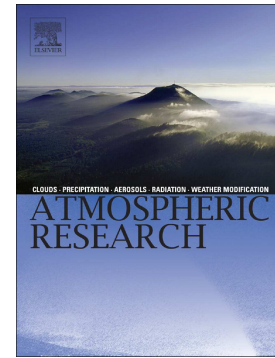


Accepted Manuscript

Multi time-scale evaluation of high-resolution satellite-based precipitation products over northeast of Austria

Ehsan Sharifi, Reinhold Steinacker, Bahram Saghafian



PII: S0169-8095(17)30712-3
DOI: doi:[10.1016/j.atmosres.2018.02.020](https://doi.org/10.1016/j.atmosres.2018.02.020)
Reference: ATMOS 4196
To appear in: *Atmospheric Research*
Received date: 4 July 2017
Revised date: 8 February 2018
Accepted date: 21 February 2018

Please cite this article as: Ehsan Sharifi, Reinhold Steinacker, Bahram Saghafian , Multi time-scale evaluation of high-resolution satellite-based precipitation products over northeast of Austria. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Atmos(2017), doi:[10.1016/j.atmosres.2018.02.020](https://doi.org/10.1016/j.atmosres.2018.02.020)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2018. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
<http://creativecommons.org/licenses/by-nc-nd/4.0/>

Multi Time-Scale Evaluation of High-Resolution Satellite-Based Precipitation Products over Northeast of Austria

Ehsan Sharifi ^{a,*}, Reinhold Steinacker ^a and Bahram Saghafian ^b

^a Department of Meteorology and Geophysics, University of Vienna, Austria; ehsan.sharifi@univie.ac.at , reinhold.steinacker@univie.ac.at

^b Department of Technical and Engineering, Science and Research Branch, Islamic Azad University, Tehran 1477893855, Iran; b.saghafian@gmail.com

* Corresponding author: ehsan.sharifi@univie.ac.at; Tel.: +43-1-427753737

Abstract: Over the years, combinations of different methods that use multi-satellites and multi-sensors have been developed for estimating global precipitation. Recently, studies that have evaluated Integrated Multi-satellite Retrievals for GPM (IMERG) Final-Run (FR) version V-03D and other precipitation products have indicated better performance for IMERG-FR compared to other similar products in different climate regimes. This study comprehensively evaluates the two GPM-IMERG products, specifically IMERG-FR and IMERG-Real-Time (RT) late-run, against a dense station network (62 stations) in northeast Austria from mid-March 2015 to the end of January 2016 using different time-scales. Both products are examined against station data in capturing the occurrence and statistical characteristics of precipitation intensity. With regard to probability density functions (PDFs), the satellite precipitation estimate (SPE) products have detected more heavy and extreme precipitation events than the ground measurements. Both precipitation products at all time-scales, except for IMERG-RT 12-hourly and daily precipitation, capture less occurrence of precipitation than the station dataset for light precipitation. This partially explains the under-detection of precipitation events. For all time-scales, both IMERG products' CDFs (Cumulative Distribution Function) are well above that of the stations' precipitation. For lower precipitation levels, IMERG-RT is slightly below the IMERG-FR whereas IMERG-RT is above IMERG-FR at higher precipitation levels. Furthermore, for entire spectrum precipitation rates ($P \geq 0.1$ mm), 1, 3, 6-hourly, IMERG-FR did not show a clear improvement of the Bias over IMERG-RT, while for 12-hourly and daily precipitation estimates, the bias in IMERG-FR has improved compared to IMERG-RT. In addition, IMERG-FR shows a considerable improvement in RMSE as compared to IMERG-RT. IMERG-FR, however, systematically underestimates moderate to extreme precipitation and overestimates light precipitation for all time scales against rain-gauges in northeast Austria. When comparing the bias, RMSE, and correlation coefficients, IMERG-FR has outperformed IMERG-RT particularly for 6-hourly, 12-hourly, and daily precipitation. Despite the general low probability of detection (POD) and threat score (TS) and the high false alarm ratio (FAR) within specified precipitation thresholds, the contingency table shows relatively acceptable values of the POD, TS and FAR for precipitation without classification.

Keywords: satellite precipitation products; GPM constellation satellite; IMERG; statistical analysis

Despite the limited data in the initial release of IMERG, results presented here show that systematic differences between the two products do exist and vary with precipitation rates (Liu, 2016). On the aggregate, it should be mentioned, we expect the direct inclusion of gauge analyses product (IMERG-FR) outperform the climatological adjustment product (IMERG-RT) while both use the same input satellite estimates. In addition, there will be differences in the statistics at the various time resolutions and precipitation classes since the monthly gauge adjustments are essentially being “decomposed” to higher temporal resolutions.

2. Study Area

To examine the accuracy of satellite-derived precipitation data in the center of Europe, we selected the northeast of Austria, because there is a rather high-density gauge network; moreover, in this area, both stratiform and convective precipitation can occur, and the precipitation pattern is not directly affected by altitude and the topography is moderate.

Austria is located in a temperate climatic zone, and due to the topographical diversity and the relatively large west-east extent there are three quite different climatic regions: the Alpine Region with alpine climate, the eastern part of the country has Pannonian climate with a continental influence and low precipitation, and the remainder of the country, referred to as transient climate is influenced by the Atlantic (in the west) and a continental-mediterranean influence in the southeast.

Altitude and distance from the rim determines the precipitation pattern; while as seen in Fig.1 and Fig. 2 high-level areas in the Alps may have a high average precipitation more than 2000 mm per year, some regions in the east and northeast of Austria have less than 600 mm annually (Hiebl et al., 2011). Moreover, according to the Köppen-Geiger Classification-updated by Rubel et. al. (Fig. 3), the climate of Austria can be classified as Cfb; and the climate of the Mountainous Regions can be classified as Dfb and Dfc (Rubel et al., 2017);

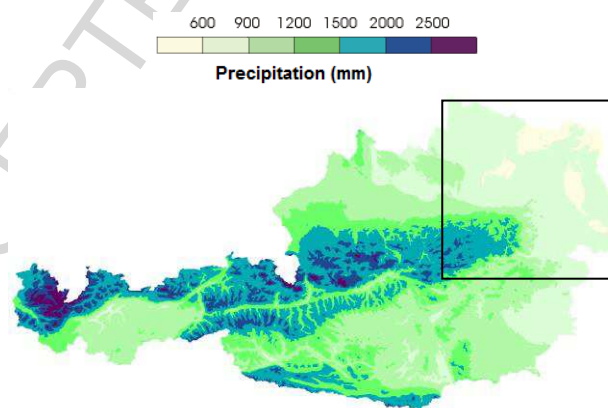


Fig 1. Mean annual precipitation of Austria (Hiebl et al., 2011)

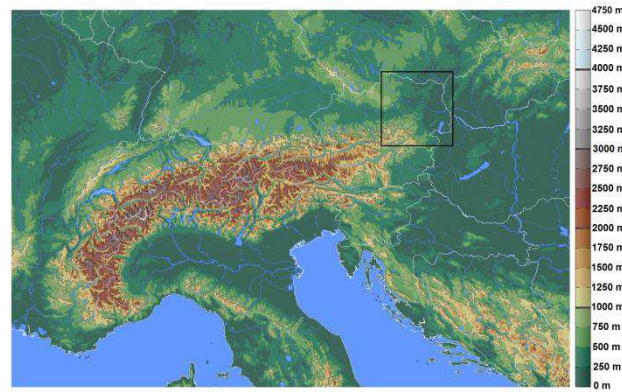


Fig 2. Elevation map of Austria

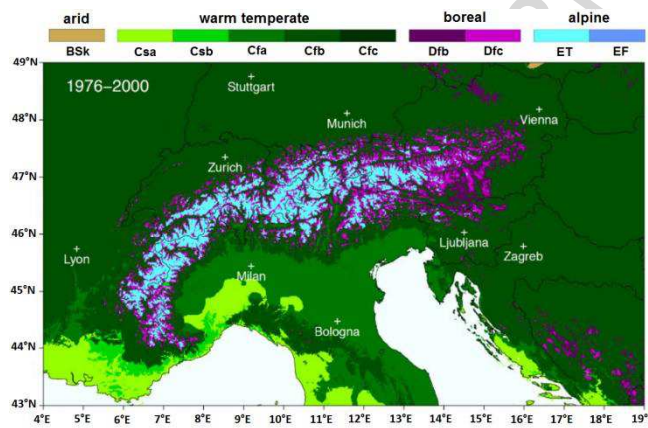


Fig 3. Climate of Austria based on Köppen climate classification-update by Rubel et. al. (2017)

3. Data and Methods

3.1. Data

a) Stations

In this research, 62 in-situ meteorological synoptic stations with 10-minute time interval observations are used, which have been provided by the Zentralanstalt für Meteorologie und Geodynamik (ZAMG)-Austria. ZAMG operates a network of semi-automatic stations across the country and provides quality-controlled precipitation data. The ZAMG stations have tipping bucket gauges and weighing rain-gauges which are equipped with a heating system (Haiden et al., 2011). Before evaluating these data, we used a second derivative filter to find unrealistic rapid and isolated changes in the time-series of precipitation (edges control). Such a derivative filter is very sensitive to noise, which should be eliminated before using the data for an objective analysis (Shapiro and Stochman, 2001). According to climatology data, the outliers are removed to avoid using these values of rain-gauge precipitation measurements. This affected only four out of the 419,890 hourly data.

b) IMERG-FR (post real-time) V03

The GPM mission's Precipitation Processing System (PPS) at NASA's Goddard Space Flight Center released the IMERG V03 data to the public in late February 2015. The data set includes precipitation rates since mid-March 2014 for IMERG-FR. Ongoing and forthcoming data sets are freely accessible to users from NASA Goddard Earth Sciences

Data and Information Services Center (GES DISC) website. Currently, the IMERG-FR V03 is available for the period between 12 March, 2014 and January 2016¹.

The IMERG product used in this study is the Level 3 multi-satellite precipitation algorithm of GPM which is intended to inter-calibrate, merge, and interpolate all constellation microwave sensors, IR-based observations from geosynchronous satellites, and monthly gauge precipitation data for the TRMM and GPM eras. The system runs several times for each time of observation, giving a quick measurement at first, and then successively providing better measurement as more data arrive. In the last step, IMERG-FR follows the TMPA approach for infusing monthly gauge information into the fine-scale precipitation estimates (Huffman et al., 2007). All of the full-resolution multi-satellite estimates are summed in a month to create a monthly multi-satellite-only field. This field is combined with the monthly Global Precipitation Climatology Centre (GPCC) precipitation rain-gauges (over land) to correct the bias of satellite retrievals and create research-level products, when available (Huffman et al., 2015a).

c) IMERG-RT late run V03

The data are available to users for the period between 7 March, 2015, and the present, thus providing IMERG-RT V03 data for the period 14 March, 2015, to January 2016. The IMERG-RT late run, which is the near real-time product in the IMERG series of products, presently runs about 16 hours after observation time. It has a climatological calibration to the monthly gauge data (unlike the Final-Run, which uses actual monthly gauge analyses).

It should be mentioned that further work is needed to evaluate the inter-annual variation of IMERG data relying on larger samples, as well as evaluate the newer versions of this product.

3.2. Methodology

The satellite products are compared with the measured precipitation data of 62 meteorological synoptic stations (Figure 4) at hourly, 3-hourly, 6-hourly, 12-hourly and daily time-scales for the northeast of Austria.

