Global Precipitation Measurement data compared to ground based disdrometer measurements

A case study in the Netherlands and Myanmar

Dorien Frederique Honingh
4208188

Delft University of Technology
Faculty of Civil Engineering and Geosciences

September 30, 2016

Supervision by:

Prof. dr. ir. Nick van de Giesen Delft University of Technology
Dr. ir. Martine Rutten Delft University of Technology

TU Delft

Contents

This report consists of two parts. The first part is the main article, which is called ‘Global Precipitation Measurement data compared to ground based (Delft-) disdrometer measurements’. The second part consists of supporting documentation for the main article and is named ‘Python scripts’.

Main article

Global Precipitation Measurement data compared to ground based disdrometer measurements

A case study in the Netherlands and Myanmar

D. F. Honingh

Delft University of Technology, Department of Watermanagement, Stevinweg 1, Delft, The Netherlands

Abstract

Rapid advancements in satellite observation technologies have resulted in an unprecedented availability of remotely sensed rainfall characteristics. These freely available precipitation measurements could be of great value for improving current weather prediction, particularly in developing countries where other data is scarce. A key parameter to derive rainfall intensities is the drop size distribution (DSD). Data of the Global Precipitation Measurement (GPM) mission provides amongst others the DSD and rainfall intensity, but it is unknown how well the device can convert the DSD to rain intensity. Therefore the main goal of this study is validating the GPM measurement results, by using ground based disdrometers. The disdrometers were located in De Bilt (The Netherlands), Yangon and Bago (Myanmar). In the Netherlands an industrial standard Thies LPM was used, whereas in Myanmar innovative Delft-disdrometers were tested. DSD's obtained from the Thies LPM showed a decreasing peak for increasing rainfall intensities, which corresponds to DSD theory. In most cases the DSD's of the Thies LPM and GPM were comparable with an average R^2 of 0.9. On the other hand, the rainfall intensities for these two datasets were not comparable with a relative error varying between 49 and 3314 %. The very limited joint occurrence of satellite flyover and rain events caused insufficient data for validation in Myanmar. The data obtained from the Delft-disdrometer in Yangon showed that, despite the limitations of this innovative device, it is possible to produce representative DSD.

Keywords: Rainfall, Drop size distribution, Satellite remote sensing, Delft-disdrometer, Thies LPM

1. Introduction

Myanmar is considered to be one of the most vulnerable countries in terms of natural disasters in Southeast Asia. The country is in second place on the Global Climate Risk Index, which is based on the number of extreme events that took place between 1995 and 2014. The annual floods during the monsoon period are one of those natural disasters that count as extreme events. For example, 460.000 people were affected and 172 people were killed by floods in 2015. Therefore, disaster risk reduction and other activities to prepare and respond to natural disasters are of great importance for Myanmar. (United Nations Humanitarian Country Team, 2015)

In Myanmar, weather forecasting is done by the Department of Meteorology and Hydrology (DMH). They give a daily weather forecast, but this is quite general and only mentions the expected number of showers for a region. For those forecasts, the DMH uses data from weather stations, satellite images and the regional weather predictions of neighbouring countries (DMH, 2016). The Weather Research Forecast (WRF) model is used to process the station data. However, the country covers 677.000 km^2 (Embassy of the Republic of the Union of Myanmar, 2016) and there are only 63 Meteorological Stations and 39 Meteorological & Hydrological Stations. Therefore the possibilities to validate and calibrate the model are very limited. As the WRF model has a resolution of 30 km by 30 km, it is not possible to make very accurate predictions. Making better use of satellite images would enable more accurate weather predictions. By using satellite images the calibration of the rainfall model can be improved with less rain gauges required. For example, National Aeronautics and Space Administration (NASA) satellites could be used to study monsoon patterns more closely and improve the weather forecast in Myanmar (NASA, 2016).

Tropical regions have the world's highest amount of rainfall. To study rainfall patterns, the Tropical Rainfall Measurement Mission (TRMM) was launched in November 1997: a joint mission between NASA and the Japan Aerospace Exploration Agency (JAXA). The rain intensity, cloud water and water vapour in the atmosphere were determined by measuring the Earth's emitted microwave energy. In order to measure this reflected energy and convert this to temperature, the TRMM Microwave Imager was used. Water surfaces emit limited microwave energy, which corresponds to 50 % of its real temperature. The measured temperature for raindrops on the other hand is almost equal to their real temperature. The contrast between cold water surfaces and warmer raindrops makes it possible to determine rainfall intensities quite accurately. Above the land surface more microwave energy is emitted, which corresponds to 90 % of its real temperature. This leads to a smaller contrast and therefore it is harder to observe the raindrops above land surface. (NASA, 2016)

The TRMM data became the space standard for precipitation measurements. Furthermore, the TRMM mission improved our knowledge on tropical cyclones, convective rainfall and human impact on rainfall. In addition, it has been useful for weather and flood forecasting. However, the mission ended in April 2015. (NASA, 2015)

The Global Precipitation Measurement (GPM) mission is the successor to the TRMM mission and data collection started in February 2014. The GPM is capable of measuring a larger area in more detail. The area covered is now between 65°S to 65°N whereas TRMM covered 50°S and 50°N. In addition to the larger coverage of the GPM, smaller intensities and other quantities such as snow can be detected. (Hou, et al., 2014)

The commonly used approach for validating satellite precipitation data is comparing it with ground observations. In this case, the ground observations are taken as the best possible standard for comparison even though they have their own uncertainties. This approach has also been used in the five field campaigns performed with Two-Dimensional Video Disdrometers. The use of disdrometers is preferred over other rain gauges because they measure both the DSD and the rain intensity. The following five campaigns have been conducted; the Light Precipitation Evaluation Experiment (LPVEx), the Mid-Continent Convective Clouds Experiment (MC3E), the GPM Cold-season Precipitation Experiment (GPCEx), the Iowa Flood Studies (IFloodS) and the Integrated Precipitation ad Hydrology Experiment (IPHEx). The results of these studies have been used to improve the GPM algorithms.

In the IFloodS study, the uncertainties in the drop size distribution (DSD) obtained by the GPM Dual-frequency Precipitation Radar (DPR) were investigated. For this, 1.5 month of GPM rain rate data was collected and compared with 11 Parsivel disdrometers. One of those comparisons was between the GPM rain rates and the rain rate data of the disdrometers. Scatter plots were created, which showed that the majority of the measurement points were near the y=x line. The correlation coefficient that was found for rain intensity was 0.987 and for the particle diameter this was 0.861. The findings of this study might not be representative for different climate regions, so therefore the authors recommend performing similar studies for different areas. (Liao L. &., 2014)

In this research project a similar study with disdrometers will be performed in the maritime climate of the Netherlands and in the monsoon climate of Myanmar. In Myanmar, the newly developed Delft-disdrometer will be installed, which is much cheaper compared to the traditional disdrometer. Having affordable disdrometers available would be a benefit for GPM ground validation (Koertellis, 2005). The reason why disdrometers are being used and preferred to other devices is that this type of rain gauge can measure the DSD. The DSD ties all the rainfall variables together and is therefore very valuable (Uijlenhoet, 1999). According to the GPM algorithms, a gamma function is used for the DSD. In order to use this gamma function, two DSD parameters are required. After obtaining the DSD, the GPM rainfall rates can be derived by combining DSD's, drop volumes and drop velocities and therefore contain multiple possible errors. From a scientific viewpoint it is therefore of great interest to assess differences between the GPM data and ground based disdrometer measurements. This will result in an indication of the added error due to the conversion from DSD to rain intensity.

1.2. Objective

The main objective of this research is to validate the GPM data in the maritime climate of the Netherlands and monsoon climate of Myanmar.

1.3. Research questions

In this study, GPM satellite data will be compared with disdrometer data. The purpose of comparing these datasets is to find out how accurate the GPM data is and where it can be improved. However, before making this comparison it is important to verify the accuracy of the Delft-disdrometer. The main research questions can thus be formulated as follows:

- *How accurate are the rainfall intensities measured by the Delft-disdrometer?*
- *Does the theoretical DSD for different rainfall intensities correspond with the measured disdrometer data?*
- *Does the theoretical DSD for different rainfall intensities correspond with the DSD's determined by GPM?*
- *What is, according to an analysis of eight GPM measurement locations, the spatial variance in common rain intensities in Myanmar?*
- *How accurate is the by GPM obtained gamma distribution?*
- *How accurate is rainfall intensity data obtained by GPM?*

2. Study locations

2.1. The Netherlands

The Netherlands has a maritime climate and the land is covering 41,543 km^2. In the Netherlands the Koninklijk Nederlands Meteorologisch Instituut (KNMI) is the Dutch knowledge institute for weather, climate and seismology. The main office of the KNMI is located in De Bilt, which is also one of the 35 measurement stations that are spread throughout the country.

De Bilt has been chosen as a study location due to the fact that this station is equipped with advanced precipitation measurement devices. One of those instruments is the Thies LPM, which can be used to measure the DSD. More information about this device can be found in section 3.1.3.

2.2. Myanmar

Myanmar covers 676,578 km^2 and has a tropical monsoon climate. There are considerable differences in the spatial average annual rain. In the delta region this amount is approximately 2500 mm (Yangon 2700 mm), in the coastal region 5000 mm, while in the dry zone the average annual rainfall is less than 1000 mm (Mandalay 850 mm). (Weather and climate information, 2016)

Two locations were used to perform ground based measurements using Delft-disdrometers, namely the Yangon Technological University (YTU) in Yangon and Irrigation Technological Centre (ITC) in Bago. The placement coordinates are given in Table 1.

Table 1: Coordinates (latitude and longitude) of the Delft-disdrometers in both Yangon and Bago, Myanmar.

Location	Latitude	Longitude
1. Yangon	N16º52'30"	E96º07'04.799"
2. Bago	N14º05'60"	E98º12'17.999"

Furthermore, eight study locations have been chosen for GPM data analysis. Those places consist out of four (major) cities and four rural areas, which are scattered throughout Myanmar. The geographical specifications of those eight places are given in Table 2.

Table 2: Coordinates of the eight study locations in Myanmar.

