



# At-site and regional frequency analysis of extreme precipitation from radar-based estimates

Edouard Goudenhoofdt<sup>1</sup>, Laurent Delobbe<sup>1</sup>, and Patrick Willems<sup>2</sup>

<sup>1</sup>The Royal Meteorological Institute of Belgium, Brussels, Belgium

<sup>2</sup>Department of Civil Engineering - Hydraulics Division, University of Leuven, Leuven, Belgium

*Correspondence to:* Edouard Goudenhoofdt[edouard.goudenhoofdt@meteo.be]

**Abstract.** In Belgium, only rain gauge time-series have been used so far to study extreme precipitation at a given location. In this paper, the potential of a 12-year quantitative precipitation estimation (QPE) from a single weather radar is evaluated. For the period 2005-2016, independent sliding 1 h and 24 h rainfall extremes from automatic rain gauges and collocated radar estimates are compared. The extremes are fitted to the exponential distribution using regression in QQ-plots with a threshold rank which minimises the mean squared error. A basic radar product used as reference exhibits unrealistic high extremes and is not suitable for extreme value analysis. For 24 h rainfall extremes, which occur partly in winter, the radar-based QPE needs a bias correction. A few missing events are caused by the wind drift of convective cells and strong radar signal attenuation. Differences between radar and gauge values are caused by spatial and temporal sampling, gauge rainfall underestimations and radar errors due to the relation between reflectivity and rain rate. Nonetheless the fit to the QPE data is within the confidence interval of the gauge fit, which remains large due to the short study period. A regional frequency analysis is performed on radar data within 20 km of the locations of 4 rain gauges with records from 1965 to 2008. Assuming that the extremes are correlated within the region, the fit to the two closest rain gauge data is within the confidence interval of the radar fit, which is small due to the sample size. In Brussels, the extremes on the period 1965-2008 from a rain gauge are significantly lower than the extremes from an automatic gauge and the radar on the period 2005-2016. For 1 h duration, the location parameter varies slightly with topography and the scale parameter exhibits some variations from region to region. The radar-based extreme value analysis can be extended to other durations.

## 1 Introduction

Very localised precipitation extremes can have a very strong impact on human activities especially in urban areas. For flood management applications (e.g. sewer system design) it is needed to know the expected maximum rainfall corresponding to a given return period. Based on the extreme value theory, a branch of statistics, several methods to fit a distribution to precipitation extremes have been developed in the literature. For infrastructure design application, one is more interested in longer return periods. Therefore a fitting method focusing on the tail of the distribution should be preferred (Willems et al., 2007). To reduce the uncertainty associated with the limited number of data at a single site, regional frequency analysis (RFA) methods have been proposed (Svensson and Jones, 2010).



There are numerous studies of RFA for rainfall extremes based on rain gauge datasets. The rain gauge network can perfectly capture rainfall extremes for widespread situations. However, they are unable to catch all rainfall extremes caused by convective storms, which often exhibit strong spatial variations over short distances. The use of high resolution gridded precipitation datasets to study rainfall extremes is still in its infancy. This can be explained by their unavailability, their processing requirements and their limited quality. Currently, the best potential is provided by radar-based quantitative precipitation estimation (QPE) products. With such gridded data, one could characterise sub-daily precipitation extremes on relatively short periods.

In a pioneer work, Overeem et al. (2009) showed that a 11-year radar data set is suitable to derive depth-duration-frequency (DDF) curves for the Netherlands. Based on a unique 23-year radar data set in Israel, Marra and Morin (2015) found that the DDF curves were generally overestimated but 60 % of them lay within the raingauge DDF confidence intervals. In Ontario (Canada), Paixao et al. (2015) demonstrate the potential to integrate radar (Digital Precipitation Array product) to raingauge analysis, especially to identify homogeneous regions of extreme rainfall. Saito and Matsuyama (2015) used a 26-year radar-gauge dataset (without RFA) to study the spatial variation of hourly precipitation in Japan. In a comprehensive study of the issues raised when using radar-based QPE to study precipitation extremes, Eldardiry et al. (2015) found for Louisiana (USA) that the relatively short period (13 years) explains the high uncertainty of the analysis, that the index flood method is recommended and that a systematic underestimation is associated with the radar products (its spatial resolution is  $4 \times 4 \text{ km}$ ). Haberlandt and Berndt (2016) found that the operational DWD product is only suitable for studies on longer durations after bias correction.