In terms of the independence of the comparison between gauge and satellite-based products, GPCC uses data from 51 Austrian synoptic stations over the northeast of the country, which accounts for a big part of the total gauges (82%) used in our study. In addition, GPM-IMERG research level product (IMERG-FR) is combined with GPCC gauge data at the monthly scale, while the comparison in our study is conducted at sub-daily and daily scales.

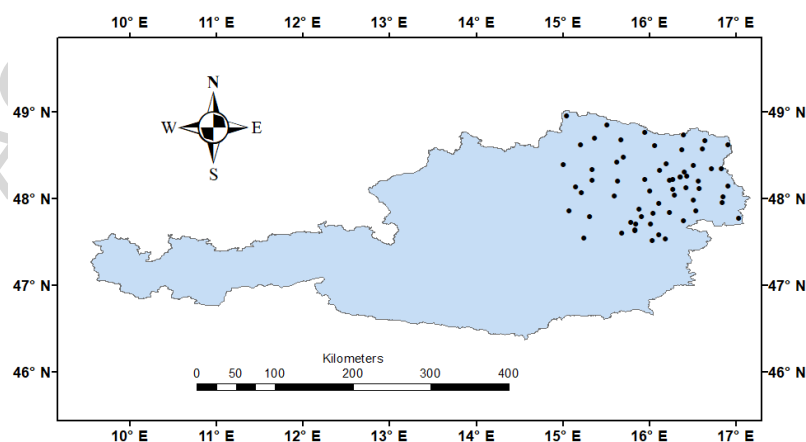


Fig 4. Map of meteorological synoptic stations distribution in northeast of Austria

¹ NASA. "GPM_3IMERGHH 03D". Available online: <http://disc.sci.gsfc.nasa.gov/SSW/> (accessed on 26 May 2016).

In this research, the data of a pixel of the satellite is compared with that corresponding to the ground point observation (i.e., the station). Only the cells where there is at least one reporting station can be selected for computation (Guo et al., 2015). In cases where the stations are within the pixels, the comparison is carried out directly between them. However, in cases where the ground station is close to the edge between two or close to the corner of four pixels (less than 0.01 degrees off the edges), an average of the two or four pixels around the station is used as the basis for comparison. Also, for a pixel with two or more rain-gauges, the areal-average precipitation is the arithmetic average of all rain-gauges located within that pixel. In addition, the statistical analysis based on regional events is done and applied to different time-scales such as hourly, 3-hourly, 6-hourly, 12-hourly and daily precipitation from meteorological synoptic stations of ZAMG. In total, there are 419,890 hourly data and 35,835 hours with precipitation available. When there is quantitative data with a wide range of possible values, it is beneficial to place these values into categories. In general, three different methods have been used to define different precipitation thresholds: (i) a relative threshold, quantiles, i.e., percentiles (ii) absolute thresholds; or (iii) return period (recurrence) values. As an example of the first case, a daily precipitation event with an amount greater than the 90th percentile of daily precipitation for all wet days can be considered as extreme (Simonović, 2012). To do that the percentile method is used and each class contains an admissible number of values.

Some researchers (Karl et al., 1995; Karl and Knight, 1998) observed a significant positive trend in the frequency of extreme rainfalls (greater than 50 mm per day) over the last few decades in the USA. For Australia, Suppiah and Hennessy (1996, 1998) showed a significant increase in the 90th and 95th percentiles, while other studies showed increases in the 99th percentile (Hennessy and Suppiah, 1999; Plummer et al., 1999). Variations in total precipitation can be caused by a change in the frequency of precipitation events, in the intensity of precipitation per event, or a combination of both. In order to ensure better understanding of precipitation process in comparison to satellite estimation, multi time-scale precipitation series must be analyzed. In particular, for the studies aimed at comprising stations and SPE products at varied classes which are referred to as light, medium, heavy, very heavy and extreme precipitation respectively, we consider the 50th, 70th, 90th, 98th percentile. Extreme events are particularly interesting as these events cause considerable damage and loss of life worldwide each year.

3.2.1. Probability Density Distribution

To understand the overall characteristics of precipitation, the precipitation frequency with different intensities is equally important as knowing the mean and spatial/temporal variation patterns of precipitation (Sohn et al., 2013; Tian et al., 2007). Because the same rainfall amount in the form of long-lasting light rain or a short-duration storm will yield quite different impacts in natural hazards, e.g., flood and landslide (Li et al., 2013; Shen et al., 2010). In this regard, a PDF can provide detailed information about the frequency of rainfall with different intensities.

In this study, a CDF is used to measure the number of observations that lie above or below a particular value in a data set. In other words, this is an indication of how often the satellite precipitation measurements are below or above the precipitation from stations (Gebere et al., 2015).

3.2.2. Statistical Analysis

In the first step, statistical indices such as bias, multiplicative bias (MBias), mean absolute error (MAE), root mean square error (RMSE), and linear correlation coefficient (CC) for different precipitation thresholds are determined.

3.2.3. Multi-category contingency table

Another assessment technique of satellite estimation is using a contingency table (see Table 2). In this way, a precipitation event is considered if at least one of the stations available in the region registered precipitation. Otherwise, if none of the stations registered precipitation, a ‘no-precipitation’ event is assigned to the whole area (Ballester and Moré, 2007). Both satellite and station data that are less than 0.1 mm are assigned to zero-precipitations, because our stations do not measure precipitation less than 0.1 mm. The statistical results are compiled and stratified by a number of criteria, including the number of time intervals (i.e., 1, 3, 6, 12 and 24 h) and particularly by precipitation thresholds. In this step analyses for observed precipitation below or above a particular threshold and satellite estimation below or above the same threshold are determined (Wilks, 2006).

Table 2. Contingency table

		Observed		
I / J		Yes	No	Total
Satellite	Yes	Hit (a)	False alarm (b)	a + b
	No	Miss (c)	Correct negative (d)	c + d
	Total	a + c	b + d	n = a+b+c+d

In order to apply these to non-probabilistic estimates that are not dichotomous, it is necessary to collapse the $I = J > 2$ contingency table into a series of 2×2 contingency tables. Each of these 2×2 tables is constructed, as indicated in Figure 5, by considering the estimated event in distinction to the complementary, not the forecast event. The advantage of this approach is that the nature of the estimation errors can more easily be diagnosed and the disadvantage is that it is more difficult to condense the results into a single number.

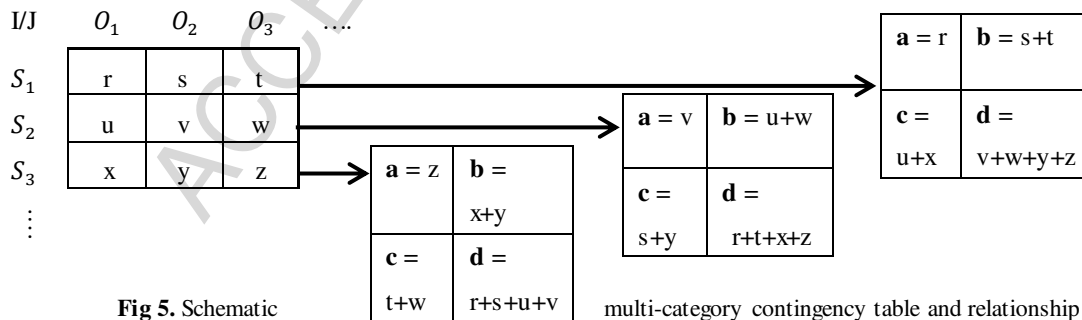


Fig 5. Schematic multi-category contingency table and relationship between counts (letters a–d) of satellite estimate/event pairs for the dichotomous non-probabilistic verification situation as displayed in a 2×2 contingency table

By using this table for different time-scale precipitation, a set of statistical indices are shown as follows (see equations 1-5):
POD responds to the question of what fraction of the observed “Yes” events is correctly estimated. The perfect score is 1.

$$POD = \frac{hits}{hits + misses} \quad (1)$$

FAR deals with the question of what fraction of the estimated “Yes” events did not occur. The ideal score is 0.

$$FAR = \frac{\text{false alarms}}{\text{hits} + \text{false alarms}} \quad (2)$$

TS is the fraction between the number of correct “Yes” estimated and the total number of times that event is observed. The perfect score is 1.

$$TS = \frac{\text{hits}}{\text{hits} + \text{misses} + \text{false alarms}} \quad (3)$$

Accuracy is the fraction of the total estimated events when the satellite products correctly estimated event and non-event; the perfect skill score is 1.

$$Accuracy = \frac{\text{hits} + \text{correct negatives}}{\text{total}} \quad (4)$$

The frequency bias (Fbias) answers the question of how the estimated frequency of “Yes” events compare to the observed frequency of “Yes” events. The Range of values is 0 to ∞ with a perfect score of 1 which means an unbiased estimation. In other words, the satellite estimates and observations have an occurrence above a given threshold the same number of times. Note that the bias provides no information about the correspondence between the individual forecasts and observations of the event on particular occasions, so that equation 5 is not an accuracy measure. An Fbias greater than unity indicates that the event is estimated more often than observed, which is called overestimated. Conversely, a bias less than unity indicates that the event is estimated less often than observed, or is underestimated.

$$Frequency\ bias = \frac{\text{hits} + \text{false alarms}}{\text{hits} + \text{misses}} \quad (5)$$

4. Results

4.1. Probability Density Distribution

With regard to PDFs, both satellite-based precipitation products (IMERG-FR and IMERG-RT) are examined against the reference data (stations) to capture the occurrence of precipitation and statistical characteristics of precipitation intensity in the northeast of Austria from March 2015 to the end of January 2016. The PDFs of hourly, 3-hourly, 6-hourly, 12-hourly and daily precipitation occurrences are computed as a ratio between the number of times the precipitation occurs inside each bin and the total number of times precipitation occurs overall. Then based on percentiles, precipitation intensities (P) are grouped into several bins for each time-scale with regard to the 50th, 70th, 90th, 98th percentile based on gauge data (Figure 6a-e).

The PDF of hourly precipitation with different intensities is shown in Figure 6-a. Both SPE products, IMERG-FR and IMERG-RT, capture less precipitating events than the reference dataset, principally between 0.1 to 0.4 mm/hr. For precipitation of 0.4 mm/hr and above, the trend is reversed, and more precipitation events are detected by both satellite products. In other words, the SPE products detected more heavy and extreme precipitation events than the ground measurements. This may shift the precipitation distribution spectrum to the higher intensity and cause great differences for the various applications, such as runoff production (Guo et al., 2015). IMERG-RT outperforms IMERG-FR for the frequency precipitation intensity in the range of 0.1 to 0.4 mm/hr when the precipitation rate is in the range of 1.1 to 1.8 mm/hr and 1.8 to 5.6 mm/hr. IMERG-RT shows the closest performance with stations. IMERG-FR (calibrated with monthly ground rain-gauges) performs even worse than its corresponding original satellite product (IMERG-RT). As demonstrated by the PDF of precipitation, the magnitude of the detected events is reduced at the expense of the missed events, skewing the intensity distribution (Tian et al., 2007). In fact, the gauge adjustment product (IMERG-FR) can modify the precipitation amounts but it cannot modify the occurrence of precipitation (Behrangi et al., 2014; Gosset et al., 2013). Both precipitation products in all time-scales, except for IMERG-RT in 12-hourly and daily precipitation,

capture less precipitating events than the station dataset for light precipitation. This partially explains the over-detection of no-precipitation events.

As seen in Figure 7.a, about 9% is reported by the station dataset in the frequency of precipitation events for an hourly time-scale. In general, both satellite-based precipitation products slightly tend to detect less precipitation events compared to the ground measurements. This indicates that satellite-based precipitation products missed some precipitation cases. Further details about all time-scales are found in Figure 7.a-e.