Location	Latitude	Longitude
1. Mandalay	N21º56'24.0"	E96º05'24.000"
2. Naypyiadaw	N19º46'04.8"	E96º06'07.200"
3. Yangon	N16º52'30.0"	E96º07'04.799"
4. Dawei	N14º05'60.0"	E98º12'17.999"
5. Mohnyin	N25º30'00.0"	E96º40'29.999"
6. Kengtung	N21º19'30.0"	E99º38'24.000"
7. Nat Ma Taung National Park	N21º15'00.0"	E93º24'18.000"
8.Lenya National Park	N11º42'18.0"	E99º15'00.000"

Figure 1 indicates the spatial distribution of those eight study locations. In this figure the cities are represented by a star and the dots stand for rural areas.

Figure 1: The spatial distribution of the eight study locations in Myanmar. In this figure cities are represented by a star and dots represent rural areas.

3. Materials and methods

3.1. Materials

Precipitation data was collected with the GPM Core Observatory satellite, disdrometers and tipping buckets. The specifications of these materials are explained below. The obtained data was processed by using Python 3.5.0. (Python, 2016) and Excel 2007 (Microsoft, 2016). Furthermore, Panoply 4.5.1 (Schmuck, 2016) was used for initial assessment of the obtained GPM data.

3.1.1. GPM

The mission started with the launch of the GPM Core Observatory satellite in Japan on February 27th, 2014. The Core Observatory satellite is the reference standard for approximately eight constellation satellites. Several of them, deliver precipitation measurements every three hours. (JAXA, 2016)

The Core Observatory satellite flies a non-Sun-synchronous orbit at an altitude of 407 km; this path covers the earth from 65°S to 65°N. The satellite orbits the earth 16 times a day, so each orbit takes 1.5 hours. For measuring precipitation, the GPM Core Observatory satellite is using the GPM Microwave Imager (GMI) and the Dual-frequency Precipitation Radar (DPR). The GMI delivers horizontal patterns and rainfall intensities, while the DPR provides information about the three dimensional structure of droplets. (NASA, 2016)

There are 3 different levels of GPM data products available, which are described in appendix A. From the available GPM data, only level 2A-DPR contains DSD parameters and was thus used in this research. Data has been accessed from the Science Team On-Line Request Module (STORM) data source portal (Lammers, 2016).

The DPR measures the reflectivity; this is translated to the DSD using the following formulas:

$$Z_e = \frac{\lambda^4}{\pi^5\, IKI^2} \int \sigma_b(D) N(D) dD \tag{1}$$

$$K = \frac{m^2 - 1}{m^2 + 2} \tag{2}$$

$$N(D) = N^* \cdot D^{\mu} \cdot \exp[-\frac{(3.67 + \mu)D}{D^*}] \tag{3}$$

In which Z_e [mm^6/m^3] is the effective radar reflectivity, λ [m] the electromagnetic wavelength, K [-] a constant defined as a function of m [-] which is the complex refractive index of scattering particles. The $\sigma_b(D)$ [-] is the particle reflection, $N(D)$ [m^3·mm] the drop size distribution, N^* [-], D^* [-] the rainfall distribution parameters and D [mm] the diameter of the droplet. The μ [-] is the shape parameter, which is fixed at 3 for the gamma distribution (Liao L. , 2014).

The following formula is used to transfer the obtained parameters to rainfall rates:

$$R = \int V(D) v(D) N(D) dD \tag{4}$$

In which the R [mm/h] is the rainfall rate, the $V(D)$ [m^3] and the $v(D)$ [mm/h] the particle volume and the particle velocity, both depending on the diameter of the droplet.

3.1.2. Delft-disdrometer

Disdrometers are capable of measuring DSD and convert these to rainfall intensities. However, the price for a conventional disdrometer starts from about €5000, while the commonly used tipping bucket is already available from €400. For this reason, a new kind of disdrometer has been designed, called the "Delft-Disdrometer". This type of rain gauge has been designed by combining the best of both. It is namely able to measure the acoustic energy (and therefore the drop size distribution) and the costs are approximately between €200 and €300. (Hut, 2013)

The smallest droplets that can be measured by the Delft-disdrometer have a diameter of 1 mm. For droplets larger than 1.75 mm, the uncertainties of the Delft-disdrometer are comparable to the uncertainties of industrial standards (Thies LPM & Ott Pasivel). Compared with a KNMI electronic rain gauge, the measurement difference falls within a 10 % interval. (Hut, 2013)

Furthermore, disdrometers can suffer from uncertainties due to a certain deadtime. This phenomenon has been investigated for the Joss-Waldvogel disdrometer that has a resonance time of several milliseconds, just like the Delft-disdrometer. The findings of this research were that in case deadtime correction was applied, the amount of precipitation of an event increased with nearly 15 % and the rainfall intensity increased with almost 35 %. (Uijlenhoet et al., 2002)

Another error is caused by stiffness variations within the disdrometer casing, which is comparable to the 'drum effect'. In appendix B, a simplified derivation of the relative error caused by this effect is given. This showed that smaller droplets suffer more from the drum effect.

3.1.3. Thies Laser Precipitation Monitor (LPM)

The Thies LPM (disdrometer) is capable of measuring precipitation particles with a diameter from 0.16 till 8 mm, a fall speed ranging from 0.2 till 20 m/s and an intensity from 0.005 mm/h till 1000 mm/h. It is capable of identifying more than 97 % of the drizzle, 99 % of the rain, 95 % of hail and 99 % of snow. (Adolf Thies GbmH & Co. KG, 2011)

The device has a measurement area of 46 cm^2, uses a laser of class 1M and has a laser diode of 785 nm. There are 440 measurement classes, namely 22 particle diameter classes and 20 speed classes. (Adolf Thies GbmH & Co. KG, 2011)

3.1.4. Tipping bucket

The tipping bucket is one of the most frequently used rain gauges for measuring rainfall. A tipping bucket has a bucket in which rainfall is collected and produces a measurement when a certain amount of water that has been collected. The used rain collector with product number 7852 has a collection area of 214 cm^2 and records every 0.2 mm of rainfall. Furthermore, it delivers continuous rainfall data, with a chosen time resolution of 5 minutes. (Rain Collector # 7852 & 7852M, 2004)

The accuracy of the tipping bucket for rain rates up to 50 mm/hr is equal to the greatest value of 4 % of the total or 0.2 mm. For rain rates from 50 mm/hr to 100 mm/hr, the accuracy is equal to $\pm$5 % of the total intensity or 0.2 mm, whichever is the biggest value. (Rain Collector # 7852 & 7852M, 2004)

3.2. Method

In order to answer the research questions, a number of steps will be carried out. The steps consist of:

1. Comparing Delft-disdrometer data with tipping bucket data for rain intensity control
2. Comparing theoretical DSD's with measured DSD's
3. Analysing historical GPM data for Myanmar
4. Validating GPM satellite data

3.2.1. Comparing Delft-disdrometer data with tipping bucket data for rainfall intensity control

A previous study with Delft-disdrometers, tipping buckets and manual rain gauges in Yangon and Bago, showed divergent results in rain intensity (Honingh, 2016). Therefore, rainfall intensities measured with the Delft-disdrometer are re-calibrated using the rainfall intensities of the tipping buckets.

In this research, the rainfall intensities measured by both the tipping bucket and Delft-disdrometer will be checked again. The time resolution for comparing the measured rain amount of Delft-disdrometer and tipping bucket data was 5 minutes.

Scatter plots were constructed to visualize differences between the two measurement methods. Furthermore, the mean bias error (MBE), the root mean squared error (RMSE) and the coefficient of determination (R^2) were calculated. Finally, a regression analysis was performed to investigate the relationship between the datasets.

The formula for calculating the mean bias error is:

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \tilde{y}_i) \quad (5)$$

In which y_i is the measurement and $\tilde{y}_i$ is the prediction. The formula for calculating the root mean squared error is:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2} \quad (6)$$

In which y_i is the measurement and $\tilde{y}_i$ is the prediction. The formula for calculating the coefficient of determination is:

$$R^2 = 1 - \frac{\sum(y_i - \tilde{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad (7)$$

In which y_i is the measurement, $\bar{y}$ is the average of the measurements and $\tilde{y}_i$ is the prediction. Finally, the formula for the regression analysis is:

$$Y = A \cdot X + B \quad (8)$$

In which Y is the dependent variable, X the independent variable, A the slope coefficient and B the intercept with the Y-axis.

3.2.2. Comparing theoretical DSD's with measured DSD's

According to Uijlenhoet (1999), DSD's as shown in Figure 2 are to be expected. This figure shows that the higher the rain intensity the higher the probability of large drop sizes.

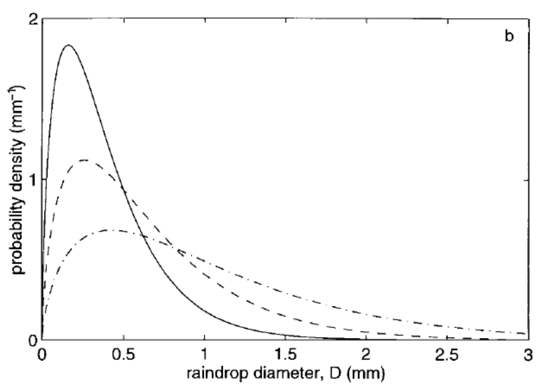

Figure 2: Theoretically expected DSD, in which the full line represents the intensity class of 1 mm/h, the striped line shows the intensity category of 10 mm/h and the dot stripe dot line stands for the rain intensity of 100 mm/h. (Uijlenhoet, 1999)

For both the Thies LPM, Delft-disdrometer and GPM, a check was performed to see if the measurement results were consistent with the theoretical expected DSD.

Both types of disdrometers produce binned measurement results. This implies that the disdrometers measure the droplet size and count how many droplets fall within a certain range. For the Thies LPM, research into binning effects on the DSD has been carried out (Garcia et al., 2014). This study showed that the effect of binning DSD data can lead to a maximum error of 5 % in rainfall and reflectivity during heavy rainfall intensities. Furthermore, a noticeable impact of binning methods on the DSD was reported.

First, research was carried out to check whether a gamma fit or a generalized gamma fit corresponded best to the binned measurement results. Second, rain intensity categories were made, with which probability density curves were created.

3.2.3. Analysing historical GPM data

Historical GPM data were analysed in order to find possible patterns. For this, the eight locations that were described in paragraph 2.2 were used. First, the measurements of all the locations are combined to see whether the by GPM measured DSD follows the in theory expected DSD. Second, the variance in measured rain intensities was compared for the different locations, for which box plots were used. This gives an indication for the possible spatial variation in precipitation.