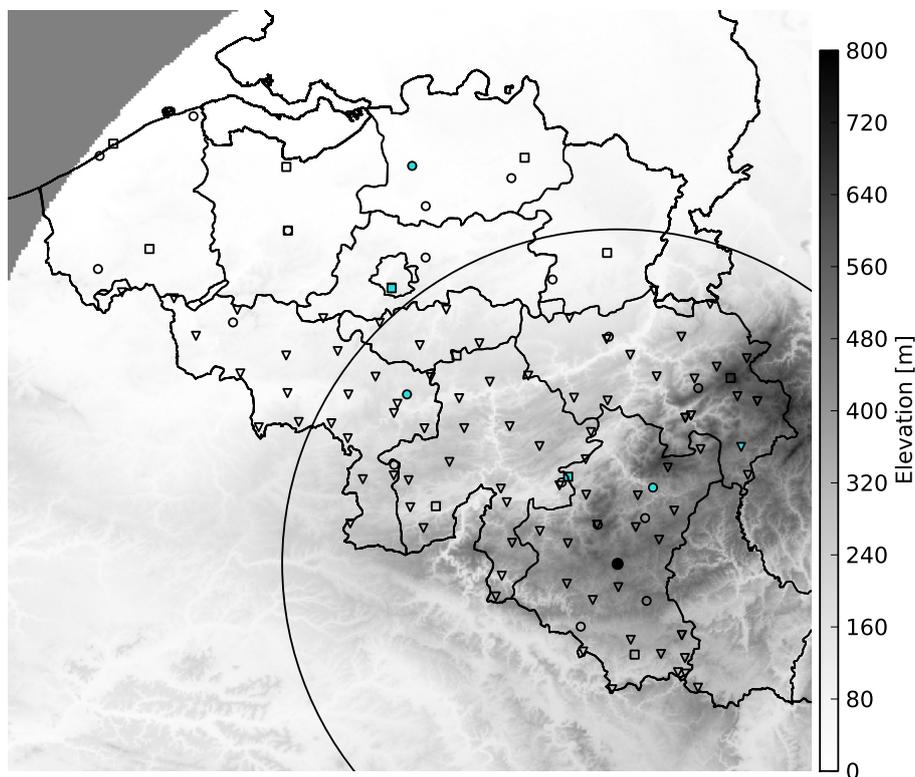
At the Royal Meteorological Institute of Belgium (RMIB), a high-resolution radar-based surface precipitation estimation is generated using all available observations. This dataset is used for various applications such as case studies and model verification. The methodology to derive the estimation from volumetric radar reflectivity data has been verified for the period 2005-2014 against an independent raingauge network (Goudenhoofdt and Delobbe, 2016). In this study, we want to demonstrate the potential of this radar-based QPE to derive point rainfall statistics. Radar-based extreme statistics of 1 h precipitation accumulation are compared with the same statistics but derived from the observations by a high-temporal resolution (10 min) rain gauge network. The radar-based extremes statistics of 24 h precipitation are compared with rain gauge data from another more dense network (hourly data). A regional frequency analysis is performed and compared to the results obtained with a network of 45-year high resolution (10 min) records. Finally, the regional approach is applied at each radar pixel on the whole of Belgium to study the spatial variations of the precipitation extremes.

## 2 Precipitation data

### 2.1 Raingauge measurements

Over the years, Belgium (Fig. 1) has been covered by several raingauge networks for different purposes.

Since the end of the 19th century, RMIB maintains a network (CLIM) of non-recording rain gauges from which precipitation measurements are taken at 8 am LT. The data are carefully controlled and used for climate applications (Journée et al., 2015).