These results may be attributed to the following aspects. First, the satellite-based precipitation products, in particular, IMERG-RT, can detect reasonably heavy and convective precipitation events. Second, the bias correction procedures also boost the amplitude of events detected to compensate for the contribution on missed events (Tian et al., 2007).

Figure 8.a-e shows the CDF of precipitation between both IMERG-FR and IMERG-RT products against the stations. Figure 8.a for hourly time-scale shows that both IMERG products' CDFs are well above that of the stations' precipitation. However, the 94% frequency level for the ground data at 2.4 mm/hr corresponds to only an 88% frequency level for IMERG-FR and IMERG-RT. Below 2.4 mm/hr, the IMERG-RT frequency is slightly lower than of IMERG-FR whereas it is above IMERG-FR at higher precipitation levels.

Moreover, the finding for the 3-hourly precipitation is almost the same as for the hourly time-scale; but, the difference between them is that for the 3-hourly time-scale at 92% frequency level stations, precipitation is less than 4.5 mm/3hr. This is the case for only 88% of IMERG-FR and IMERG-RT. However, until 4.5 mm/3hr, the IMERG-RT frequency is slightly lower than the IMERG-FR frequency, and above this value, the IMERG-RT frequency is higher than of IMERG-FR. This indicates that the IMERG-FR observation is closer to the station values until a 4.5 mm/3hr precipitation rate is reached. This result shows an overestimated frequency of precipitation by both IMERG-FR and IMERG-RT (see Figure 8.b).

In order to evaluate the 6-hourly precipitation, as can be seen in figure 8.c, the result shows that the CDF of IMERG-FR is above that of the stations' precipitation at every point. However, IMERG-RT is above the station values until the 48% frequency level and almost at the same frequency level of ground data till the 77% frequency level when the precipitation rate is 2.7 mm/6hr. After this frequency level, IMERG-RT is again below the station value. This shows an overestimated frequency of precipitation by both IMERG-FR and IMERG-RT except for IMERG-RT in the medium precipitation range. Here it is more or less in the same frequency range as the stations are.

As shown in Figure 8.d for the 12-hourly time-scale, the CDF of IMERG-FR is above that of the stations' precipitation whereas the IMERG-RT is above the ground data until the 31% frequency level. This corresponds to a precipitation of less than 0.47 mm/12hr. For higher precipitation intensities in the range of heavy precipitation, the IMERG-RT is below the frequency of stations at the level of 78% frequency. This corresponds to a ground precipitation of less than 4 mm/12hr. It dips below the frequency of IMERG-FR at the 85% frequency level, corresponding to a precipitation of less than 7 mm/12hr. Furthermore, an overestimation of the frequency of IMERG-FR in comparison to IMERG-RT can be observed when the precipitation rate is lower than 4.5 mm/12hr and there is an underestimation above a precipitation rate of 4.5 mm/12hr.

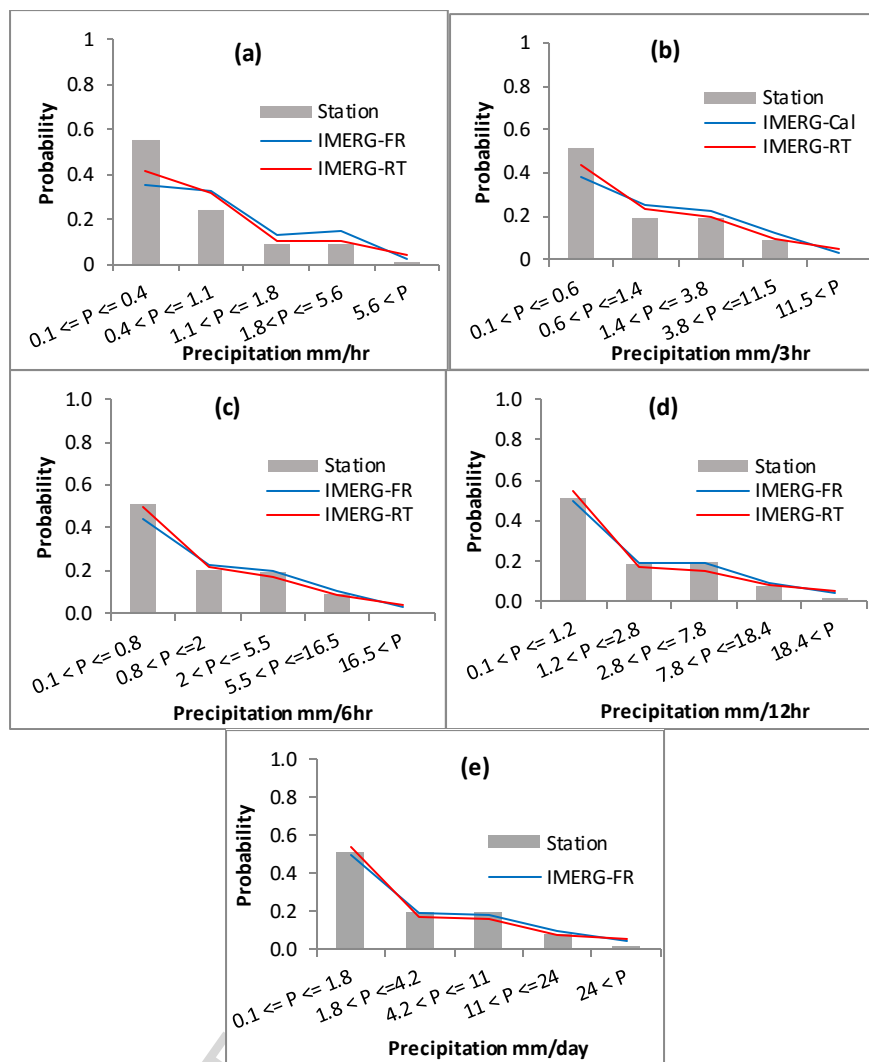
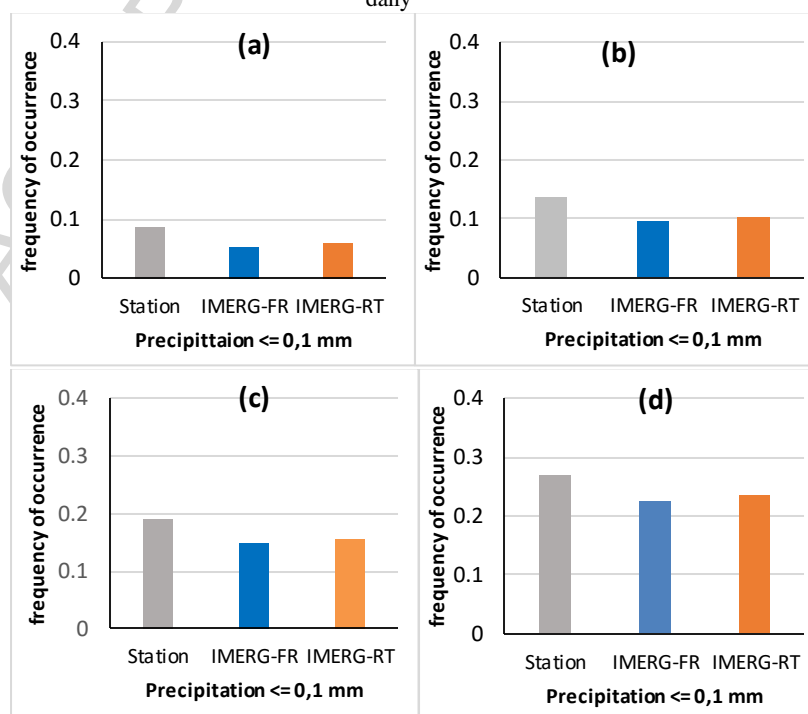


Fig 6. PDFs of precipitation events with different intensities a) hourly b) 3-hourly c) 6-hourly d) 12-hourly and e) daily



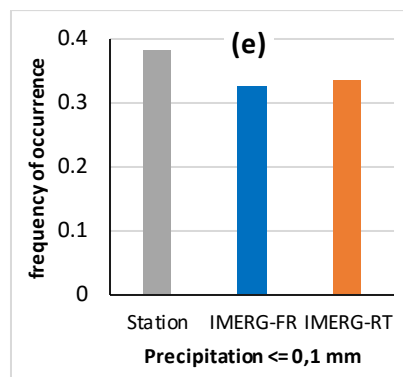


Fig 7. Fraction of detected precipitation events a) hourly b) 3-hourly c) 6-hourly d) 12-hourly and e) daily

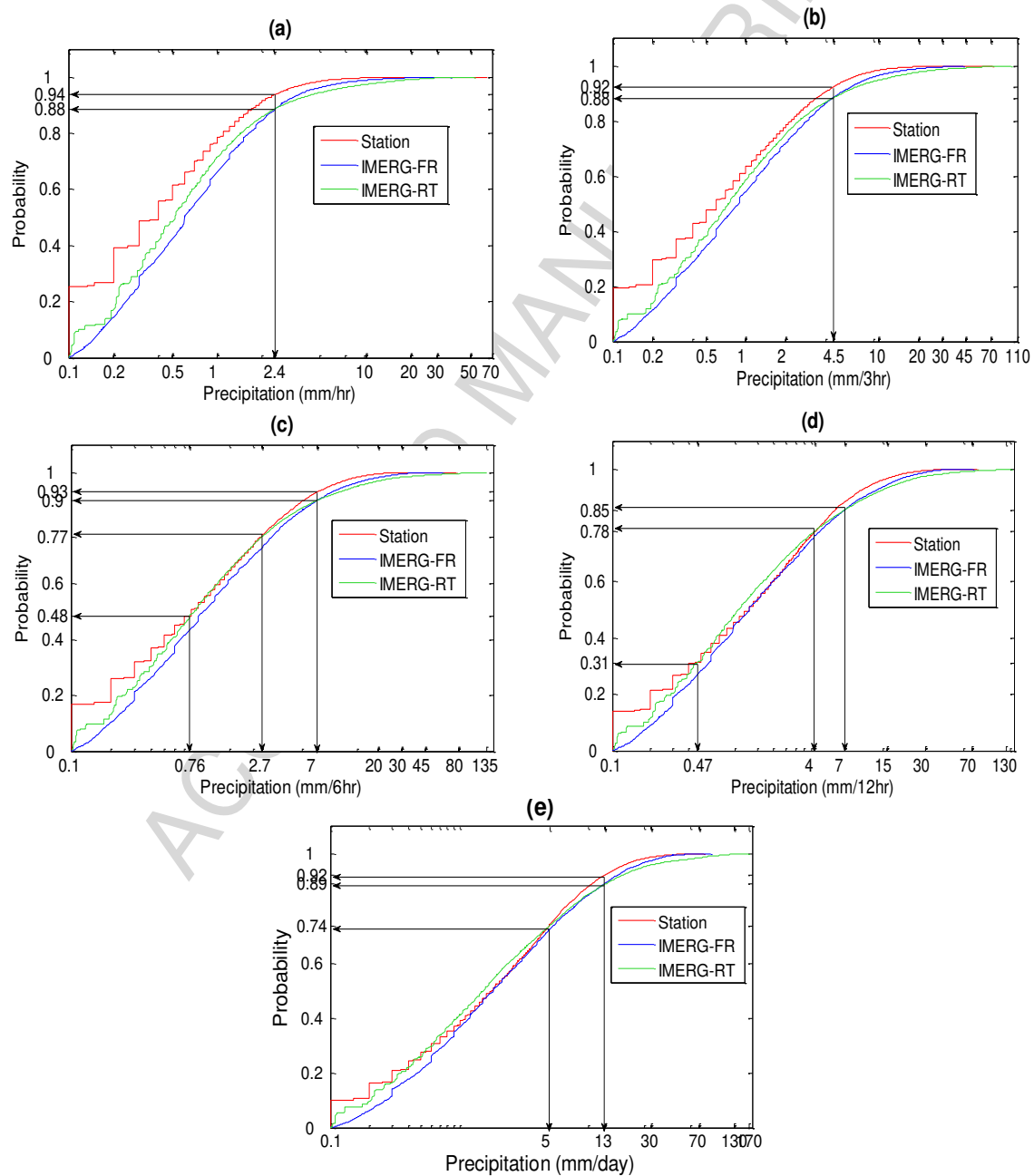


Fig 8. CDFs of precipitation events with different intensities a) hourly b) 3-hourly c) 6-hourly d) 12-hourly and e) daily

Figure 8.e shows that for daily precipitation the IMERG-FR product's CDF is above that of the stations' precipitation whereas the IMERG-RT is above only until the 74% frequency level when station data are less than 5 mm/day. For higher precipitation intensities in the range of heavy precipitation, the IMERG-RT drops below the frequency of stations at the level of 74%, corresponding to ground precipitation less than 5 mm/day. It even dips below the frequency of IMERG-FR at the 89% frequency level, corresponding to a precipitation of less than 13 mm/day. In other words, IMERG-FR underestimates the frequency of precipitation in comparison to IMERG-RT when the precipitation rate is above 13 mm/day.