3.2.4. Validating GPM satellite data

For GPM satellite data validation ground measured data, both the Thies LPM and the Delft-disdrometer data, are taken as the 'ground truth'. The one minute findings of the GPM are compared with the measurement results of the ground devices. For this, plots were made for DSD comparison. Beside this, the MBE, the RMSE, the R^2 and the regression coefficients were determined. For comparing the rain intensities, the relative error was calculated. The formula for the relative error is:

$$E_{relative} = \frac{I\,(y_i - \tilde{y}_i)\,I}{\tilde{y}_i}$$

In which y_i is the prediction and $\tilde{y}_i$ is the actual value.

4. Results and discussion

4.1. Comparing tipping bucket data with Delft-disdrometer data for rain intensity control

For the tipping bucket and Delft-disdrometer, the amount of precipitation for every five minutes was compared. In order to present the differences between these, scatter plots have been constructed, which are shown in Figure 3 and Figure 4.

In Figure 4, it can be seen that for low rainfall intensities the y=x line is a good fit. However, the error increases seriously for higher rainfall intensities. For Figure 3, the error is more constant for different rain intensities.

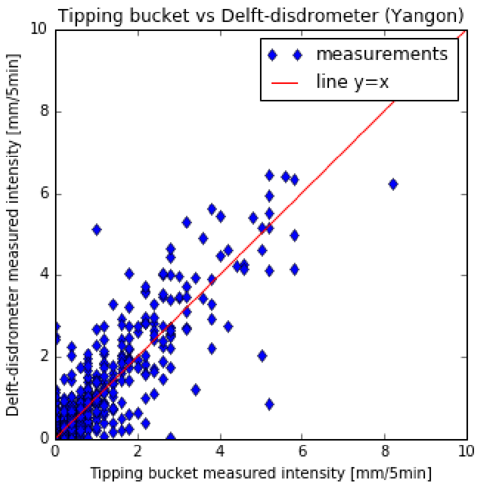

Figure 3: Scatter plot of tipping bucket & Delft-disdrometer comparison, locations: Yangon (left) and Bago (right).

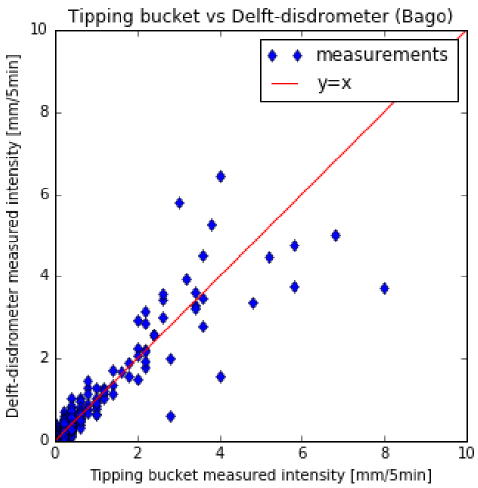

Figure 4: Scatter plot of tipping bucket & Delft-disdrometer comparison, location: Bago.

In addition, a statistical analysis has been performed, of which the results can be seen in Table 3. In this table, it can be seen that the results for the tipping bucket and Delft- disdrometer comparison are the best for location Yangon.

Table 3: Comparison of rain amount measured by the tipping bucket and the Delft-disdrometer in Yangon and Bago. In this table, the mean biased error (MBE), Root mean squared error (RMSE), coefficient of determination (R^2), the slope coefficient of the regression line (A) and the regression line its intercept with the Y-as (B) are given.

Location:	Yangon	Bago
Time span	11July-5Sept.	11-21 July
MBE [mm]	$-4.4\cdot10^{-3}$	$6.9\cdot10^{-4}$
RMSE [mm]	$1.3\cdot10^{-1}$	$2.1\cdot10^{-1}$
R^2 [-]	$8.3\cdot10^{-1}$	$7.8\cdot10^{-1}$
A [-]	$9.8\cdot10^{-1}$	$7.8\cdot10^{-1}$
B [mm]	$5.2\cdot10^{-3}$	$1.6\cdot10^{-2}$

4.2. Comparing theoretical DSD's with measured DSD's – Thies LPM

One year of Thies LPM data for the location De Bilt was analysed in order to compare the theoretical DSD with the measured DSD of the Thies LPM. In the Netherlands, high rain intensities are less common. Therefore different rain intensity classes are chosen compared to the ones used in Figure 2, but still a comparable figure is expected. The chosen intensity classes are 1 mm/h, 5 mm/h, 10 mm/h and 50 mm/h.

First, it was determined whether a gamma or a generalized gamma function would fit the measured data best. The results showed that in all cases the generalized gamma function outperformed the gamma function.

In Figure 5, the average bins of the different intensity classes are plotted to get a visual impression of the binned data of the LPM. Also, it shows how the data and the corresponding fits relate to each other. Both functions provide a good fit, but the generalized gamma results were slightly better. The reason why the generalized gamma fit gives a better result is because this fit has more parameters, which leads to more freedom to develop.

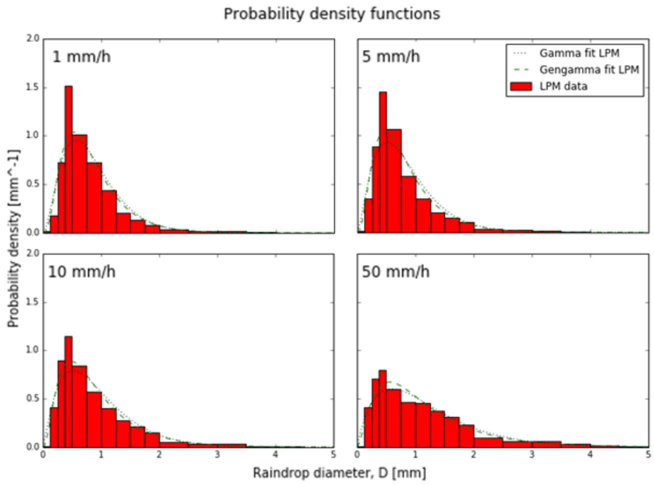

Figure 5: Comparing for every intensity class, the average bins of the LPM data with the gamma and generalized gamma fit. The average results for the R^2 of the gamma fit were respectively 0.71, 0.73, 0.78 and 0.84. The average results for the R^2 of the generalized gamma fit were respectively 0.84, 0.82, 0.84 and 0.85.

Subsequently, probability density curves with 90 % confidence bands were made using fitted generalized gamma distributions on the binned data. The results can be found in Figure 6, where it can be observed that larger droplets were measured for increasing intensities. There is no overlap in the peaks of the 90 % confidence bands, this leads to an increased confidence in that the DSD's follow the theory. Furthermore, it can be seen that the bands for 10 and 50 mm/h are bigger compared to the bands of 1 and 5 mm/h. This is because there are more measurement results for 1 and 5 mm/h, which results in smaller confidence bands.

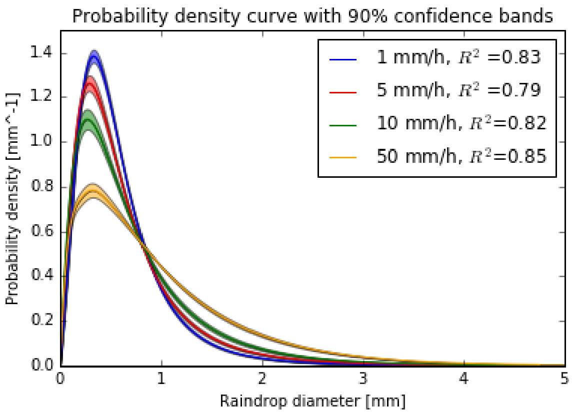

Figure 6: Probability density curve of the generalized gamma fit on the LPM data with 90 % confidence bands, the bands are set as follows; 0.99-1.01, 4.50-5.50, 9.50-10.50,40.0-60.0 mm/h.

The mean curve was compared with all the binned data, of which the R^2 is given in the legend of Figure 6. The best relation between the mean confidence band and the binned data was found for the intensity class of 50 mm/h, with an R^2 of 0.85.

4.3. Comparing theoretical DSD's with measured DSD's – Delft-disdrometer

Yangon

Two months of Delft-disdrometer data for the location Yangon was analysed. In Myanmar, higher rainfall intensities are more common than in the Netherlands. For this reason, the intensity of 100 mm/h replaced the intensity of 5mm/h. The data collection period was from 5 July 2016 until 5 September 2016.

The Delft-disdrometer is not capable of measuring droplets with a diameter smaller than 1 mm. Due to the recalibration, it can be observed that also bins larger than 1 mm are not well represented when comparing this to the theoretical DSD.

Deriving the droplets that have a diameter between 0 and 1 mm is therefore challenging when using the Delft-disdrometer. From Figure 7 it can be observed that directly fitting the generalised gamma and gamma distributions to the data resulted in shifted DSD's.

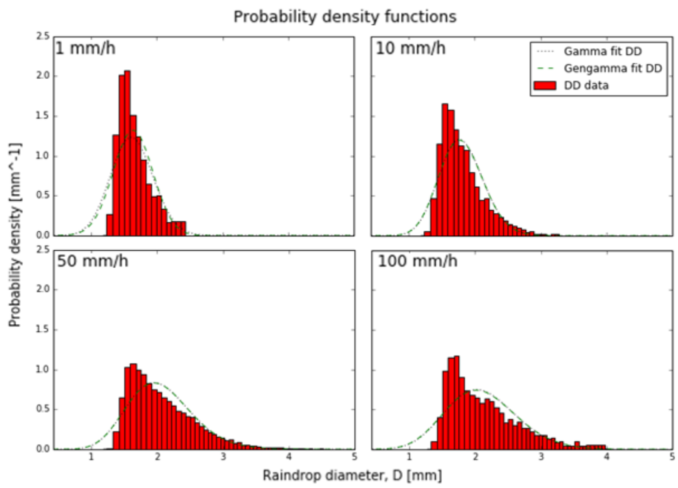

Figure 7: Comparing the binned Delft-disdrometer data with the gamma and generalized gamma fit. Droplets smaller than 1 mm could not be detected by the Delft-disdrometer.

A different method of fitting gamma and generalised gamma functions was therefore used. Due to the fact that the Delft-disdrometer only provides drop sizes larger than one mm, a function has to be fitted on the tail of the data. With this information it is possible to derive an estimate for drop size frequencies between 0 and 1 mm.