**Figure 1.** Elevation map centered on Belgium with the Wideumont radar (black dot) covering 240 km range (circle is 120 km) with AWS (square), SPW (triangle) and BUL (circle) rain gauge networks. The gauge locations selected in this paper are in cyan. Country borders with France, Luxembourg, Germany and the Netherlands are also displayed.

A Hellmann-Fuess pluviograph has been in operation in Uccle (RMIB) from 1898 to 2008 and has enabled the compilation of a continuous time series of 10 min precipitation (Demarée, Gaston, 2003). The 10 min precipitation values had to be manually extracted from line graphs on papers. Starting from the fifties, additional rain gauges were installed to constitute a network (BUL) for hydrological research. Since the rain gauges underestimate the rainfall by 5-10% due to its mechanism, its records  
5 have been calibrated using a collocated gauge from the CLIM network.

For weather forecast purposes, the RMIB maintains a network of automatic weather stations (AWS) in Belgium. These stations provide precipitation measurements at very high temporal resolution ; 10 min accumulations are available from the database. The tipping-bucket gauges are progressively replaced by weighted gauges (the first one was installed in Uccle on 10 February 2009). The data are available since 2002-2004 and have been quality controlled.

10 The hydrological service of the Walloon Region (SPW) maintains a dense network of hourly (every 5 min since 2012) precipitation measurements. The tipping bucket gauges are progressively replaced by weighting gauges since 2015. The data have been quality controlled by RMIB since April 2004.



It is important to know the limitations of the respective rain gauges in case of extreme precipitation. It is known (Duchon and Biddle, 2010; Nystuen, 1999) that tipping buckets underestimate high rainfall rates. The use of weighting gauges for extreme precipitation is discussed in Colli et al. (2012). Every 10 mm, the pluviograph has to be emptied which results in an underestimation in case of extreme precipitation. The calibration of the pluviograph is probably not sufficient for sub-daily extremes. Finally, the quality controls, albeit conscientious, can never be considered as perfect.

## 2.2 Radar estimation

The quantitative precipitation estimation (QPE) available on a 1 km grid every 5 min is made using an elaborated processing chain from the radar volumetric reflectivity measurements. The quality of the volume is controlled using static clutter and beam blockage maps and clutter identification based on vertical gradients, horizontal texture and satellite observations. A maximum threshold for reflectivity is set to 55 dBZ to mitigate higher reflectivity values due to hail. The rainfall rate estimates are obtained using stratiform-convective classification, a 40 min averaged vertical profile of reflectivity (VPR), a bright band identification and a specific transformation to rain rates for the two precipitation regimes. The detailed procedure is described in Goudenhoofd and Delobbe (2016). As a reference for the QPE product, the CAP product is defined as the interpolation at 800 m above the radar level. It makes use of a standard Z R relationship, which comes from the hypothesis that the drop size distribution follows the distribution of Marshall-Palmer, as discussed in (Uijlenhoet and Pomeroy, 2001). Consecutive rainrate estimates are integrated to obtain 10 min accumulations (5 min gaps are tolerated) to match the lowest resolution of the rain gauge data. Hourly accumulations are combined with the SPW gauges using a mean field bias correction (MFB). A more complex merging method (i.e. external drift Kriging) was tested but found to be unstable for some time moments.

It is important to mention the limitations of the radar products in case of extreme precipitation. The most important impact of the QPE processing on extreme values is the 55 dBZ reflectivity threshold used to mitigate hail. This corresponds to a maximum rainfall rate of 80 mm/hour and hence 13.33 mm/10 min. Higher values of about 100 mm/hour are possible when the standard Z R relationship is used for stratiform areas. This can only happen close to the radar where convective precipitation can not be identified. This thresholding underestimates very rare (if any) rainfall rate exceeding the threshold. Even after thresholding an overestimation due to hail is possible. The second most important error is related to signal attenuation especially in case of well organised convective systems. This is why extremes might be underestimated the further the distance from the radar. In addition, the increasing radar sample volume will give lower extreme values.