In general, with respect to CDFs of 3-hourly, 6-hourly and 12-hourly precipitation, the CDFs of IMERG-RT is slightly lower than the IMERG-FR whereas IMERG-RT is above IMERG-FR at higher precipitation levels.

It should be noted that the results in Figures 6-8 are according to all stations and they are not based on an individual station.

4.2. Statistical Analysis

To evaluate the IMERG precipitation products comprehensively, ten metrics are selected. These can be divided into three categories: the first category includes the CC, describing the agreement between satellite estimates and gauge observations; the second category includes the MAE, the MBias and the RMSE, which are used to describe the error and bias of satellite estimates compared with gauge observations; and, the third category includes the POD, the FAR, TS, accuracy and frequency Bias, which are used to describe the contingency of SPEs (Yong et al., 2010). All metrics are illustrated in table A.a-e.

In table A, the statistical summary of the metrics for IMERG-FR and IMERG-RT products at hourly, 3-, 6-, 12-hourly and daily resolution at different thresholds over northeast Austria is shown. According to the results (Table A.a), precipitation shows a weak correlation to both products at the hourly time-scale for all precipitation classifications. Only precipitation without classification ($P \geq 0.1$ mm/hr) yields values of 0.30 and 0.29 for IMERG-FR and IMERG-RT respectively, showing a better agreement than for all other precipitation classes. At a 3-hourly time-scale (Table A.b), a correlation value of 0.46 and 0.40 for IMERG-FR and IMERG-RT respectively in the range of equal or above 0.1 mm/3hr is documented while in other classified precipitation values, the correlations are rather weak. The 6-hourly correlation coefficients (Table A.c) for the precipitation amount at different thresholds vary from 0.06 to 0.55, indicating that nearly all the classified precipitation considered in this time-scale is not homogenous. In the range of precipitation rates $P \geq 0.1$ mm/6hr, IMERG-FR and IMERG-RT show a correlation of 0.55 and 0.21 respectively. Moreover, in the range of heavy precipitation ($5.5 \text{ mm/6hr} < P \leq 16.5 \text{ mm/6hr}$) the correlation to the values of 0.38 and 0.29 for IMERG-FR and IMERG-RT respectively are higher than for other thresholds. Also, as shown in Table A.d-e, the value of the CC in the range of light and moderate rain is low while in the entire range of precipitation ($P \geq 0.1$ mm/12hr or $P \geq 0.1$ mm/24hr) it is higher than for other classified thresholds for both 12-hourly and daily precipitation with the corresponding values of 0.61 and 0.49 for 12-hourly and 0.65 and 0.51 for daily precipitation for IMERG-FR and IMERG-RT.

Generally, the CC of IMERG-FR yields rather better than IMERG-RT due to the fact that the general performance of the CC is rather weak for different thresholds and time-scales. Several factors could contribute to the relatively low CC of IMERG products over such areas: (1) the topography and climate of this area is partly complex, posing a great challenge for accurate SPE (Dinku et al., 2007); (2) The gauges are used for the production of GPCC through monthly gauge analyses while our evaluation is on a daily and sub-daily

time-scale, resulting in the quality of IMERG products being potentially degraded; and (3) rain-gauge measurement error.

Generally, IMERG-FR and IMERG-RT overestimate hourly precipitation in the range of light rain and with a slight trend to underestimate medium, heavy, very heavy and extreme precipitation over northeast Austria (Fig. 9 table A.a). Maximum values of RMSE can reach up to 5.97 mm/hr for IMERG-FR for extreme precipitation intensity and the minimum value is close to 0 mm for IMERG-RT for precipitation without defined classification ($P \geq 0.1$ mm/hr). As shown in Table A.a and Figure 10.a, the RMSE value can be as high as 9.07 mm/hr for IMERG-RT in the range of extreme precipitation and 1.02 mm/hr for IMERG-FR for light precipitation.

3-hourly light precipitation indicates rather an overestimation for both products while only IMERG-FR shows an underestimation for other intensities (Table A.b and Figure 9.b). IMERG-FR has a total bias of +0.17 mm/3hr in comparison to +0.25 mm/3hr for IMERG-RT in the range of light precipitation intensity with a tendency to underestimate precipitation at higher intensities. In addition, in the range of extreme precipitation ($P > 11.5$ mm/3hr), IMERG-FR has a larger negative total bias (-8.50 mm/3hr) in comparison to IMERG-RT (-2.28 mm/3hr). However, MAE and RMSE of IMERG-RT, 14.79 mm/3hr and 18.92 mm/3hr respectively show larger/higher values in comparison to IMERG-FR of 10.66 mm/3hr and 12.89 mm/3hr respectively.

As shown in Table A.c and Figure 9.c, for 6-hourly precipitation and in the range of extreme precipitation, IMERG-FR underestimates precipitation with the highest Bias (-8.02 mm/6hr) and the lowest Mbias (0.64 mm/6hr) in comparison to IMERG-RT. This overestimates precipitation with a Bias and Mbias of 3.42 mm/6hr and 1.15 mm/6hr respectively. Meanwhile, MAE and RMSE for IMERG-FR and IMERG-RT correspondingly increase from light to extreme precipitation intensities (Fig. 10.c). On the other hand, based on Bias and Mbias indices, IMERG-FR tends to underestimate from light to extreme precipitation while IMERG-RT tends to underestimate from light to heavy precipitation and overestimate from very heavy to extreme precipitation.

Based on metrics of Table A.d, at 12-hourly time-scale precipitation, the performance of both satellite products can be compared. They vary between each other and for different thresholds. For instance, both products overestimate at the lowest precipitation intensity range ($0.1 \text{ mm/12hr} \leq P \leq 0.8 \text{ mm/12hr}$) and underestimate at precipitation intensities of $2 \text{ mm/12hr} < P \leq 5.5 \text{ mm/12hr}$ (Fig. 9).

Moreover, for the daily time-scale, IMERG-FR shows a smaller/lower RMSE than IMERG-RT particularly for heavy and extreme precipitation over northeast Austria, which is consistent with Table A.e and Fig. 10.e. As seen in Fig. 9.e, the total bias and its components show that IMERG-FR has a total bias to the extent of 0.64 mm/day in comparison to 0.85 mm/day for IMERG-RT in the range of light precipitation intensity. This is higher than that for medium precipitation intensity. In addition, in the range of extreme precipitation ($P > 24$ mm/day), IMERG-FR has a larger negative total bias (-7.09 mm/day) in contrast to IMERG-RT (-0.17 mm/day). However, MAE of IMERG-RT (23.76 mm/day) shows a larger value in comparison to IMERG-FR (15.37 mm/day). The low Bias compared to MAE is probably caused by the cancellation of positive/negative values.

Next, an analysis of observed precipitation above or between particular thresholds and satellite estimation above or between the same thresholds are carried out (Wilks, 2006). Multi-category contingency table evaluation (Table A.a, Table B1 and Fig. 11.a) for hourly precipitation reveals a low POD, a lower TS and high FAR for both products in all precipitation intensity classes, meaning both products are not able to detect many events on an hourly time-scale. The low POD may have been influenced by the dominance of convective storms and missed precipitation over this region (Nasrollahi, 2015; Sharifi et al.,

2016). However, relatively high FAR values for classified precipitation might be due to spurious events detected by satellites (Blacutt et al., 2015) and/or disability of satellite products to detect precipitation in their exact precipitation categories. However, they may detect the amount of precipitation somewhat lower or higher than the specified intensities. Moreover, a frequency bias of light, moderate and heavy precipitation less than unity (<1) for both products indicate that satellite estimation underestimates the frequency, while for extreme precipitation ($P > 5.6 \text{ mm/hr}$) an overestimation is shown. Despite this, at this time-scale, the accuracy values of both products are similar in each intensity and they show nearly a perfect skill score in the range of extreme precipitation.

With respect to the 3-hourly time-scale with different precipitation intensities (Table A.b, Table B2 and Fig.11.b), IMERG-RT shows better POD than IMERG-FR with 0.46 and 0.34 for IMERG-RT and 0.32 and 0.24 for IMERG-FR in the range of heavy and extreme precipitation. FAR values of both products are relatively high and there is no meaningful difference between the two products. Lower POD and high FAR may be associated with systematic errors of the products algorithms, missed precipitation over this region, poor performance to detect short-lived convective precipitation, and precipitation over the surface covered by snow. Moreover, in the range of extreme precipitation intensity, IMERG-FR and IMERG-RT both tend to overestimate the number of events with a bias frequency of 1.9 and 3.4 respectively, but the value of IMERG-FR is close to unity, even in the range of heavy precipitation. IMERG-FR with the bias frequency value of 0.96 outperformed IMERG-RT with a value of 0.79.

Table A.c, Table B3 and Figure 11.c show the 6-hourly precipitation statistical analysis and details. With respect to the contingency table in the range of light to very heavy precipitation, the behavior of both products is similar and shows a low POD and TS and a high FAR, while the POD of extreme precipitation intensity ($P > 16.5 \text{ mm/6hr}$) yields relatively high values for both products. However, IMERG-RT yields better results than IMERG-FR for the POD values of 0.44 and 0.38 respectively. Nevertheless, in general, indices related to the contingency table show high POD and TS and a low FAR but they underestimate the frequency of events for precipitation equal or above 0.1 mm/6hr . It is notable that there is no meaningful difference between both products.

For 12-hourly precipitation, IMERG-FR and IMERG-RT tended to underestimate the number of events corresponding to light, moderate, heavy and very heavy precipitation while they tended to slightly overestimate the extreme precipitation with the bias frequency of 1.6 and 2.27 for IMERG-FR and IMERG-RT respectively. Moreover, again both products yield relatively poor results with respect to POD, TS and FAR for light, moderate, heavy and very heavy precipitation. Instead, extreme precipitation ($P > 18.4 \text{ mm/12hr}$) yields better. However, in this threshold, IMERG-FR yields better than IMERG-RT to detect precipitation with higher POD and TS and lower FAR, 0.51, 0.24 and 0.68 in contrast to 0.47, 0.17 and 0.79 for IMERG-FR and IMERG-RT respectively (see Tables A.d, B4 and Fig. 11.d).