In order to determine which function would be the preferable one to use several tests were performed on

Thies LPM data. Only the functions with R^2 above 0.99 for the tail (D >1 mm) section were accepted. Subsequently, the best function was then selected based on the R^2 of data in the head (D < 1 mm) of the probability density functions. Results of rain intensities 1, 10 and 50 mm/h showed that the generalized gamma (R^2; 0.74, 0.66, 0.78) outperformed the gamma distribution (R^2; 0.10, 0.25, 0.51).

For rain intensities of 1, 10 and 50 mm/h, the Thies LPM in De Bilt classified 15, 24 and 37 % of all the rain drop diameters bigger than 1 mm. Furthermore, Uijlenhoet (1999) showed that approximately 45 % of the droplets during 100 mm/h intensity had a diameter bigger than 1 mm.

Combining these findings, the tail of the measured DSD of the Delft-disdrometer was selected and multiplied by its corresponding reduction factor. After this data transformation, new generalized gamma distributions were created. Only the function with a coefficient of determination larger than 0.99 with respect to the transformed data, were selected (see Figure 8).

From the figure below, it can be seen that the largest spread in each of the intensity classes is located in the part without measurements. This is because the tail of the distribution was measured, while the peak was not measured at all.

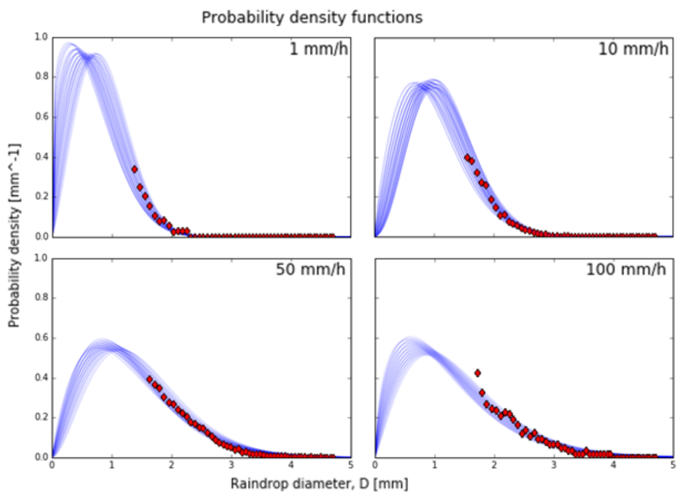

Figure 8: Representative probability density curves for the Delft-disdrometer in Yangon. Generalized gamma distributions were fitted on the measurement result for drop diameters bigger than 1 mm. Only the distributions which had on the measured interval a R^2 bigger than 0.99 were taken into account.

The benefit of plotting all the functions with a coefficient of determination larger than 0.99 compared to plotting only the best fit, is that the spread in potential DSD is visible. However, it is possible that the reduction factor for the Netherlands and Myanmar is different, which directly changes the tail and thus all potential fits. Therefore, additional research should be conducted with industrial standard disdrometers in several climates to determine the variation in the percentage of droplets with a diameter bigger than 1 mm.

Bago

For the location Bago, a remarkable spread in the droplets diameter was detected compared to theory and the previous results. Therefore, it can be concluded that recalibration of the Delft-disdrometer in Bago was unsuccessful. For this reason, the results found in Bago are not further used.

4.4. Analysing historical GPM data for 8 locations in Myanmar

In Table 4, the number of rain events measured by the GPM satellite are given for each study location. Data was collected between March 9 2014 and August 9 2016.

Table 4: Number of GPM rain measurements between March 9 2014 and August 8 2016, for 8 locations throughout Myanmar.

Location	Abbreviation	# GPM satellite rain detections
1. Mandalay	MAN	10
2. Naypyiadaw	NAP	9
3. Yangon	YAN	17
4. Dawei	DAW	23
5. Mohnyin	MOH	9
6. Kengtung	KEN	15
7. Nat Ma Taung National Park	NAT	31
8.Lenya National Park	LEN	18

In Figure 9 the measured DSD's of all eight locations are combined. Measured rain intensities were, for this historical analysis, on average quite low. For this reason the rain intensity was binned in the following ranges; 0-1 mm/h, 1-5 mm/h and > 5 mm/h. It can be seen that on average the diameter of the droplets increases and the height of the probability density curve thus decreases with increasing precipitation intensity. This is as expected and follows from the theory discussed in 3.2.2.

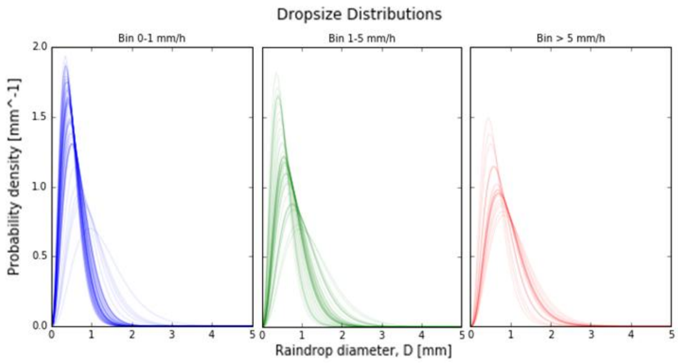

Figure 9: Combined drop size distributions of all 8 study locations in Myanmar. Data was collected between March 9 2014 and August 8 2016.

In Figure 10, box plots show the spread in measured rain intensity by the GPM for every study location. In Mandalay, which is located in the dry zone, very low rain intensities were measured. Interestingly, results from other locations outside the dry zone show very low rain intensities too. Due to the limited GPM measurements, it is possible that high rain events are common, but remain undetected. Therefore, it is not possible to draw reliable conclusions about possible differences in common rain intensities or DSD's.

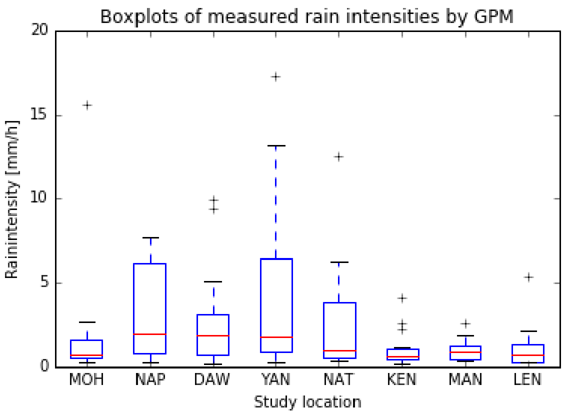

Figure 10: Box plots of the rain intensities measured by the GPM for the 8 study locations in Myanmar. Measurement period was between March 9 2014 and August 8 2016.

4.5. Comparison GPM data with LPM ground measurements of the KNMI (De Bilt, Netherlands)

Measurement statistics concerning the Thies LPM in De Bilt are given in the Table 5. Between March 2014 and March 2016, the GPM has 242 measurements for the location De Bilt. Of those measurements, only 11 times precipitation was measured. In one of those cases, the Thies LPM did not measure precipitation at all. Vice versa, in 15 measurements where rain was detected by the Thies LPM, no rain was detected by the GPM.

Research showed that larger errors in GPM products occur for higher latitudes and during the winter period (AghaKouchak et al., 2012). However, the measured errors are not well documented in literature. Most papers only mention that the validation results were used to improve the GPM algorithms.

Table 5: Measurement statistics of GPM and Thies LPM comparison for location De Bilt. Data was collected between March 2014 and March 2016.

	2014	2015	2016
Number of GPM satellite measurements with rain	4	6	1
Number of GPM satellite measurements without rain	89	134	8
Number of measurements where the LPM measured rain while no rain was measured by the GPM satellite	3	12	0

During the GPM satellite measurement of 12-05-2014, the LPM in De Bilt was undergoing maintenance. Measurements of both the LPM and GPM were only taken into account when two values could be used for comparison. Furthermore, the measurement of 30-01-2015 where solid precipitation was measured was also not taken into account. Without those, eight events were compared and the results are shown in Table 6 and Figure 11.

For all eight events, the LPM data was best modelled by the generalized gamma fit. For three of the eight compared measurements, the measured rain intensities were of the same order of magnitude. In the other situations, both under and over estimation of GPM occurred. The relative errors that were made by the under-and overestimations in rain intensity vary between 49 and 3314 %.

For six results (08/07/2014, 16-10-2014, 01/02/2015, 25-03-2015, 16-10-2015, 28-11-2015), the GPM DSD's matched with the LPM DSD's. However, for only three (08/07/2014, 16-10-2014, 01/02/2015) of those measurements the rain intensity matched, while for the other cases the rain intensity was underestimated by the GPM. During one event (07-03-2016), where the GPM had measured a lot of small droplets and fewer big droplets compared to the LPM measurements, the corresponding rain intensity was remarkably enough overestimated. For the last discussed event (23/03/2014), both the DSD and rain intensity were overestimated.

Table 6: GPM rain measurement overview and comparison to LPM findings (Location De Bit). In this table, the mean biased error [mm], Root mean squared error [mm], coefficient of determination [-], the slope coefficient of the regression line [-] and the regression line its intercept with the Y-as [mm] are given.