## 2.3 Comparison framework

In this study, we will only consider validated rain gauge data. Given that the SPW network is used for merging, the radar dataset for 2005-2016 is used. To perform a direct comparison, the gauge data of AWS and SPW for the same period are used. For comparison against the reference BUL network, the gauge data for the period 1965-2010 are used. The timeseries of the BUL and CLIM networks have been tested for homogeneity by Van de Vyver (2012) and a selection of useful stations has been made. Gellens (2000) and Vannitsem and Naveau (2007) found that the vast majority of the CLIM and BUL time series are



stationary for summer precipitation. However, the existence of a multi-decadal oscillation in precipitation extremes has been found in the Uccle time series (Ntegeka and Willems, 2008).

The 10 min precipitation accumulation from the gauge networks (AWS, BUL) and radar products (QPE, CAP) are summed to obtain sliding 1 h precipitation accumulations. Such duration is associated with convective storms, which can only be properly  
5 seen on radar images. The hourly bias obtained by the MFB method could be applied to the 10 min accumulations. However, it will not be used due to the possible risk of representativity errors related to convective storms.

The hourly precipitation from the radar products (CAP, QPE, MFB) are summed to obtain sliding 24 h precipitation accumulations. Such duration is mainly associated with widespread precipitation for which the benefit of merging methods is clear.

10 It should be noted that using the lowest available duration for each network would result in an underestimation of the extremes due to the discrete time sampling. Additionally, random errors and time sampling difference can be compensated by performing the sum. For both the radar and the gauge, no missing data is tolerated in the sum to avoid underestimation. Furthermore, only timestamps with both radar and gauge data are kept.

Due to the amount of stations, it is not possible to analyse in details the results at each station. Therefore a few stations  
15 are picked at different distances from the radar (see Tab. 1 and Fig. 1). The Uccle station is chosen because it is included in the 3 networks, which makes intercomparison possible. The availability of the 1 h accumulation data is about 95 % for the radar products and close to 100 % for the AWS gauges. The radar availability of the 24h accumulation is lower than the 1 h accumulation because a significant part of the intervals without data are short. The availability of the SPW gauges is around 90 % but this is mainly due to the removal of snow events, when no extreme rainfall is expected. The availability of the BUL  
20 stations for the period 1965-2010 is highest at Uccle with 96.3 %, then about 86 % at Deurne and Gosselies. The station of Nadrin has only 60 % of availability (for the period 1965-2010) because it was started in 1978.

### 3 At-site frequency analysis

#### 3.1 Methodology

Extremes are often extracted using block maxima of one year but it is not recommended for small sample size. The peak-over-  
25 threshold (POT) method, where values exceeding a given threshold are kept, is preferred here. It has been shown by Pickands III (1975) that the extreme values converge asymptotically to a generalized Pareto Distribution (GPD) :

$$F_{(\xi, \mu, \sigma)}(x) = \begin{cases} 1 - \left(1 + \frac{\xi(x-\mu)}{\sigma}\right)^{-1/\xi} & \text{for } \xi \neq 0, \\ 1 - \exp\left(-\frac{x-\mu}{\sigma}\right) & \text{for } \xi = 0. \end{cases} \quad (1)$$

with  $\xi$ ,  $\mu$  and  $\sigma$  commonly defined as the shape, location and scale parameters. The special case when the shape parameter is equal to zero is defined as the Exponential distribution (EXP).



The choice of the threshold has an important impact on the estimation of the distribution parameters. When the number of selected values increases, the variance naturally decreases but the bias increases (due to the deviation from the theoretical distribution). It is more practical to use a threshold rank instead of a threshold value to control the sample size.