With respect to different precipitation intensities for the daily time-scale (Tables A.e, B5 and Fig. 12.e), IMERG-FR generally reveals better results than IMERG-RT in all precipitation categories. However, in the range of extreme precipitation ($P > 24 \text{ mm/day}$) and precipitation with no classified intensity ($P \geq 0.1 \text{ mm/day}$) both products show a relatively high POD with 0.54 and 0.65 for IMERG-FR and 0.49 and 0.66 for IMERG-RT. The FAR values of both products are relatively high for precipitation with classified intensities and low for no classified precipitation and indicate 0.25 and 0.26 for IMERG-FR and IMERG-RT, respectively (Fig. 11.e). It is notable that there is no big difference between the FAR values of the two products.

The advantage of the multi-category contingency table approach is that the nature of the forecast errors can more easily be diagnosed but the disadvantage is that it is more difficult to

condense the results into a single number. Despite this, in this study, a low POD and TS, and a high FAR for classified precipitation means that satellite products are not able to detect precipitation in their exact precipitation categories, but they are able to detect the amount of precipitation somewhat lower or higher than the specified intensities. Moreover, as for the other time-scale evaluations in the range of extreme precipitation intensity ($P > 24$ mm/day), both IMERG-FR and IMERG-RT tend to overestimate the number of events with a bias frequency of 1.7 and 2.31 respectively. For more information about the frequency of events with different intensities see Appendix B1-5.

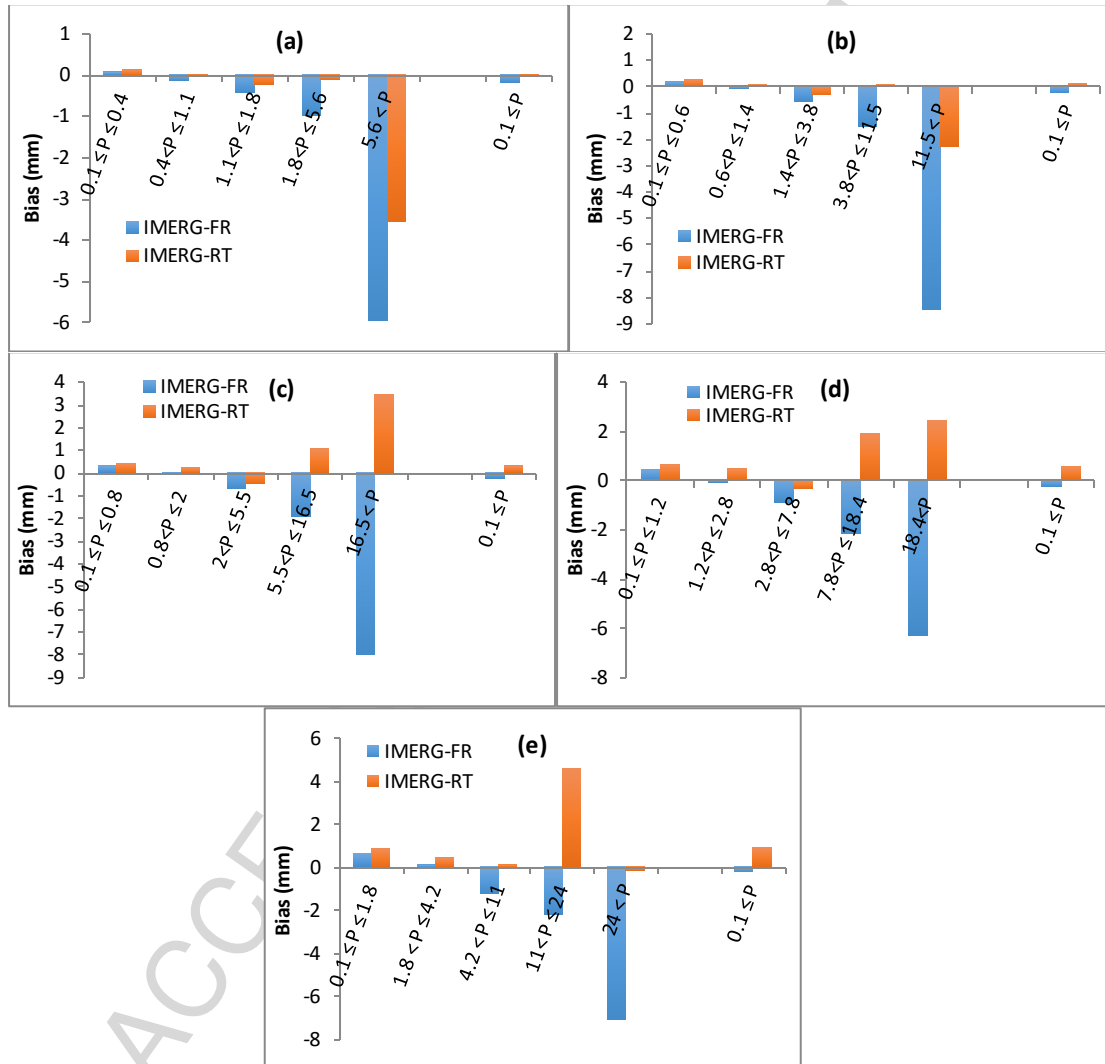


Fig 9. The bias of precipitation events for different time-scale and intensities a) hourly b) 3-hourly c) 6-hourly d) 12-hourly and e) daily over northeast Austria for the period March 2015 to January 2016.

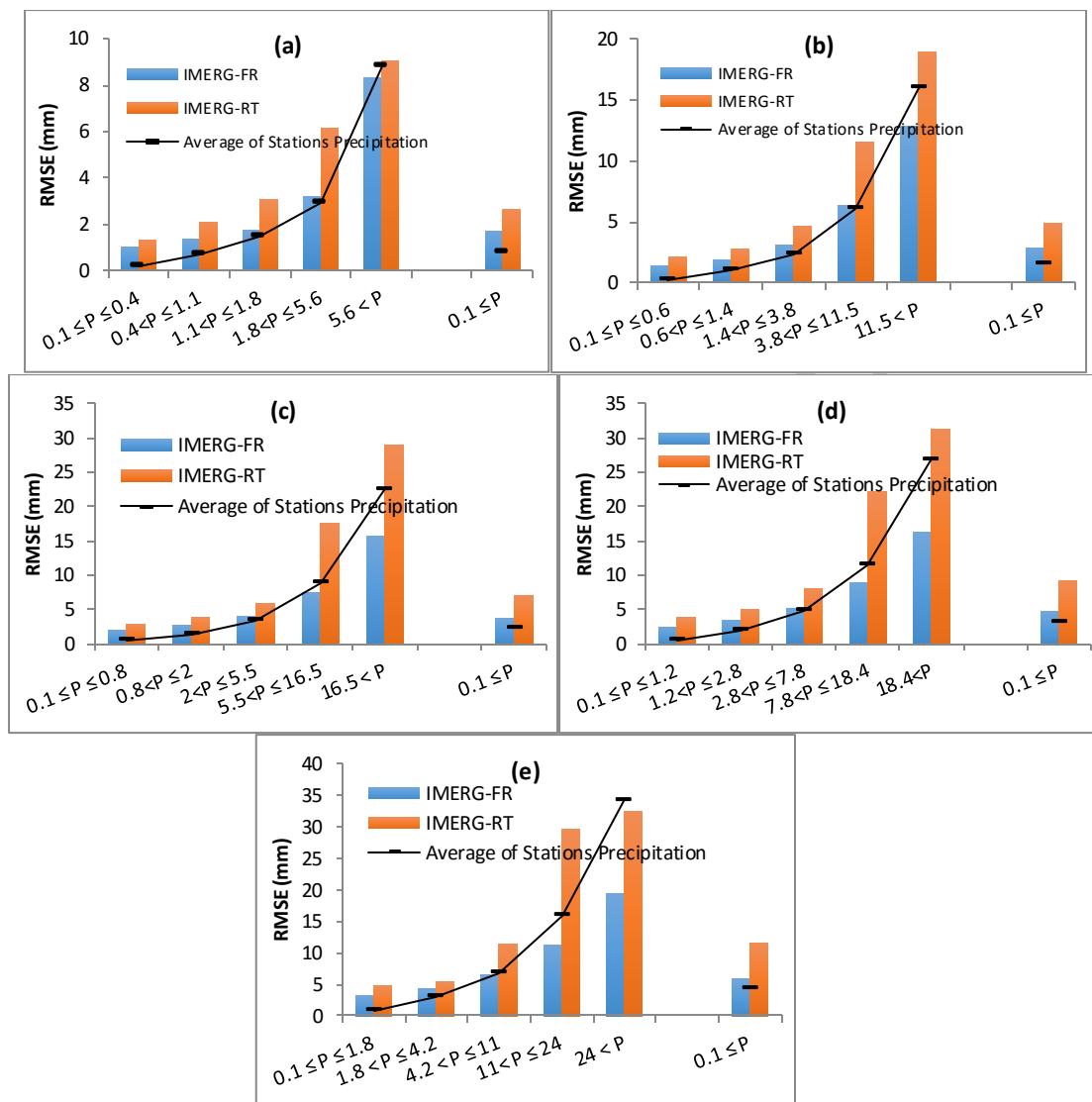
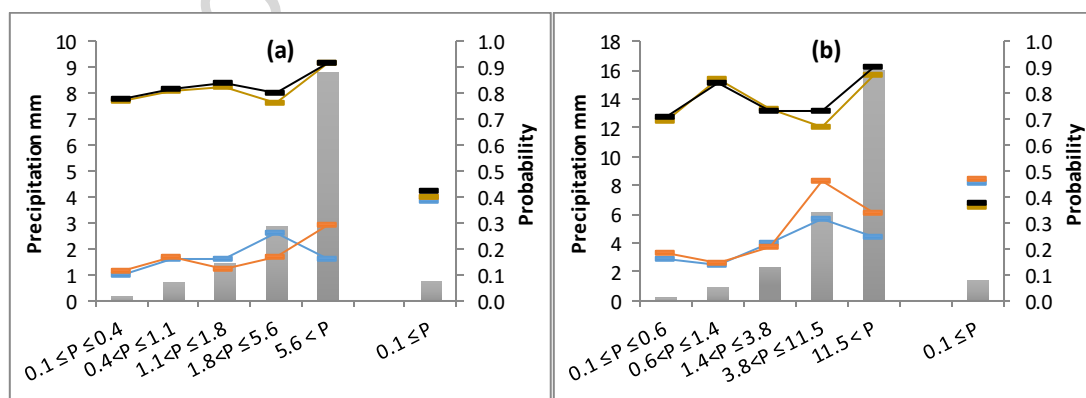


Fig 10. RMSE of precipitation events for different time-scale and intensities a) hourly b) 3-hourly c) 6-hourly d) 12-hourly and e) daily over northeast Austria for period March 2015 to January 2016.