Date and time	GPM rain rate [mm/h]	LPM rain rate [mm/h]		Gamma fit LPM data	Generalized gamma fit LPM data	GPM gamma distribution
23-03-2014 12:35	2.97	0.087	MBE	$1.9\cdot10^{-3}$	$6.1\cdot10^{-4}$	$8.7\cdot10^{-2}$
			RMSE	$1.3\cdot10^{-1}$	$1.1\cdot10^{-1}$	$5.0\cdot10^{-1}$
			R^2	$9.4\cdot10^{-1}$	$9.6\cdot10^{-1}$	$3.0\cdot10^{-1}$
			A	$6.3\cdot10^{-1}$	$6.4\cdot10^{-1}$	$2.6\cdot10^{-1}$
			B	$1.1\cdot10^{-1}$	$1.1\cdot10^{-1}$	$1.8\cdot10^{-1}$
08-07-2014 05:24	0.31	0.371	MBE	$2.3\cdot10^{-2}$	$2.2\cdot10^{-2}$	$-2.4\cdot10^{-2}$
			RMSE	$3.0\cdot10^{-1}$	$3.0\cdot10^{-1}$	$1.9\cdot10^{-1}$
			R^2	$7.6\cdot10^{-1}$	$7.7\cdot10^{-1}$	$8.8\cdot10^{-1}$
			A	$6.4\cdot10^{-1}$	$7.0\cdot10^{-1}$	$1.1\cdot10^{0}$
			B	$1.0\cdot10^{-1}$	$8.6\cdot10^{-2}$	$-1.4\cdot10^{-2}$
16-10-2014 07:24	1.34	1.380	MBE	$1.9\cdot10^{-3}$	$6.1\cdot10^{-4}$	$1.7\cdot10^{-3}$
			RMSE	$1.3\cdot10^{-1}$	$1.1\cdot10^{-1}$	$1.4\cdot10^{-1}$
			R^2	$9.4\cdot10^{-1}$	$9.6\cdot10^{-1}$	$9.4\cdot10^{-1}$
			A	$8.3\cdot10^{-1}$	$8.6\cdot10^{-1}$	$8.1\cdot10^{-1}$
			B	$5.3\cdot10^{-2}$	$4.3\cdot10^{-2}$	$5.9\cdot10^{-2}$
01-02-2015 23:24	0.51	0.504	MBE	$9.4\cdot10^{-4}$	$4.7\cdot10^{-4}$	$-3.2\cdot10^{-2}$
			RMSE	$1.3\cdot10^{-1}$	$1.1\cdot10^{-1}$	$1.7\cdot10^{-1}$
			R^2	$8.8\cdot10^{-1}$	$9.2\cdot10^{-1}$	$9.3\cdot10^{-1}$
			A	$7.9\cdot10^{-1}$	$8.5\cdot10^{-1}$	$1.3\cdot10^{0}$
			B	$6.4\cdot10^{-2}$	$4.6\cdot10^{-2}$	$-6.1\cdot10^{-2}$
25-03-2015 01:32	0.68	2.678	MBE	$1.4\cdot10^{-2}$	$1.1\cdot10^{-2}$	$2.0\cdot10^{-2}$
			RMSE	$2.2\cdot10^{-1}$	$1.9\cdot10^{-1}$	$2.4\cdot10^{-1}$
			R^2	$9.0\cdot10^{-1}$	$9.2\cdot10^{-1}$	$8.8\cdot10^{-1}$
			A	$7.5\cdot10^{-1}$	$7.7\cdot10^{-1}$	$7.1\cdot10^{-1}$
			B	$7.5\cdot10^{-2}$	$6.5\cdot10^{-2}$	$8.5\cdot10^{-2}$
16-10-2015 13:18	0.72	1.4	MBE	$6.0\cdot10^{-3}$	$5.1\cdot10^{-3}$	$-6.3\cdot10^{-3}$
			RMSE	$1.3\cdot10^{-1}$	$1.1\cdot10^{-1}$	$1.3\cdot10^{-1}$
			R^2	$9.4\cdot10^{-1}$	$9.6\cdot10^{-1}$	$9.3\cdot10^{-1}$
			A	$8.3\cdot10^{-1}$	$8.7\cdot10^{-1}$	$8.7\cdot10^{-1}$
			B	$4.9\cdot10^{-2}$	$3.5\cdot10^{-2}$	$5.0\cdot10^{-2}$
28-11-2015 00:56	0.30	3.453	MBE	$2.0\cdot10^{-2}$	$1.7\cdot10^{-2}$	$2.8\cdot10^{-2}$
			RMSE	$3.0\cdot10^{-1}$	$2.7\cdot10^{-1}$	$2.8\cdot10^{-1}$
			R^2	$7.3\cdot10^{-1}$	$7.9\cdot10^{-1}$	$7.5\cdot10^{-1}$
			A	$5.5\cdot10^{-1}$	$6.3\cdot10^{-1}$	$6.4\cdot10^{-1}$
			B	$1.3\cdot10^{-1}$	$1.1\cdot10^{-1}$	$9.5\cdot10^{-2}$
07-03-2016 02:36	0.29	0.097	MBE	$1.3\cdot10^{-3}$	$1.2\cdot10^{-3}$	$-7.3\cdot10^{-2}$
			RMSE	$1.7\cdot10^{-1}$	$1.6\cdot10^{-1}$	$3.7\cdot10^{-1}$
			R^2	$8.1\cdot10^{-1}$	$8.4\cdot10^{-1}$	$5.3\cdot10^{-1}$
			A	$7.0\cdot10^{-1}$	$7.5\cdot10^{-1}$	$1.0\cdot10^{0}$
			B	$8.5\cdot10^{-2}$	$7.2\cdot10^{-2}$	$6.7\cdot10^{-2}$

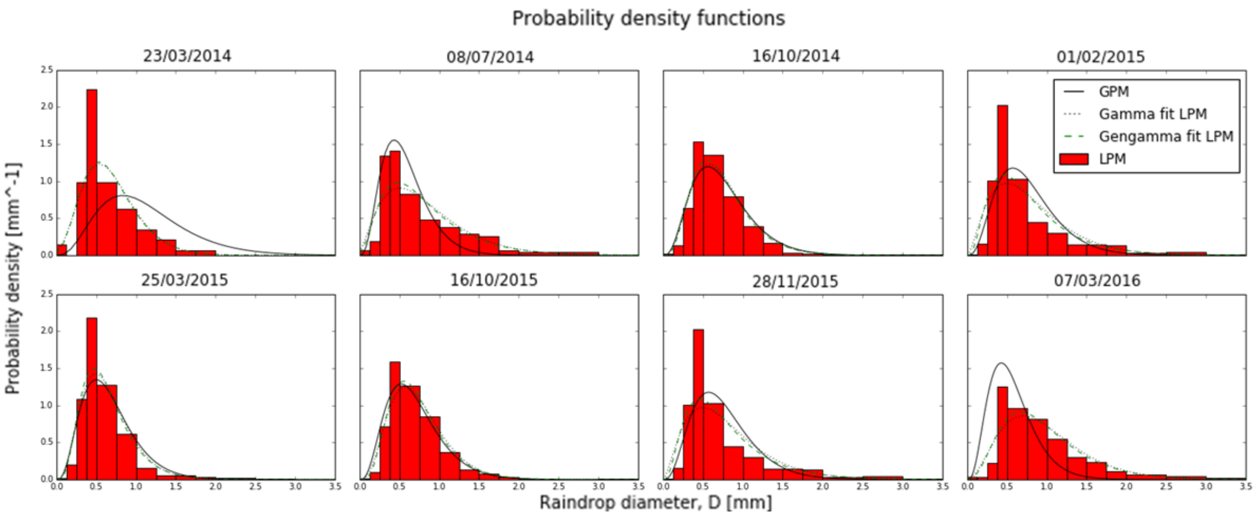

Figure 11: Comparison of probability density curves LPM and GPM (Location: De Bilt).

4.6. Comparison GPM data with Delft-disdrometer measurement results (Yangon, Myanmar)

During GPM flyover, no rain was measured by the Delft-disdrometer and tipping bucket in Yangon. For this reason, it is not possible to compare Delft-disdrometer measurement results with GPM findings.

See Table 7 for the data collection time span, the number of times GPM fly over and the number of rain measurements by the GPM and Delft-disdrometer.

Table 7: Comparing GPM and Delft-disdrometer statistics for Yangon, between 5 July and 5 September 2016.

	2016
Number of times the GPM fly over	14
Number of rain events measured by the GPM at the location	3
Number of rain events measured by the Delft-disdrometer during GPM fly over	0

During two of the three events where the GPM measured rain, the intensities where around 0.5 mm/h. However, during the last event, a rain intensity of 9.4 mm/h was measured.

5. Summary and conclusions

Validation of the GPM results for DSD's and rain intensities was the main objective in this study. For this, two different ground based disdrometers were used. Before comparing the GPM and disdrometer results, three steps were carried out to assess the results of each device. Firstly, rain amounts measured by the Delft-disdrometer were compared to tipping bucket measurements. Secondly, ground based measured DSD's of the Delft-disdrometer and Thies LPM were compared to the theoretical DSD's. Thirdly, space borne GPM data was studied. Finally, the satellite measurement results of the GPM were compared to the disdrometers measurements.

The accuracy that was found for the Delft-disdrometer in Yangon was better than the accuracy of the one in Bago. The results of comparing the Delft-disdrometer intensity data with tipping bucket intensity data were a R^2 of 0.83 for Yangon and a R^2 of 0.78 for Bago. Interestingly enough, the Delft-disdrometer in Bago measured the low rainfall intensities quite accurately, but had a significant larger deviation that occurred for rain intensities higher than 3 mm/h. For Yangon, both low and high rainfall intensities had only small discrepancies between the two types of measurements.

The DSD's measured by the Thies LPM in De Bilt were similar to the theoretical DSD's. The Delft-disdrometer in Yangon was unable to measure drop diameters smaller than 1 mm. Therefore a significant amount of droplets were not detected, for instance the Thies LPM in De Bilt classified for different intensities 85 till 63 % of all rain drop as smaller than 1mm. This resulted in an uncertainty in the height and exact location of the probability density curve peak. Nevertheless, the peak for smaller droplets decreased by increasing rain intensities, which corresponds with the theory.

For the Delft-disdrometer in Bago, the spread in measured droplets was much bigger than for both Yangon and De Bilt. This is probably due to calibration errors which impacts the sensitivity of the Delft-disdrometer. As a result, the fits were not comparable with the theory. For this reason the DSD results of Bago were not taken into account.

In addition to ground based measurements this paper showed the DSD measured by the GPM. Only low rainfall intensities were measured, but the produced DSD's already showed a decrease in peak height for increasing intensities, what corresponds with the theory. The detected spread in rainfall intensity is lower than expected, because for most locations, only low intensities were found. However, for the locations in the delta and coastal region (Yangon, Dawei and Lenya National Park), relatively higher rainfall intensities were detected. This is in line with the expectations based on the annual rainfall for different locations in Myanmar

Furthermore, the measurement results of the Thies LPM were compared to the findings of the GPM. For most events, the measured DSD's were comparable while the rain intensity only matched for half of the events. So it can be concluded, that in case of correct parameters the GPM gamma distribution can correspond well to the DSD measured by the Thies LPM. However, this doesn't deliver automatically the corresponding rain intensity. This means that especially the determination of the rain intensity could be improved.

Unfortunately it was not possible to compare the measurement results of the GPM and Delft-disdrometers, because there was no measurement with both devices simultaneously detecting rain. On the other hand it was remarkable that three measurements indicated the GPM measured rain, even though no rain was detected by the ground based devices.

Acknowledgements

Financial support for this research has been provided by the Delft University of Technology in combination with the funding of the EMMA exchange program between Europe and Asia.

The support, feedback and advice of Prof. dr. ir. Nick van de Giesen and Dr. ir. Martine Rutten within this research are greatly appreciated.

Finally, I would like to thank the KNMI for sharing the data of the Thies LPM in De Bilt.