To apply the extreme value theory, the quantiles have to be independent (i.e. not in a cluster). The 1 h extremes are caused by convective storms, which have been analysed based on radar volume data in Goudenhoofd and Delobbe (2013). Mesoscale convective systems can last more than one day, but due to their motion, they affect a particular region for several hours only. Therefore an interval of 12 h, as in Ntegeka and Willems (2008), is chosen to consider that two values are independent. In practice, the maxima in a sliding window of 24 h are selected. For 24-h durations, we use an interval of 3 days which corresponds to the synoptic scale.

The maximum likelihood is the most widely used fitting method but for small samples it can lead to unrealistic parameter estimates. This problem is partially addressed with the generalised MLE proposed by Martins and Stedinger (2000). The popular method of L-moments is preferred in case of small samples (Overeem et al., 2009). However, all those methods do not focus on the tail of the distribution, which is the most relevant for risk analysis. Therefore in this study we use a method based on regression in Q-Q plots (QQR) proposed by Willems et al. (2007). In this method the optimal threshold rank is found by minimization of the mean squared error (MSE) of the calibration. With our datasets, this rank is chosen between 18 and 30 considering the uncertainties and the relatively short period, respectively. The EXP distribution is preferred since estimating the shape parameter is very uncertain due to the short period. Moreover, earlier research has shown that the shape parameter for rainfall extreme in Belgium does not significantly differ from zero (Willems, 2000).

Confidence intervals for the scale parameter are computed using a parametric bootstrap technique. Practically the fitting is reproduced 1000 times on randomly generated values up to the corresponding optimal rank. The 10 and 90 percentiles of the scales obtained are taken as confidence intervals.

### 3.2 Comparison with AWS gauges - 1h

The 10 highest extremes are compared between the radar and the gauges (table 2). For problematic events, the underlying precipitation patterns are analysed using the radar images. This also allows to identify the weaknesses of the gauge and radar datasets.

The maximum at Humain has been observed by both the radar and the gauge on 7 June 2016. This relatively high value can be due to randomness and the short period of records. But it is also possible that the other quantiles are underestimated (the maximum was recorded by the new weighted gauge). There is generally a good match between the radar and the gauge quantiles except for the following events :

- event 2 : the radar underestimates globally
- event 6 : the gauge is located at the boundary of a convective cell (most probably with hail)
- event 11 : the radar signal is strongly attenuated by a mesoscale convective system.



- event 13 : there was probably snow in the gauge
- event 14 : the gauge is located at the boundary of a convective cell.

The second highest quantile at Uccle has been observed by both the radar and the gauge on the 7th of October 2009. There is generally a good match between the two datasets. A few events are problematic :

- 5 – event 1,4 : the gauge is at the boundary of a cell (most probably with hail)
  - event 9 : there is a stationary storm underestimated by the gauge
  - event 10 : the gauge is at the boundary of a cell and the radar is attenuated (same as event 2 in Humain)
  - event 11 : the radar signal is strongly attenuated (same as event 11 in Humain)
  - event 13 : the radar is attenuated
- 10 Missing events at cell boundaries can be explained by the fact that precipitation, which is estimated at a given height by the radar, is subject to wind drift. This effect increases with the distance to the radar.

Figure 2 shows the results of the extreme value analysis for 1 h precipitation accumulation. Numerical values can be found in table 3. The percentage of independent peaks (amongst peaks exceeding the threshold) is around 20 % for both the radar and the gauges at the two locations. This is what we expect from 1 h accumulation available every 10 min.

- 15 The empirical quantiles of the QPE product are systematically slightly lower than those for the AWS gauges. This may be expected as we compare point rainfall observations with rainfall averaged on a 1 km square. However, the underestimation of very high rainfall rate by tipping-bucket gauges can compensate for this effect. One also notes small groups of similar values for both the radar and the gauge. This can be explained by the hail thresholding and the rainfall rate limit, respectively.

- 20 The fit of the EXP distribution is relatively good for the two locations with a relatively low MSE (not shown). Except for the AWS in Uccle, the extremes tend to be heavy tailed but this can be at least partially explained by the observation biases described above. The scale parameter tends to be higher for the gauge data than the radar data. In general, the uncertainty for the scale parameter remains high and this results in wide confidence intervals for higher return periods.