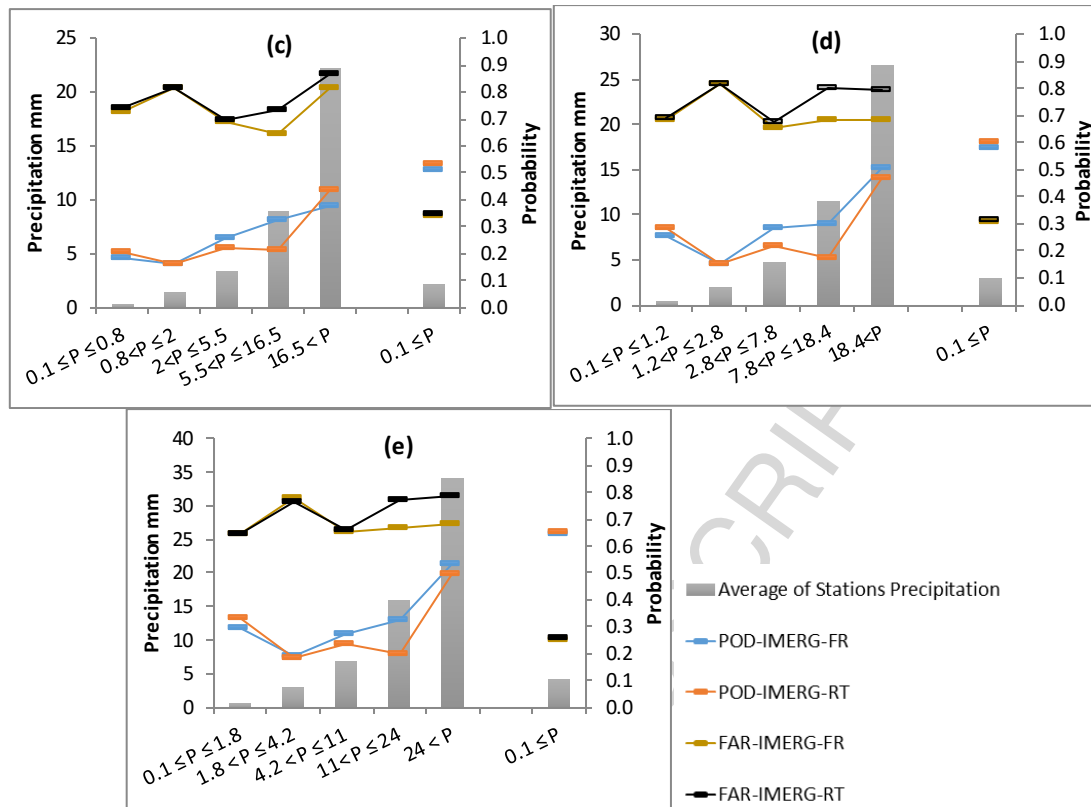


Fig 11. Verification statistics (POD and FAR) between IMERG-FR / IMERG-RT and observed gauge observation for different time-scale and intensities a) hourly b) 3-hourly c) 6-hourly d) 12-hourly and e) daily over northeast Austria for period March 2015 to January 2016.

5. Conclusions

Accurate high-resolution multi-satellite precipitation estimates have wide applications in hydrology, meteorology, water resource management and numerical weather prediction model verification. In this study, a comprehensive assessment has been performed to examine the performance of the Day-1 IMERG-FR and IMERG-RT late-run version V03 against 62 meteorological synoptic station data on a multi time-scale such as daily and sub-daily time resolution over northeast Austria from mid-March 2015 till end of January 2016. IMERG-FR uses monthly GPCC gauge data over land for bias correction; and the IMERG-RT late-run product also uses a coefficient of monthly climatology data.

It is found that in general for hourly precipitation and for entire precipitation rate ($P \geq 0.1$), both products indicate a negative bias. However, for other time steps, IMERG-FR tends to underestimate while IMERG-RT tends to overestimate precipitation over the study area.

Based on the entire range of precipitation, hourly, 3- and 6-hourly IMERG-FR did not show a clear improvement of the bias over IMERG-RT while for 12-hourly and daily precipitation estimates, the bias in IMERG-FR has improved compared to IMERG-RT. In addition, IMERG-FR shows a considerable improvement in RMSE compared to IMERG-RT. However, IMERG-FR systematically underestimates moderate to extreme and overestimates light precipitation for the entire range of precipitation intensities ($P \geq 0.1$ mm/time interval). When comparing the bias, RMSE, and correlation coefficients, IMERG-FR outperforms IMERG-RT particularly for 6-hourly, 12-hourly and daily precipitation. Despite the

generally low POD and TS and high FAR skill scores within specified precipitation thresholds, the multi-categorical contingency table shows relatively good values of the POD, TS and FAR for precipitation without classification. This means these two satellite products are relatively well able to detect precipitation without classification but they have poor results to detect precipitation in their exact precipitation categories over northeast Austria.

Also, in agreement with earlier studies, there is a different statistical relation between satellite and in-situ precipitation depending on elevation, land surface characteristics and snow-rain phase of precipitation. Therefore, further work on these issues is needed. This research is to our knowledge one of the first studies to evaluate GPM-IMERG products over Austria. Further intensive detailed studies are needed when improved and fully GPM-IMERG retrospectively processed data starting from 1998-2000 become available to understand inter-annual variability which is very important because sensors and satellites used in multi-satellite products can be out of service and new sensors can be added at any time (Prakash et al., 2016).

In summary, although the newly introduced sensors and upgraded calibration algorithms have undoubtedly improved the GPM constellation satellites' accuracy, some challenging issues in satellite retrieval processes will continue to remain open for the satellite community, providing the impetus for more research and development.

The overestimation or underestimation over forest areas indicates that accurate estimation by satellite-based precipitation products even though they are improved still remains a challenge particularly for short-time precipitation (i.e., hourly precipitation). Such inaccuracy may be rooted in the following possible causes of uncertainty of evaluation of SPEs against rain-gauges:

1. The spatial resolution of the satellite product, since precipitation within a region, may occur on smaller scales than the pixel size of satellites.
2. An inadequate number of gauges, provided by the Global Precipitation Climatology Centre (GPCC) and used for bias correction in satellite products.
3. Short-time precipitation (i.e., daily and sub-daily accumulation) is much more variable than monthly precipitation and regional effects like topography and local circulation play an important role in rainfall generation. Accordingly, monthly precipitation from GPCC which are used for the calibration of IMERG-FR would not be well smoothed if sub-daily precipitation fields are used.
4. Algorithm error
5. Rain-gauge measurements error due to wind-induced undercatching, precipitation type and particle falling velocities, rainfall intensity and the aerodynamic properties of a particular type of gauge

As more IMERG data become available, more detailed studies of GPM-IMERG applications in water, weather, and climate studies are possible in the near future. We expect that the regional analysis of GPM constellation satellite-based precipitation estimates reported here can give users a better understanding of the features associated with currently available IMERG precipitation estimates and a broader perspective.

Acknowledgments: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

ACCEPTED MANUSCRIPT

Appendix A

(a)

	0.1≤P≤0.4		0.4<P≤1.1		1.1<P≤1.8		1.8<P≤5.6		5.6<P		0.1≤P	
	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT
RMSE mm	1.02	1.31	1.38	2.11	1.77	3.08	3.20	6.14	8.33	9.07	1.72	2.67
MAE mm	0.35	0.40	0.79	0.90	1.24	1.56	2.31	3.35	6.76	7.35	0.79	0.97
Bias mm	0.07	0.12	-0.14	-0.05	-0.44	-0.23	-0.99	-0.11	-5.97	-3.56	-0.18	-0.01
Mbias	1.37	1.64	0.81	0.93	0.70	0.84	0.66	0.96	0.32	0.60	0.76	0.99
CC	0.09	0.07	0.09	0.07	0.05	0.06	0.15	0.14	0.00	0.05	0.30	0.29
POD	0.10	0.11	0.16	0.17	0.16	0.12	0.26	0.17	0.16	0.29	0.39	0.40
FAR	0.77	0.78	0.81	0.81	0.83	0.84	0.76	0.80	0.91	0.91	0.40	0.43
TS	0.07	0.08	0.09	0.10	0.09	0.07	0.14	0.10	0.06	0.07	0.31	0.31
Accuracy	0.94	0.94	0.97	0.97	0.99	0.99	0.99	0.99	1.00	1.00	0.93	0.92
Frequency bias	0.41	0.52	0.85	0.91	0.92	0.75	1.09	0.85	1.92	3.39	0.64	0.70

(b)

	0.1≤P≤0.6		0.6<P≤1.4		1.4<P≤3.8		3.8<P≤11.5		11.5<P		0.1≤P	
	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT
RMSE mm	1.43	2.13	1.88	2.79	3.11	4.68	6.33	11.55	12.89	18.92	2.88	4.93
MAE mm	0.52	0.60	1.15	1.32	2.11	2.53	4.48	6.48	10.66	14.79	1.39	1.80
Bias mm	0.17	0.25	-0.08	0.08	-0.58	-0.33	-1.52	0.03	-8.50	-2.28	-0.24	0.12
Mbias	1.68	2.00	0.92	1.08	0.75	0.86	0.75	1.01	0.47	0.86	0.83	1.08
CC	0.11	0.07	0.08	0.09	0.14	0.12	0.26	0.22	-0.01	0.05	0.46	0.40
POD	0.16	0.18	0.13	0.15	0.22	0.20	0.32	0.46	0.24	0.34	0.45	0.47
FAR	0.69	0.71	0.85	0.84	0.74	0.73	0.67	0.73	0.87	0.90	0.36	0.38

TS	0.12	0.13	0.08	0.08	0.14	0.13	0.19	0.13	0.09	0.08	0.36	0.36
Accuracy	0.93	0.93	0.96	0.96	0.96	0.96	0.98	0.98	1.00	1.00	0.89	0.89
Frequency bias	0.51	0.61	0.92	0.91	0.83	0.76	0.96	0.79	1.90	3.40	0.71	0.76

(c)

	0.1≤P≤0.8		0.8<P≤2		2<P≤5.5		5.5<P≤16.5		16.5<P		0.1≤P	
	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT
RMSE mm	1.97	2.89	2.67	3.82	4.07	5.90	7.51	17.53	15.68	28.96	3.75	7.04
MAE mm	0.69	0.80	1.60	1.86	2.90	3.45	5.95	9.87	12.72	21.21	1.88	2.53
Bias mm	0.29	0.41	-0.05	0.22	-0.72	-0.49	-1.96	1.04	-8.02	3.42	-0.25	0.29
Mbias	1.92	2.31	0.97	1.16	0.79	0.86	0.78	1.12	0.64	1.15	0.88	1.14
CC	0.11	0.09	0.06	0.06	0.15	0.13	0.38	0.29	0.06	0.20	0.55	0.21
POD	0.18	0.21	0.16	0.16	0.26	0.22	0.33	0.21	0.38	0.44	0.51	0.53
FAR	0.73	0.74	0.82	0.82	0.69	0.70	0.65	0.73	0.81	0.87	0.34	0.35
TS	0.12	0.13	0.09	0.09	0.16	0.15	0.20	0.13	0.14	0.11	0.40	0.42
Accuracy	0.87	0.87	0.94	0.94	0.95	0.95	0.98	0.98	1.00	0.99	0.86	0.86
Frequency bias	0.67	0.79	0.87	0.88	0.83	0.72	0.92	0.78	2.02	3.33	0.78	0.82

(d)

	0.1≤P≤1.2		1.2<P≤2.8		2.8<P≤7.8		7.8<P≤18.4		18.4<P		0.1≤P	
	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT
RMSE mm	2.44	3.80	3.45	5.05	5.18	8.04	8.93	22.14	16.23	31.29	4.71	9.14
MAE mm	0.91	1.11	2.18	2.71	3.87	4.83	7.26	12.92	12.69	21.40	2.48	3.49
Bias mm	0.41	0.62	-0.01	0.47	-0.90	-0.35	-2.19	1.88	-6.32	2.41	-0.27	0.53

