG with other nine state-of-the-art high resolution precipitation products, six satellite-based precipitation products (TRMM 3B42, 0.25°/3 hour; CMORPH, 0.25°/3 hour; PERSIANN-CDR, 0.25°/1 day; GSMaP 0.1°/1 hour; CHIRPS, 0.05°/1 day; SM2RAIN, 0.25°/1 day) and three reanalysis datasets (ERA5, ~0.25°/1 hour; ERA-Interim, ~0.75°/3 hour; MERRA2~0.5° × 0.625°/1 hour), from 2000 to 2018, and found that the IMERG product generally outperformed other datasets, except the

Global Satellite Mapping of Precipitation (GSMaP), which was adjusted at the daily scale by the gauge analysis (0.5°/daily) from the CPC (Mega et al., 2014). Therefore, we have compared the AIMERG with GSMaP, in case of the typhoon Chan-hom, which is coordinated with those in Figure 11. As shown in Fig. 14, though the spatial patterns of the GSMaP are similar with those of the AIMERG, the AIMERG provides much more details than GSMaP, especially over the northeastern Zhejiang Province. Meanwhile, AIMERG significantly overwhelms GSMaP in terms of both bias and random errors. For instance, GSMaP underestimates the precipitation (bias ~ -31%) twice as large as AIMERG (bias ~ -15%), and the random errors of GSMaP (MAE ~ 1.97 mm/hour, RMSE ~ 3.26 mm/hour) are also significantly larger than those of AIMERG (MAE ~ 1.44 mm/hour, RMSE ~ 2.50 mm/hour). Compared with the original IMERG in Figure 11, though the random errors of GSMaP are relatively larger, the bias of GSMaP (~ -31%) is significantly smaller than that of the original IMERG (~ -50%), which owes to the calibrations on the GSMaP at the daily scale. In future, we also encourage researchers to comprehensively evaluate and compare the AIMERG with other high resolution precipitation products at various spatio-temporal scales.

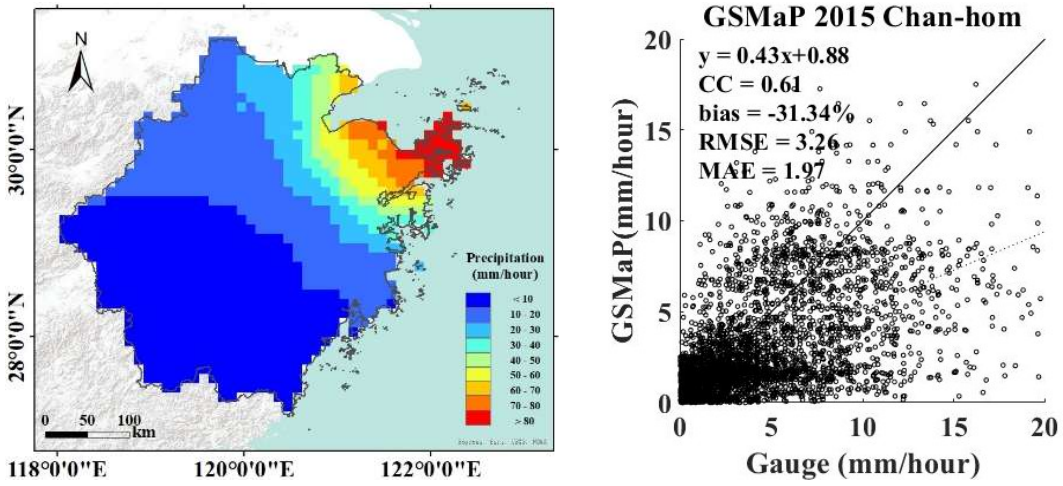


Figure 14. The typhoon, Chan-hom, is selected as an example for assessing the quality of the Gauge adjusted GSMaP, occurred in the typical period 0 a.m., – 11 a.m., June, 11, 2015, in Zhejiang Province, which is coordinated with those in Figure 11. The background map used in this study was provided by Esri, USGS and NOAA (http://goto.arcgisonline.com/maps/World_Terrain_Base, last access: 17 January 2020)."