When using the CAP product, the higher quantiles are overestimated especially for Uccle. This can be mainly attributed to the effect of hail. This results in an overestimation of the scale parameter.

### 25 3.3 Comparison with SPW gauges - 24h

The 10 highest extremes for the radar (MFB) and the gauge (SPW) can be seen in table 4. For Uccle, most extreme values occurred during summer and are therefore associated with convective storms. There is a good match between the gauge and the radar except for a few events:

- event 8, 11 : the gauge is at the boundary of a convective cell
- 30 – event 13 : strong radar attenuation by a mesoscale convective system



- event 14 : snow episode probably underestimated by the radar

For Saint-Vith, the extreme values occurred either in summer or in winter with therefore a mix of convective and widespread precipitation episodes. The match is very good except for the following events :

- event 2 : at the boundary of a cell (probably with hail)
- 5
- event 3 : might be a radar overestimation due to snow melting
  - event 13 : at the boundary of a cell

Figure 3 shows the results of the extreme value analysis for the 24 h precipitation accumulation. Numerical values can be found in table 5. The percentage of independent peaks (amongst peaks exceeding the threshold) is between 6 % and 9 % for the two locations and for all datasets. This is what we expect from 24 h accumulation available every hour.

- 10
- For Uccle there are not many differences between QPE and MFB because most events are associated with convective storms. Compared to the gauge quantiles, the radar quantiles are lower below 1-year and higher between 1-year and 5-year return periods. This can be attributed mainly to hail overestimation by the radar and gauge losses. It results in a higher scale for the radar, which is close the upper bound of the gauge confidence interval.

- 15
- For Saint-Vith, there is a clear effect of the bias correction (MFB) to remove the underestimation of the QPE product. As for Uccle, the radar quantiles are higher for return periods higher than 2 years but the effect is limited because less convective storms are involved. The final result is a good match of the two distributions for this station.

For Uccle, the CAP product overestimates the scale parameter and underestimates the location parameter due to hail and VPR errors, respectively. For Saint-Vith, the quantiles (not shown) are similar to QPE except for a very high unrealistic maximum.

## 4 Regional frequency analysis

### 20 4.1 Methodology

- One possibility to decrease the uncertainty of at-site analysis is to perform a regional frequency analysis (RFA). The RFA is characterised by the selection of the regions and the parameter estimation approach applied to each region (Buishand, 1991). The index flood approach, which consider that only the location parameter varies in the region, is very popular (Gellens, 2000; Rulfova et al., 2014). Uboldi et al. (2014) used a bootstrap technique to randomly select data from neighbouring locations
- 25
- with a probability depending on the distance and altitude variation to the target location. Different RFA approaches for radar datasets are tested in Eldardiry et al. (2015), who defines a region as a square window of 44 km size. Overeem et al. (2009) applied the index flood method for the whole of the Netherlands. In this study, we consider that the distribution parameters are the same within a radius of 20 km around the target location. This choice of a neighborhood provides a sufficiently large data sample.

- 30
- Even for 1 h duration, precipitation maxima exhibit spatial correlation (Vannitsem and Naveau, 2007). Reed et al. (1999) propose a technique to deal with spatially correlated data. Here we remove the dependent values by taking the maximum within



a certain distance. A time window of 12 h is used as in the at-site frequency analysis. A first distance of 10 km is tested, which corresponds to the maximum expected size of a convective cell (Goudenhoofdt and Delobbe, 2013). We also test the hypothesis that convection is always organised at the meso-scale and hence consider that all values are dependent within 50 km. In the text, we will refer to these two datasets by the names R10 and R50. The time span of the pooled dataset is reduced according to the percentage of spatially independent peaks, amongst peaks exceeding the threshold. The large number of peaks available from the radar data allows us to choose a higher threshold rank. This increase in sample size leads to a more reliable extreme value analysis, which is the final goal of this radar-based RFA. Applying the QQR method on the pooled dataset, one can safely choose a rank between 30 and 100.