References

Adolf Thies GbmH & Co. KG. (2011). *Laser Precipiation Monitor.* Gottingen: Thies CLIMA.

AghaKouchak, A., Mehran, A., Norouzi, H., & Behrangi, A. (2012). *Systematic and random error components in setellite precipitation data sets.* Irvine: Geophysical research letters.

Bagree, R. (2009). *Characterization and design of a readout circuit dor a piezoelectric-based acoustic disdrometer.* Delft: TU Delft.

Ciach, J. (2010). *Local Random Errors in Tipping-Bucket Rain gauge Measurments.* Lowa City: American Meteorological Society.

DMH. (2016). *Meteorological Division*. Retrieved 07 03, 2016, from Department of meteorology and hydrology (Myanmar): http://www.dmh.gov.mm/meteorological-division

Embassy of the Republic of the Union of Myanmar. (2016, July 13). *Basic facts about Myanmar*. Retrieved from Embassy of the Republic of the Union of Myanmar: http://www.myanmar-embassy-tokyo.net/about.htm

Garcia, R., Tokay, A., & Tabiador, F. (2014). *Binning effects on in-situ raindrop size distribution measurements.* Karlsruhe: Atmospheric Measurement Techniques Discussions.

Honingh, D. (2016, June 13). *Innovative rainfall measurement devices in tropical regions.* Retrieved from TU Delft repository: http://resolver.tudelft.nl/uuid:66fd52ee-ce26-4059-86f7-1c8f26fe8ff8

Hou, A., Kakar, M., Neeck, S., Azarbarzin, A., Kummerow, C., Kojima, M., & ... Iguchi, T. (2014). *The Global Precipitation Measurement Mission.* Boston: American Meteorological Society.

Hut, R. (2013). *New Observational Tools and Datasources for Hydrology.* Delft: TU Delft.

JAXA. (2016, September 5). *Global Precipitation Measurement/Dual-frequency Precipitation Radar*. Retrieved from Japan Aerospace Exploration Agency: http://global.jaxa.jp/projects/sat/gpm/index.html

Koertellis, A. (2005). *Disdrometer Calibration Using an Adaptive Signal Processing Algorithm.* Orlando: ASRC Aerospace.

Lammers, M. (2016, 8 28). *Data access*. Retrieved from STORM: https://storm.pps.eosdis.nasa.gov/storm/data/Service.jsp?serviceName=RestrictedOrder

Lanza, L. (n.d.). *On the quality of tipping bucket rain intensity measurements.* Genova (Italy): University of Genova, Dept. of Environmental Engineering.

Liao, L. &. (2014). *Uncertainties of GPM DPR Rain Estimates Caused by DSD Parameterizations.* Greenbelt: Journal of applied meteorology and climatology .

Liao, L. (2014). *Uncertainties of GPM/DPR rain estimates caused by DSD parameterizations.* Morgan: Journal of Applied Meteorology and Climatology.

Microsoft. (2016, August 25). *Excel*. Retrieved from Microsoft: https://products.office.com/nl-nl/excel

NASA. (2015, March 6). *TRMM*. Retrieved July 9, 2016, from NASA: http://trmm.gsfc.nasa.gov/

NASA. (2016, July 13). *GPM*. Retrieved from NASA: https://pmm.nasa.gov/data-access/downloads/gpm

NASA. (2016, July 10). *GPM data downloads*. Retrieved from Precipitation measurement mission: https://pmm.nasa.gov/data-access/downloads/gpm

NASA. (2016). *GPM Data Downloads*. Retrieved 07 04, 2016, from Precipitation Measurement Mission: https://pmm.nasa.gov/data-access/downloads/gpm

NASA. (2016, 02 13). *GPM Spacecraft and Instruments*. Retrieved 07 04, 2016, from NASA: http://www.nasa.gov/mission_pages/GPM/spacecraft/index.html

NASA. (2016). *Monsoons: Wet, Dry, Repeat*. Retrieved 07 03, 2016, from Precipitation measurement mission: https://pmm.nasa.gov/articles/monsoons-wet-dry-repeat

NASA. (2016, September 6). *TMI*. Retrieved from Precipitation measurement mission: https://pmm.nasa.gov/TRMM/TMI

Python. (2016, August 25). *Python is a programming language that lets you work quickly and intergrate systems more effectively*. Retrieved from Python: www.python.org

(2004). *Rain Collector # 7852 & 7852M.* Hayward U.S.A.: Davis Instruments.

Royale Geographical Society. (n.d.). *Spearman's Rank Correlation Coefficient - Excel Guide.* London: Royale Geographical Society.

Schmuck, R. (2016, August 25). *Panoply netCDF, HDF and GRIB Data Viewer*. Retrieved from NASA: http://www.giss.nasa.gov/tools/panoply/

TFA DOSTMANN. (2016). *Operating Instructions.* Ottersberg (Germany): TFA Dostmann GmbH & Co. KG.

Uijlenhoet, R. (1999). *A consistent rainfall parameterization based on the exponential raindrop size distribution.* Wageningen: Elsevier.

Uijlenhoet, R., Steiner, M., & Smith, J. (2002). *Influence of disdrometer deadtime correction on self-consistent analytical parameterizations for raindrop size distributions.* Wageningen: ERAD.

United Nations Humanitarian Country Team. (2015). *Humanitarian response plan.* Myanmar: United Nations.

Weather and climate information. (2016). *Weather averages and climate in Myanmar (Burma)*. Retrieved from Weather and climate information: https://weather-and-climate.com/average-monthly-Rainfall-Temperature-Sunshine-in-Myanmar-Burma

Appendix A: Overview of the three GPM data products levels

	Available products	Data description	Spatial resolution	Temporal resolution
Level 0	Raw instrument data			
Level 1	1A-GMI	GMI unprocessed instrument data, but with ancillary information like calibration coefficients and georeferencing parameters	4 km x 4km	16 orbits per day
	1B-GMI	Radiometrically and geolocated corrected 1A data	Varies by channel	16 orbits per day
	1C-GMI 1C-R	Common intercalibrated brightness temperature	Varies by channel	16 orbits per day
	1C-constellation	Same but with partner radiometers	Varies by satellite	Varies by satellite
Level 2	2A-GROF-GMI	Radar enhanced (RE) precipitation retrievals	4 km x 4km	16 orbits per day
	2A-GROF-constellation	Radar enhanced (RE) precipitation retrievals from 1C	Varies by satellite	Varies by satellite
	2A-Ku	DPR products Scan swath of 245 km	5.2 km x 5.2 km	16 orbits per day
	2A-Ka	Two scans/outputs: One scan: for comparing the central 25 beams of 2A-Ku Second scan: highly sensitive for light rain& snow	5.2 km x 5.2 km	16 orbits per day
	2A-DPR	Reflectivity Sigma Zero DSD Precipitation with vertical structure	5.2 km x 5.2 km	16 orbits per day
	2B-CMB	Precipitation Available time span: past two weeks	5 km x 5 km	
Level 3	3-GROF	2A-GROF data transformed in gridded data	0.25º x0.25º	Daily -monthly
	3-DPR	Gridded data of computed low resolution and high resolution data	0.25º - 0.5º	Daily -monthly
	3-CMB	Combined precipitation data (gridded)	0.25º - 0.5º	Daily -monthly
	IMERG	Merged product of GMI, partner radiometer and IR	0.1º	30 min - monthly

Table 8: Overview of the three GPM data products levels (NASA, 2016).

Appendix B: Assessing the Drum Effect - Impact of rain droplet

Terminal fall velocity:

$$V_t = \sqrt{\frac{2mg}{\rho A C_d}}$$

In which V_t [m/s] is the terminal fall velocity, m [kg] is the mass of the droplet, g [$\frac{m}{s^2}$] the acceleration due to gravity, ρ [$\frac{kg}{m^3}$] the density of the air, A [m^2] the projected area, C_d [-] the drag coefficient.

Kinetic energy:

$$E_k = \frac{m\, v^2}{2}$$

In which E_k[J] is the kinetic energy.

Conservation of energy:

$$\frac{m\, v^2}{2} \longrightarrow \frac{k\, x^2}{2}$$

In which k [$\frac{kg}{s^2}$] is the stiffness and x [m] is the maximum displacement from its equilibrium position. This physical process can be seen as a damped forced vibration.

Natural frequency:

$$f_n = \frac{1}{2\,\pi}\sqrt{\frac{k}{m}}$$

In which f_n[Hz] is the undamped natural frequency. This has to be converted to a damped natural frequency by the following formula:

$$f_d = f_n\sqrt{1-\zeta^2}$$

In which f_d [Hz] is the damped natural frequency and ζ [-] is the damping ratio of the mass spring damper model.

Damping ratio:

$$\zeta = \frac{c}{2\sqrt{km}}$$

In which c [$\frac{kg}{s}$] is the damping coefficient.

This relationship study was preformed with the following parameters:

- *Droplet diameters: 1, 2.5 and 5 mm*
- *Damping coefficient: 1* $\frac{kg}{s}$
- *Stiffness: k ~ EI · h, while the thickness h is assumed to be constant. Furthermore, it is assumed that the stiffness follows a trapezoidal distribution and the* $EI_{edge} = ¼ * EI_{centre}$

Figure 12: Stiffness distribution of piezo disk disdrometer. Note that EI_{max} is the stiffness at x = 20 mm.

Undamped natural frequency: $\frac{1}{2\,\pi}\sqrt{\frac{k}{m}} = \frac{1}{2\,\pi}\sqrt{\frac{k}{\mathrm{m}}}$

Damping ratio: $\frac{c}{2\sqrt{km}} = \frac{1}{2\sqrt{k \cdot \mathrm{m}}}$

Damped natural frequency: $f_n\sqrt{1-\zeta^2} = \frac{1}{2\,\pi} \cdot \sqrt{\frac{k}{\mathrm{m}}} \cdot \sqrt{1-(\frac{1}{2\sqrt{k \cdot \mathrm{m}}})^2}$

The resulting normalized damped natural frequency is shown in Figure 13.