Mbias	1.96	2.43	0.99	1.24	0.81	0.93	0.81	1.16	0.76	1.09	0.91	1.18
CC	0.14	0.11	0.06	0.01	0.15	0.10	0.36	0.23	0.12	0.17	0.61	0.49
POD	0.26	0.29	0.15	0.15	0.28	0.22	0.30	0.18	0.51	0.47	0.58	0.60
FAR	0.68	0.69	0.82	0.81	0.65	0.68	0.68	0.80	0.68	0.79	0.31	0.31
TS	0.16	0.18	0.09	0.09	0.18	0.15	0.18	0.10	0.24	0.17	0.46	0.47
Accuracy	0.82	0.81	0.92	0.92	0.93	0.93	0.97	0.97	0.99	0.99	0.82	0.82
Frequency bias	0.81	0.93	0.84	0.80	0.81	0.67	0.95	0.88	1.60	2.27	0.84	0.87

(e)												
	0.1≤P≤1.8		1.8<P≤4.2		4.2<P≤11		11<P≤24		24<P		0.1≤P	
	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT	IMERG-FR	IMERG-RT
RMSE mm	3.25	4.77	4.35	5.37	6.52	11.40	11.28	29.68	19.39	32.49	5.95	11.63
MAE mm	1.32	1.54	2.88	3.39	5.15	6.88	9.29	18.41	15.37	23.76	3.30	4.75
Bias mm	0.64	0.85	0.01	0.41	-1.24	0.10	-2.21	4.56	-7.09	-0.17	-0.24	0.90
Mbias	2.00	2.34	1.00	1.14	0.82	1.01	0.86	1.29	0.79	1.00	0.94	1.21
CC	0.17	0.15	0.13	0.12	0.19	0.17	0.27	0.15	0.14	0.28	0.65	0.51
POD	0.30	0.33	0.19	0.18	0.27	0.24	0.33	0.20	0.54	0.49	0.65	0.66
FAR	0.64	0.64	0.78	0.77	0.65	0.66	0.67	0.77	0.69	0.79	0.25	0.26
TS	0.19	0.21	0.11	0.11	0.18	0.16	0.20	0.12	0.25	0.17	0.53	0.54
Accuracy	0.76	0.75	0.89	0.90	0.91	0.91	0.96	0.95	0.99	0.98	0.78	0.78
Frequency bias	0.83	0.93	0.86	0.77	0.79	0.70	0.99	0.88	1.70	2.31	0.86	0.88

Table A: Summary of evaluation metrics for IMERG-FR and IMERG-RT products at a)hourly , b)3-hourly, c)6-hourly, d)12-hourly and e)daily time-scales over northeast Austria. The metrics are calculated based on all hours of all stations.

Appendix B

(a)

IMERG-FR

Station		P<0.1	0.1≤P≤0.4	0.4<P≤1.1	1.1<P≤1.8	1.8<P≤5.6	5.6<P	Total
	P<0.1	374858	4697	3011	729	645	115	384055
	0.1≤P≤0.4	14827	1928	1916	644	595	98	20008
	0.4<P≤1.1	4921	987	1425	697	754	98	8882
	1.1<P≤1.8	1362	304	610	437	587	66	3366
	1.8<P≤5.6	866	311	541	427	834	246	3225
	5.6<P	52	28	50	59	107	58	354
	Total	396886	8255	7553	2993	3522	681	419890

(b)

IMERG-RT

Station		P<0.1	0.1≤P≤0.4	0.4<P≤1.1	1.1<P≤1.8	1.8<P≤5.6	5.6<P	Total
	P<0.1	373435	5953	3099	711	656	201	384055
	0.1≤P≤0.4	14468	2296	1955	542	557	190	20008
	0.4<P≤1.1	4851	1268	1503	535	528	197	8882
	1.1<P≤1.8	1296	473	735	337	377	148	3366
	1.8<P≤5.6	814	333	741	424	553	360	3225
	5.6<P	39	28	64	57	63	103	354
	Total	394903	10351	8097	2606	2734	1199	419890

Table B1: Frequency of events with different intensities in corresponding to GPM-IMERG products a) IMERG-FR and b) IMERG-RT for hourly precipitation

		(a)						
		IMERG-FR						
Station		P<0.1	0.1≤P≤0.6	0.6<P≤1.4	1.4<P≤3.8	3.8<P≤11.5	11.5<P	Total
	P<0.1	115980	2850	1267	645	152	44	120938
	0.1≤P≤0.6	6851	1281	889	602	179	30	9832
	0.6<P≤1.4	1886	534	491	540	198	18	3667
	1.4<P≤3.8	1358	429	533	802	487	54	3663
	3.8<P≤11.5	291	113	165	416	530	165	1680
	11.5<P	11	10	18	37	66	46	188
	Total	126377	5217	3363	3042	1612	357	139968

		(b)						
		IMERG-RT						
Station		P<0.1	0.1≤P≤0.6	0.6<P≤1.4	1.4<P≤3.8	3.8<P≤11.5	11.5<P	Total
	P<0.1	115464	5953	3099	711	656	201	384055
	0.1≤P≤0.6	14468	2296	1955	542	557	190	20008
	0.6<P≤1.4	4851	1268	1503	535	528	197	8882
	1.4<P≤3.8	1296	473	735	337	377	148	3366
	3.8<P≤11.5	814	333	741	424	553	360	3225
	11.5<P	39	28	64	57	63	103	354
	Total	394903	10351	8097	2606	2734	1199	419890

Table B2: Frequency of events with different intensities in corresponding to GPM-IMERG products a) IMERG-FR and b) IMERG-RT for 3-hourly precipitation

		(a)						
		IMERG-FR						
Station		P<0.1	0.1≤P≤0.8	0.8<P≤2	2<P≤5.5	5.5<P≤16.5	16.5<P	Total
	P<0.1	53090	2397	726	344	71	17	56645
	0.1≤P≤0.8	4357	1241	665	408	108	23	6802
	0.8<P≤2	1238	487	432	374	140	17	2688
	2<P≤5.5	760	350	389	645	347	31	2522
	5.5<P≤16.5	154	98	128	295	389	130	1194
	16.5<P	3	8	5	18	49	50	133
	Total	59602	4581	2345	2084	1104	268	69984

		(b)						
		IMERG-RT						
Station		P<0.1	0.1≤P≤0.8	0.8<P≤2	2<P≤5.5	5.5<P≤16.5	16.5<P	Total
	P<0.1	53090	2397	726	344	71	17	56645
	0.1≤P≤0.8	4357	1241	665	408	108	23	6802
	0.8<P≤2	1238	487	432	374	140	17	2688
	2<P≤5.5	760	350	389	645	347	31	2522
	5.5<P≤16.5	154	98	128	295	389	130	1194
	16.5<P	3	8	5	18	49	50	133
	Total	59602	4581	2345	2084	1104	268	69984

Station	P<0.1	52822	2792	639	266	110	16	56645
	0.1≤P≤0.8	4196	1412	663	342	156	33	6802
	0.8<P≤2	1183	550	431	330	153	41	2688
	2<P≤5.5	689	523	438	551	221	100	2522
	5.5<P≤16.5	140	115	189	304	251	195	1194
	16.5<P	2	3	9	21	40	58	133
	Total	59032	5395	2369	1814	931	443	69984

Table B3: Frequency of events with different intensities in corresponding to GPM-IMERG products a) IMERG-FR and b) IMERG-RT for 6-hourly precipitation

		(a) IMERG-FR						Total
		P<0.1	0.1≤P≤1.2	1.2<P≤2.8	2.8<P≤7.8	7.8<P≤18.4	18.4<P	
Station	P<0.1	23135	1859	379	160	43	14	25590
	0.1≤P≤1.2	2741	1236	459	300	65	21	4822
	1.2<P≤2.8	692	391	271	322	77	16	1769
	2.8<P≤7.8	423	351	296	528	238	42	1878
	7.8<P≤18.4	66	78	77	190	224	111	746
	18.4<P	2	6	8	18	58	95	187
	Total	27059	3921	1490	1518	705	299	34992

		(b) IMERG-RT						Total
		P<0.1	0.1≤P≤1.2	1.2<P≤2.8	2.8<P≤7.8	7.8<P≤18.4	18.4<P	
Station	P<0.1	23022	2106	280	123	46	13	25590
	0.1≤P≤1.2	2633	1387	393	259	117	33	4822
	1.2<P≤2.8	654	454	265	236	118	42	1769
	2.8<P≤7.8	396	425	366	407	181	103	1878
	7.8<P≤18.4	62	90	108	208	132	146	746
	18.4<P	2	4	9	23	61	88	187
	Total	26769	4466	1421	1256	655	425	34992

Table B4: Frequency of events with different intensities in corresponding to GPM-IMERG products a) IMERG-FR and b) IMERG-RT for 12-hourly precipitation

		(a) IMERG-FR						Total
		P<0.1	0.1≤P≤1.8	1.8<P≤4.2	4.2<P≤11	11<P≤24	24<P	
n	P<0.1	9426	1145	174	65	20	6	10836

0.1≤P≤1.8	1756	1007	380	207	44	19	3413
1.8<P≤4.2	364	354	242	250	53	9	1272
4.2<P≤11	217	272	230	354	201	27	1301
11<P≤24	21	48	59	132	176	100	536
24<P	2	4	6	16	36	74	138
Total	11786	2830	1091	1024	530	235	17496

Station	(b)						
	IMERG-RT						
	P<0.1	0.1≤P≤1.8	1.8<P≤4.2	4.2<P≤11	11<P≤24	24<P	Total
P<0.1	9342	1290	119	62	18	5	10836
0.1≤P≤1.8	1712	1133	284	175	83	26	3413
1.8<P≤4.2	355	381	229	201	87	19	1272
4.2<P≤11	206	310	248	311	137	89	1301
11<P≤24	23	52	93	149	107	112	536
24<P	2	3	9	19	37	68	138
Total	11640	3169	982	917	469	319	17496

Table B5: Frequency of events with different intensities in corresponding to GPM-IMERG products a) IMERG-FR and b) IMERG-RT for daily precipitation

References

- Aires, F., Prigent, C., Bernardo, F., Jiménez, C., Saunders, R., Brunel, P., 2011. A Tool to Estimate Land-Surface Emissivities at Microwave frequencies (TELSEM) for use in numerical weather prediction. *Q.J.R. Meteorol. Soc.* 137 (656), 690–699.
- Ballester, J., Moré, J., 2007. The representativeness problem of a station net applied to the verification of a precipitation forecast based on areas. *Met. Apps* 14 (2), 177–184.
- Behrangi, A., Tian, Y., Lambrigtsen, B.H., Stephens, G.L., 2014. What does CloudSat reveal about global land precipitation detection by other spaceborne sensors? *Water Resour. Res.* 50 (6), 4893–4905.
- Blacutt, L.A., Herdies, D.L., de Gonçalves, Luis Gustavo G., Vila, D.A., Andrade, M., 2015. Precipitation comparison for the CFSR, MERRA, TRMM3B42 and Combined Scheme datasets in Bolivia. *Atmospheric Research* 163, 117–131.
- Carr, N., Kirstetter, P.-E., Hong, Y., Gourley, J.J., Schwaller, M., Petersen, W., Wang, N.-Y., Ferraro, R.R., Xue, X., 2015. The Influence of Surface and Precipitation Characteristics on TRMM Microwave Imager Rainfall Retrieval Uncertainty. *J. Hydrometeor* 16 (4), 1596–1614.
- Collier, C.G., 2009. On the propagation of uncertainty in weather radar estimates of rainfall through hydrological models. *Met. Apps* 16 (1), 35–40.
- Dezfuli, A.K., Ichoku, C.M., Huffman, G.J., Mohr, K.I., Selker, J.S., van de Giesen, N., Hochreutener, R., Annor, F.O., 2017. Validation of IMERG Precipitation in Africa. *J. Hydrometeor.* 18 (10), 2817–2825.
- Dinku, T., Ceccato, P., Grover-Kopec, E., Lemma, M., Connor, S.J., Ropelewski, C.F., 2007. Validation of satellite rainfall products over East Africa's complex topography. *International Journal of Remote Sensing* 28 (7), 1503–1526.
- Ebert, E.E., Janowiak, J.E., Kidd, C., 2007. Comparison of Near-Real-Time Precipitation Estimates from Satellite Observations and Numerical Models. *Bull. Amer. Meteor. Soc.* 88 (1), 47–64.
- Gebere, S., Alamirew, T., Merkel, B., Melesse, A., 2015. Performance of High Resolution Satellite Rainfall Products over Data Scarce Parts of Eastern Ethiopia. *Remote Sensing* 7 (9), 11639–11663.