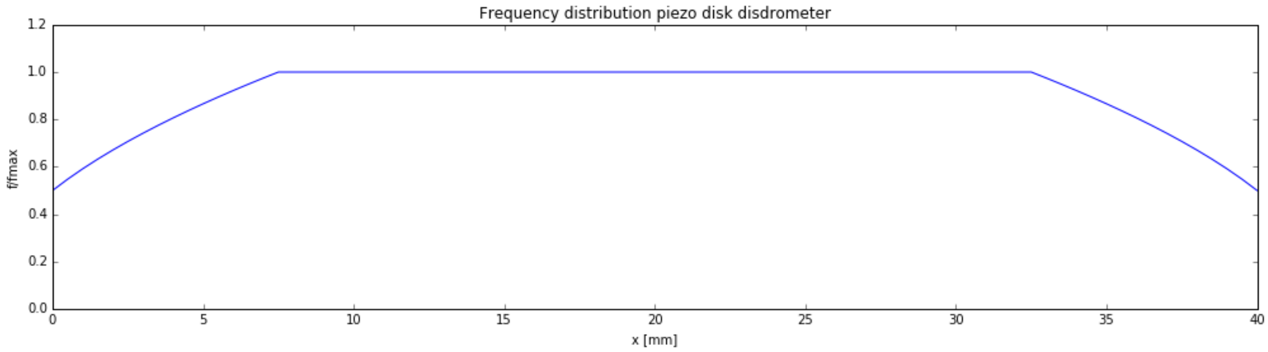

Figure 13: Frequency distribution piezo disk disdrometer. Note that f_{max} is the damped natural frequency at x = 20 mm.

This normalized frequency plot was very similar for different drop diameters. The figure below was constructed, to assess the relative errors for different drop sizes. This figure shows that the relative error is larger for smaller droplets.

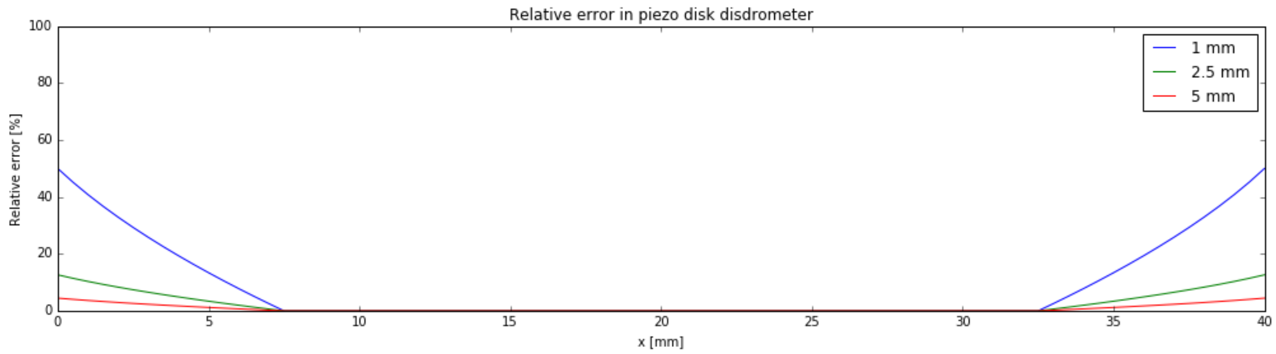

Figure 14: Relative error made by the piezo disk of the disdrometer

Python scripts

Script 1: Script for tipping bucket – Delft-disdrometer comparison

```
import matplotlib.pyplot as plt
import numpy as np
import scipy.stats
import pandas as pd
from pandas import read_csv, DataFrame
from sklearn import datasets, linear_model
from scipy import stats

datayangon = read_csv('yangon11July5Sept.csv',',')
databago  = read_csv('bago11July21July.csv',',')
tipyangon = np.zeros(len(datayangon))
disyangon = np.zeros(len(datayangon))
tipbago  = np.zeros(len(databago))
disbago  = np.zeros(len(databago))
for q in range(len(datayangon)):
  tipyangon[q] = datayangon.tip[q]
  disyangon[q] = datayangon.dis[q]
for p in range(len(databago)):
  tipbago[p] = databago.tip[p]
  disbago[p] = databago.dis[p]

def statistics(y,j):
  MBE = (1.0/len(y))*np.sum((y-j))
  RMSE = np.sqrt((1.0/len(y))*np.sum((y-j)**2.0))
  Rsqrt = ((scipy.stats.linregress(y,j))[2])**2
  slope = stats.linregress(y,j)[0]
  intercept = stats.linregress(y,j)[1]
  return MBE,RMSE,Rsqrt,slope,intercept

print ('Yangon', statistics(tipyangon,disyangon))
print ('Bago', statistics(tipbago,disbago))

plt.figure(figsize=[5,5])
plt.plot(tipyangon, disyangon,'d', label='measurements')
plt.xlim(0,10)
plt.ylim(0,10)
x = np.linspace(0,10,100)
y = np.linspace(0,10,100)
plt.plot(x,y, color='red',label='line y=x')
plt.xlabel('Tipping bucket measured intensity [mm/5min]')
plt.ylabel('Delft-disdrometer measured intensity [mm/5min]')
plt.title('Tipping bucket vs Delft-disdrometer (Yangon)')
plt.legend(loc=1)

plt.savefig("Yangon comparison.png")
```

Script 2: Script for Thies LPM data analysis

```
import matplotlib.pyplot as plt
import numpy as np
import scipy.stats
import pandas as pd
from pandas import read_csv, DataFrame
from scipy import stats
from scipy.interpolate import spline
import filemapper as fm


wf1 = read_csv('KNMI1gem.csv',',')
wf5 = read_csv('KNMI5gem.csv',',')
wf10 = read_csv('KNMI10gem.csv',',')
wf50 = read_csv('KNMI50gem.csv',',')


bins = [0,0.125,0.250,0.375,0.500,0.750,1.0,1.25,1.5,1.75,2,2.5,3,3.5,4,4.5,5]
binav= []
for q in range(len(bins)-1):
  binav.append(((bins[q+1]-bins[q])*0.5)+bins[q])


MainMatrix = []
for i in range(len(wf5)):
  row = [0]
  for column in wf5:
    row.append(wf5[column][i])
  MainMatrix.append(row)

output = []
xbin = [0]
for i in range(len(MainMatrix[0])):
  for j in range(MainMatrix[0][i]):
    xbin.append(binav[i])
entries, binedg, pat=plt.hist(xbin,bins,normed='True','r',label='LPM data')
output.append(entries)


plt.figure()
for i in range(len(output)):
  plt.plot(binav,output[i],'d')


lnspc = np.linspace(0.,5.,len(xbin))
ag,bg,cg = stats.gamma.fit(xbin)
pdf_gamma = stats.gamma.pdf(lnspc, ag, bg,cg)
pdfdatagamma = stats.gamma.pdf(binav,ag, bg,cg)

a,b,c,d = stats.gengamma.fit(xbin)
pdf_gengamma = stats.gengamma.pdf(lnspc, a, b, c,d)
pdfdatagengamma = stats.gengamma.pdf(binav, a, b, c,d)


plt.plot(lnspc, pdf_gamma, ':', color='k', label="Gamma fit LPM")
plt.plot(lnspc, pdf_gengamma, '--', color='g', label="Gengamma fit LPM")
plt.xlim(0.,5.)
plt.ylim(0,2.)
plt.xlabel('Raindrop diameter, D [mm]')
plt.ylabel('Probability density [mm^-1]')
```

```
plt.title('Probability density function - 01/02/2015')
plt.legend(loc='best')
plt.savefig("5.png")

def statistics(y,j):
  MAE = (1.0/len(y))*np.sum(np.abs(y-j))
  RMSE = np.sqrt((1.0/len(y))*np.sum((y-j)**2.0))
  Rsqrt = ((scipy.stats.linregress(y,j))[2])**2
  slope = stats.linregress(y,j)[0]
  intercept = stats.linregress(y,j)[1]
  return MAE,RMSE,Rsqrt,slope,intercept

print('gengamma fit',statistics(output,pdfdatagengamma))
print('gamma fit',statistics(output,pdfdatagamma))
```

Script 3: Script for the probability density curve of Thies LPM, with 90 % confidence bands

```
import matplotlib.pyplot as plt
import numpy as np
import scipy.stats
import pandas as pd
from pandas import read_csv, DataFrame
from scipy import stats
from matplotlib import style
from scipy.interpolate import spline
import filemapper as fm


wf1 = read_csv('Intensity1Data.csv',',')
wf5 = read_csv('Intensity5Data.csv',',')
wf10 = read_csv('Intensity10Data.csv',',')
wf50 = read_csv('Intensity50Data.csv',',')


def mean_confidence_interval(data, confidence=0.80):
  a = 1.0*np.array(data)
  n = len(a)
  m, se = np.mean(a), scipy.stats.sem(a)
  h = se * scipy.stats.t._ppf((1+confidence)/2., n-1)
  return m, m-h, m+h


MainMatrix = []
for i in range(len(wf1)):
  row = [0]
  for column in wf1:
    row.append(wf1[column][i])
  MainMatrix.append(row)


bins = [0,0.125,0.250,0.375,0.500,0.750,1.0,1.25,1.5,1.75,2,2.5,3,3.5,4,4.5,5]
binav= [0]
for q in range(len(bins)-1):
  binav.append(((bins[q+1]-bins[q])*0.5)+bins[q])


MainMatrixGenGamma = []


for p in range(len(MainMatrix)):
  x = MainMatrix[p]

  xbin = [0]
  for i in range(len(binav)-1):
    for j in range(x[i]):
      xbin.append(binav[i])

  lnspc = np.linspace(0.,5.,len(xbin))
  a,b,c,d = stats.gengamma.fit(xbin)
  pdf_gengamma = stats.gengamma.pdf(lnspc, a, b, c,d)
  pdfdatagengamma = stats.gengamma.pdf(binav, a, b, c,d)

  MainMatrixGenGamma.append(pdfdatagengamma)
```

```
MainBands = []

for k in range(len(binav)):
  data = []
  for x in range(len(MainMatrixGenGamma)):
    data.append(MainMatrixGenGamma[x][k])
  bands = mean_confidence_interval(data)
  MainBands.append(bands)


MeanBand = [0]
LowBand = [0]
HighBand = [0]
for t in range(1,len(binav)):
  MeanBand.append(MainBands[t][0])
  LowBand.append(MainBands[t][1])
  HighBand.append(MainBands[t][2])


lin = np.linspace(0.,5.,len(xbin))
MeanBandSM = spline(binav,MeanBand,lin)
LowBandSM = spline(binav,LowBand,lin)
HighBandSM = spline(binav,HighBand,lin)


plt.plot(lin,MeanBandSM,color='blue',label="1 mm/h, Rsqrt =0.83")
#plt.plot(lin,LowBandSM,color='blue',linestyle='--')
#plt.plot(lin,HighBandSM,color='blue',linestyle='--')
plt.fill_between(lin,MeanBandSM,LowBandSM,
where = MeanBandSM > LowBandSM,facecolor='blue',alpha=0.5 )
plt.fill_between(lin,MeanBandSM,HighBandSM,
where = MeanBandSM < HighBandSM,facecolor='blue',alpha=0.5)


plt.legend(loc='best')
plt.ylim(0,1.5)
plt.xlabel('Raindrop diameter [mm]')
plt.ylabel('Probability density [mm^-1]')
plt.title('Probability density curve with 90 % confidence bands')


plt.savefig("KNMI_plotlineband.png")
```

Script 4: Script for Delft-disdrometer data analysis

```
import matplotlib.pyplot as plt
import numpy as np
import scipy.stats
import pandas as pd
from pandas import read_csv, DataFrame
from scipy import stats
from scipy.interpolate import spline
import filemapper as fm

binavp = read_csv('binav.csv',',')

data = read_csv('Measurementpoints.csv',',')

def statistics(y,u):
  MAE = (1.0/len(y))*np.sum(np.abs(y-u))
  RMSE = np.sqrt((1.0/len(y))*np.sum((y-u)**2.0))
  Rsqrt = ((scipy.stats.linregress(y,u))[2])**2
  slope = stats.linregress(y,u)[0]
  intercept = stats.linregress(y,u)[1]
  return MAE,RMSE,Rsqrt,intercept,slope

lnspc = np.linspace(0,5,100)
a = np.linspace(0.6,1.0,10)
b = np.linspace(1.8,2.2,10)
c,d = 0,1
#pdf_gengamma = []
#pointsgengam = []

for i in range(len(a)):
  for j in range(len(b)):
    pdf_gengamma = stats.gengamma.pdf(lnspc, a[i], b[j], c, d)
    pointsgengam = stats.gengamma.pdf(binavp, a[i], b[j], c, d)
    #print (statistics(data, pointsgengam)[2])
    if statistics(data, pointsgengam)[2] > 0.993:
      #print(statistics(data, pointsgengam)[2])
      #print(a[i],b[j])
      #plt.figure()
      plt.plot(binavp,data,'rd')
      plt.plot(lnspc,pdf_gengamma)
      plt.tick_params(
      axis='x',
      labelbottom='off')
    plt.savefig("Yangon1gg.png")
```

Script 5: Script for DSD's of the eight study locations in Myanmar

```
import matplotlib.pyplot as plt
import numpy as np
import scipy.stats
import h5py as h5py
import math
import pandas as pd
from pandas import read_csv, DataFrame
from numpy.random import normal
from scipy.stats import norm
from scipy.integrate import simps
from scipy import stats
from os import listdir
from os.path import isfile, join
import filemapper as fm


#Import Data
all_files = fm.load('Data')


##Set up figure 1
fig1=plt.figure()
fig1,ax = plt.subplots(nrows=1,ncols=3,sharey=True,sharex=True,figsize=(10,5))
ax[0].set_title('Bin 0-1 mm/h',fontsize='medium')
ax[1].set_title('Bin 1-5 mm/h',fontsize='medium')
ax[2].set_title('Bin > 5 mm/h',fontsize='medium')


for n in range(len(all_files)):
  data = read_csv(all_files[n],',')
  D0 = data.Dparam
  N0 = data.Nparam
  R = data.Intensity
  Dist = data.Distance

  #Prepare for calc
  mu = 3.0
  D = np.linspace(0,5,100)
  MainN = np.zeros([100,len(N0)])
  BinBoundary1 = 1
  BinBoundary2 = 5

  #Count number of measurements inside each bin
  a = 0
  b = 0
  c = 0
  for t in range(len(N0)):
    if R[t]<=BinBoundary1:
      a += 1
    if R[t]>BinBoundary1 and R[a]<= BinBoundary2:
      b += 1
    if R[t]> BinBoundary2:
      c +=1

  #Main Loop, calc + plot
  MainN1 = np.zeros([100])
  MainN2 = np.zeros([100])
  MainN3 = np.zeros([100])
```

```
  for k in range(len(N0)):
    Nparam = N0[k]
    Dparam = D0[k]
    def DSD(Nparam,Dparam):
      N = []
      for i in range(100):
        N.append(Nparam*D[i]**mu*np.e**(-((3.67+mu)*D[i])/Dparam))
      return N

    int=simps(DSD(Nparam,Dparam),D)

    Nnorm = np.zeros([100])

    for j in range(100):
      Nnorm[j]=((Nparam*D[j]**mu*np.e**(-((3.67+mu)*D[j])/Dparam))/int)

    #Plot1
    if R[k] <= BinBoundary1:
      ax[0].plot(D,Nnorm,color='blue',linewidth=1, alpha=0.1)
    if R[k] > BinBoundary1 and R[k] <= BinBoundary2:
      ax[1].plot(D,Nnorm,'g',linewidth=1, alpha=.1)
    if R[k] > BinBoundary2:
      ax[2].plot(D,Nnorm,color='red',linewidth=1, alpha=.1)

    if R[k] <= BinBoundary1:
      MainN1 = MainN1 + Nnorm/a
    if R[k] > BinBoundary1 and R[k] <= BinBoundary2:
      MainN2 = MainN2 + Nnorm/b
    if R[k] > BinBoundary2:
      MainN3 = MainN3 + Nnorm/c

  meanvalues1 = []
  meanvalues2 = []
  meanvalues3 = []
  for m in range(100):
    meanvalues1.append(np.mean(MainN1[m]))
    meanvalues2.append(np.mean(MainN2[m]))
    meanvalues3.append(np.mean(MainN3[m]))


#Plot and save
for ax in fig1.get_axes():
  ax.set_xlabel("Raindrop diameter, D [mm]")
for ax in fig1.get_axes():
  ax.set_ylabel("Probability density [mm^-1]")
fig1.tight_layout()
plt.suptitle("Dropsize Distributions",size=12)
fig1.subplots_adjust(top=0.88)

plt.savefig("AllCities.png")
```

Script 6: Script for box plots of measured rain intensities by GPM

```
import matplotlib.pyplot as plt
import numpy as np
import h5py as h5py
import pandas as pd
from pandas import read_csv, DataFrame
import filemapper as fm


#Import Data
all_files = fm.load('Data')


R = []
for n in range(len(all_files)):
  data = read_csv(all_files[n],',')
  R.append(data.Intensity)


plt.boxplot(R)

labels = ( 'MOH', 'NAP', 'DAW', 'YAN', 'NAT', 'KEN', 'MAN','LEN')

plt.xticks(range(1,9),labels)
plt.ylim(0,20)
plt.xlabel('Study location')
plt.ylabel('Rainintensity [mm/h]')
plt.title('Boxplots of measured rain intensities by GPM')


plt.savefig("Boxplot intensities.png")
```

Script 7: Script for comparing the data results of Thies LPM and GPM

```
import matplotlib.pyplot as plt
import numpy as np
import scipy.stats
from numpy.random import normal
from scipy.stats import norm
from scipy.integrate import simps
from scipy import stats


#LPM Bins
bins = [0,0.125,0.250,0.375,0.500,0.750,1.0,1.25,1.5,1.75,2,2.5,3,3.5,4,4.5,5]
binav= []
for q in range(len(bins)-1):
  binav.append(((bins[q+1]-bins[q])*0.5)+bins[q])

x = [0,6,12,17,37,15,12,10,8,5,3,3,0,1,0,0]

xbin = [0]
for i in range(len(binav)):
  for j in range(x[i]):
    xbin.append(binav[i])

entries, bin_edges, patches = plt.hist(xbin,bins,'r',normed='True',label="LPM")


#LPM fit
data = []
for i in range(len(entries)):
  data.append(entries[i])

lnspc = np.linspace(0.,5.,len(xbin))
ag,bg,cg = stats.gamma.fit(xbin)
pdf_gamma = stats.gamma.pdf(lnspc, ag, bg,cg)
pdfdatagamma = stats.gamma.pdf(binav,ag, bg,cg)

a,b,c,d = stats.gengamma.fit(xbin)
pdf_gengamma = stats.gengamma.pdf(lnspc, a, b, c,d)
pdfdatagengamma = stats.gengamma.pdf(binav, a, b, c,d)


#GPM
Nnul = 33.81
Dnul = 1.06
mu = 3.0
D = np.linspace(0,5,len(xbin))

def DSD(Nnul,Dnul):
  N = []
  for i in range(len(xbin)):
    N.append(Nnul*D[i]**mu*np.e**(-((3.67+mu)*D[i])/Dnul))
  return N
def DSD2(Nnul,Dnul):
  N = []
  for i in range(len(binav)):
    N.append(Nnul*binav[i]**mu*np.e**(-((3.67+mu)*binav[i])/Dnul))
  return N
integrate=simps(DSD(Nnul,Dnul),D)
```

```
int2=simps(DSD2(Nnul,Dnul),binav)
Nnorm = []
for j in range(len(xbin)):
  Nnorm.append((Nnul*D[j]**mu*np.e**(-((3.67+mu)*D[j])/Dnul))/integrate)
GPMdata = np.zeros(len(data))
for k in range(len(binav)):
  GPMdata[k] = (Nnul*binav[k]**mu*np.e**(-((3.67+mu)*binav[k])/Dnul))/int2


#plot
plt.plot(D,Nnorm,color='k',label="GPM")
plt.plot(lnspc, pdf_gamma,':',color='k', label="Gamma fit LPM")
plt.plot(lnspc, pdf_gengamma,'--',color='g', label="Gengamma fit LPM")

plt.xlim(0,3.5)
plt.ylim(0,2.5)

plt.xlabel('Raindrop diameter, D [mm]')
plt.ylabel('Probability density [mm^-1]')
plt.title('Probability density function - 01/02/2015')
plt.legend(loc=1)

plt.savefig("1-2-15.png")


#stats
def statistics(y,j):
  MAE = (1.0/len(y))*np.sum(np.abs(y-j))
  RMSE = np.sqrt((1.0/len(y))*np.sum((y-j)**2.0))
  Rsqrt = ((scipy.stats.linregress(y,j))[2])**2
  slope = stats.linregress(y,j)[0]
  intercept = stats.linregress(y,j)[1]
  return MAE,RMSE,Rsqrt,intercept,slope

print('gengamma fit',statistics(data,pdfdatagengamma))
print('gamma fit',statistics(data,pdfdatagamma))
print('GPM',statistics(data,GPMdata))
```