- Germann, U., Berenguer, M., Sempere-Torres, D., Zappa, M., 2009. REAL-Ensemble radar precipitation estimation for hydrology in a mountainous region. *Q.J.R. Meteorol. Soc.* 135 (639), 445–456.
- Germann, U., Galli, G., Boscacci, M., Bolliger, M., 2006. Radar precipitation measurement in a mountainous region. *Q. J. R. Meteorol. Soc.* 132 (618), 1669–1692.
- Gosset, M., Viarre, J., Quantin, G., Alcoba, M., 2013. Evaluation of several rainfall products used for hydrological applications over West Africa using two high-resolution gauge networks. *Q.J.R. Meteorol. Soc.* 139 (673), 923–940.
- Guo, H., Chen, S., Bao, A., Hu, J., Gebregiorgis, A., Xue, X., Zhang, X., 2015. Inter-Comparison of High-Resolution Satellite Precipitation Products over Central Asia. *Remote Sensing* 7 (12), 7181–7211.
- Haiden, T., Kann, A., Wittmann, C., Pistotnik, G., Bica, B., Gruber, C., 2011. The Integrated Nowcasting through Comprehensive Analysis (INCA) System and Its Validation over the Eastern Alpine Region. *Wea. Forecasting* 26 (2), 166–183.
- Hakuba, M.Z., Folini, D., Sanchez-Lorenzo, A., Wild, M., 2014. Spatial representativeness of ground-based solar radiation measurements-Extension to the full Meteosat disk. *J. Geophys. Res. Atmos.* 119 (20), 12.
- Hennessy, K.J., Suppiah, R., 1999. Australian rainfall changes, 1910–1955. *Australian Meteorology Magazine* 48, 1–13.
- Hiebl, J., Reisenhofer, S., Auer, I., Böhm, R., Schöner, W., 2011. Multi-methodical realisation of Austrian climate maps for 1971–2000. *Adv. Sci. Res.* 6, 19–26.
- Hou, A.Y., Kakar, R.K., Neeck, S., Azarbarzin, A.A., Kummerow, C.D., Kojima, M., Oki, R., Nakamura, K., Iguchi, T., 2014. The Global Precipitation Measurement Mission. *Bull. Amer. Meteor. Soc.* 95 (5), 701–722.
- Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K.L., Joyce, R.J., 2015a. Algorithm Theoretical Basis Document (ATBD) Version 4.5: NASA Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG).
- Huffman, G.J., Bolvin, D.T., Nelkin, E.J., 2015b. Day 1 IMERG Final Run Release Notes, 9 pp.

- Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Wolff, D., Adler, R.F., Gu, G., Hong, Y., Bowman, K.P., Stocker, E.F., 2007. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *J. Hydrometeor* 8 (1), 38–55.
- Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., 2004. CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. *Journal of Hydrometeorology* 5.
- Karl, T.R., Knight, R.W., 1998. Secular Trends of Precipitation Amount, Frequency, and Intensity in the United States. *Bull. Amer. Meteor. Soc.* 79 (2), 231–241.
- Karl, T.R., Knight, R.W., Plummer, N., 1995. Trends in high-frequency climate variability in the twentieth century. *Nature* 377, 217–220.
- Krakauer, N., Pradhanang, S., Lakhankar, T., Jha, A., 2013. Evaluating Satellite Products for Precipitation Estimation in Mountain Regions: A Case Study for Nepal. *Remote Sensing* 5 (8), 4107–4123.
- Li, Z., Yang, D., Hong, Y., 2013. Multi-scale evaluation of high-resolution multi-sensor blended global precipitation products over the Yangtze River. *Journal of Hydrology* 500, 157–169.
- Liu, Z., 2016. Comparison of Integrated Multisatellite Retrievals for GPM (IMERG) and TRMM Multisatellite Precipitation Analysis (TMPA) Monthly Precipitation Products: Initial Results. *J. Hydrometeor* 17 (3), 777–790.
- Nasrollahi, N., 2015. False Alarm in Satellite Precipitation Data, in: Nasrollahi, N. (Ed.), *Improving Infrared-Based Precipitation Retrieval Algorithms Using Multi-Spectral Satellite Imagery*. Springer International Publishing, Cham, pp. 7–12.
- Plummer, N., Salinger, M.J., Nicholls, N., 1999. Changes in Climate Extremes Over the Australian Region and New Zealand During the Twentieth Century. *Climatic Change* 42, 183–202.
- Prakash, S., Mitra, A.K., AghaKouchak, A., Liu, Z., Norouzi, H., Pai, D.S., 2016. A preliminary assessment of GPM-based multi-satellite precipitation estimates over a monsoon dominated region. *Journal of Hydrology*.
- Rubel, F., Brugger, K., Haslinger, K., Auer, I., 2017. The climate of the European Alps: Shift of very high resolution Köppen-Geiger climate zones 1800–2100. *metz* 26 (2), 115–125.

- Shapiro, L., Stochman, G. (Eds.), 2001. Computer Vision: Filtering and Enhancing Images. Prentice Hall PTR Upper Saddle River, NJ, USA, 61 pp.
- Sharifi, E., Steinacker, R., Saghafian, B., 2016. Assessment of GPM-IMERG and Other Precipitation Products against Gauge Data under Different Topographic and Climatic Conditions in Iran: Preliminary Results. *Remote Sensing* 8 (2), 135.
- Shen, Y., Xiong, A., Wang, Y., Xie, P., 2010. Performance of high-resolution satellite precipitation products over China. *J. Geophys. Res.* 115 (D2).
- Simonović, P.S., 2012. Floods in a changing climate. Risk management, 1st ed. Cambridge Univ. Press, Cambridge.
- Sohn, B.J., Ryu, G.-H., Song, H.-J., Ou, M.-L., 2013. Characteristic Features of Warm-Type Rain Producing Heavy Rainfall over the Korean Peninsula Inferred from TRMM Measurements. *Mon. Wea. Rev.* 141 (11), 3873–3888.
- Sorooshian, S., Hsu, K.L., Gao, X., Gupta, H.V., Imam, B., Braithwaite, D., 2000. Evaluation of PERSIANN System Satellite-Based Estimates of Tropical Rainfall. *Bulletin of the American Meteorological Society* 81.
- Sungmin, O., Foelsche, U., Kirchengast, G., Fuchsberger, J., 2016. Validation and correction of rainfall data from the WegenerNet high density network in southeast Austria. *Journal of Hydrology*.
- Sungmin, O., Foelsche, U., Kirchengast, G., Fuchsberger, J., Tan, J., Petersen, W.A., 2017. Evaluation of GPM IMERG Early, Late, and Final rainfall estimates with WegenerNet gauge data in southeast Austria. *Hydrol. Earth Syst. Sci. Discuss.*, 1–21.
- Suppiah, R., Hennessy, K.J., 1996. Trends in the intensity and frequency of heavy rainfall in tropical Australia and links with the Southern Oscillation. *Australian Meteorology Magazine* 45, 1–17.
- Suppiah, R., Hennessy, K.J., 1998. Trends in total rainfall, heavy rain events and number of dry days in Australia, 1910–1990. *International Journal of Climatology* 10, 1141–1164.
- Tian, Y., Peters-Lidard, C.D., Choudhury, B.J., Garcia, M., 2007. Multitemporal Analysis of TRMM-Based Satellite Precipitation Products for Land Data Assimilation Applications. *J. Hydrometeor* 8 (6), 1165–1183.

- Vila, D.A., de Goncalves, Luis Gustavo G., Toll, D.L., Rozante, J.R., 2009. Statistical Evaluation of Combined Daily Gauge Observations and Rainfall Satellite Estimates over Continental South America. *J. Hydrometeor* 10 (2), 533–543.
- Villarini, G., Krajewski, W.F., Ciach, G.J., Zimmerman, D.L., 2009. Product-error-driven generator of probable rainfall conditioned on WSR-88D precipitation estimates. *Water Resour. Res.* 45 (1), n/a-n/a.
- Villarini, G., Mandapaka, P.V., Krajewski, W.F., Moore, R.J., 2008. Rainfall and sampling uncertainties: A rain gauge perspective. *J. Geophys. Res.* 113 (D11), 385.
- Wang, K., Augustine, J., Dickinson, R.E., 2012. Critical assessment of surface incident solar radiation observations collected by SURFRAD, USCRN and AmeriFlux networks from 1995 to 2011. *J. Geophys. Res.* 117 (D23), n/a-n/a.
- Wilks, D.S., 2006. *STATISTICAL METHODS IN THE ATMOSPHERIC SCIENCES*, Second Edition ed. Elsevier, United States of America, 649 pp.
- Wong, W.F.J., Chiu, L.S., 2008. Spatial and Temporal Analysis of Rain Gauge Data and TRMM Rainfall Retrievals in Hong Kong. *Annals of GIS* 14 (2), 105–112.
- Xie, P., Arkin, P.A., 1995. An Intercomparison of Gauge Observation and Satellite Estimates of Monthly Precipitation. *Journal of Applied Meteorology* 34.
- Yong, B., Ren, L.-L., Hong, Y., Wang, J.-H., Gourley, J.J., Jiang, S.-H., Chen, X., Wang, W., 2010. Hydrologic evaluation of Multisatellite Precipitation Analysis standard precipitation products in basins beyond its inclined latitude band: A case study in Laohahe basin, China. *Water Resour. Res.* 46 (7), n/a-n/a.
- Zhang, Y., Long, C.N., Rossow, W.B., Dutton, E.G., 2010. Exploiting diurnal variations to evaluate the ISCCP-FD flux calculations and radiative-flux-analysis-processed surface observations from BSRN, ARM, and SURFRAD. *J. Geophys. Res.* 115 (D15), 38.

Highlights

- Evaluation of GPM-IMERG version-3 products, Real-Time late-run and Final-Run:
- With regard to probability density functions (PDFs) the satellite precipitation products detected more heavy and extreme precipitation events than the ground measurements.
- IMERG-FR, systematically underestimates moderate to extreme and overestimates light precipitation for the entire range of precipitation ($P \geq 0.1$ mm) over Northeastern Austria.
- Despite the general low probability of detection (POD) and thread score (TS) and high false alarm ratio (FAR) within specified precipitation thresholds, the contingency table showed relatively good values of POD, TS and FAR for precipitation without precipitation classification ($P \geq 0.1$ mm).