## 4.2 Comparison with rain gauges

Figure 4 and 5 shows the results of the regional frequency analysis for 1 h precipitation accumulation at the 4 locations selected from the BUL network. The results of the at-site frequency analysis for the gauge and collocated radar pixels are showed as reference. Numerical values can be found in table 6. The percentage of temporally independent extremes for the gauge is close to 30 % for Deurne and Uccle while it is slightly above 20 % for the two others stations. This suggests that there are larger clusters which might be related to altitude. Above the threshold, the percentage of spatially independent extremes (50 km) ranges from 1.1 % (Uccle) to 2.6 % (Nadrin). The corresponding period for the pooled dataset is between 200 and 500 years. Using a decorrelation distance of 10 km results in twice more data suggesting that convection is often organised.

The radar images associated with each maximum of the radar-based RFA is analysed :

- Deurne and Uccle (28 July 2006) : several supercells on the whole of Belgium
- Gosselies (22 August 2011) : a squall line moving parallel to the flow
- Nadrin (26 July 2008) : a stationary convective cell

The highest extremes exhibit abrupt variations in the form of steps for both the gauge and radar. This could be explained by the siphonage of the gauge and hail thresholding, respectively. Since Nadrin is close to the radar, the standard Z-R relationship is used and allows for higher rain rates (i.e. 100 mm/hour).

The gauge extremes are relatively low at Deurne and Uccle compared to Nadrin and Gosselies. The radar extremes are lower at Deurne compared to the other stations. This can be at least partially attributed to the large sample volume at this range. The match between the gauge and the radar (R10 and R50) is good except at Uccle with much higher radar extremes. This can be partially attributed to hail but the similar 4 highest extremes suggest a gauge limitation. It is also striking that half of the 20 highest gauge extremes occurred during the period 1999-2008 (not shown). This positive trend for Uccle is possibly related to the urban heat island effect (Hamdi and Van de Vyver, 2011).

The uncertainty of the radar fit is low because of the larger sample size, due to which a higher rank can be chosen. Furthermore, the fit is less impacted by the potentially large errors of the highest extremes. The location parameter (corresponding to the threshold) increases for the successive products due to the increased sample size.



Except for the Uccle station, the scale parameter is the lowest for the QPE dataset due to the bias as a result of the small sample size. The scale parameter of the pooled radar datasets is slightly higher at Deurne and significantly higher in Uccle. For Gosselies and Nadrin, the R10 and BUL data have similar scales while it is slightly higher for the R50 data. The fit to the R10 and R50 data is within the uncertainty bound of the fit to the BUL data. For those two stations, the fit to the BUL data is even in the small uncertainty bound of the fit to the R50 data.

### 4.3 Spatial maps

In Belgium, Van de Vyver (2012) derived a spatial GEV model depending linearly on the altitude. Rulfova et al. (2014) found for 6-hourly precipitation in the Czech Republic that the assumption of a linear model might be too restrictive, especially for convective precipitation. Here we are able to use the radar data to study directly the spatial variations of the extremes. It also allows to verify our hypothesis that the distribution parameters do not vary on a 40 km scale.

In practice, we performed the regional frequency analysis at all pixel locations but using a radius of 10 km (with a decorrelation distance of 50 km). Besides the reduced computation cost, it improves the resolution of the maps. A few pixels having too much (50) peaks over the threshold (12 mm/hour) are considered as residual clutter and removed. To make the comparison easier, we choose a fixed threshold rank (60).

Figure 6 and 7 show respectively for Belgium the location (i.e. the 60th highest extreme) and scale parameters. There are no values beyond the 180 km range because the quality of the radar QPE tends to decrease. For the location parameter, there is some correlation with topography and the mean annual rainfall (Journée et al., 2015) but the variations are small. The scale parameter exhibits higher variations which are partially correlated with the location parameter. The small scale variability in the study area can be explained by uncertainties due to the sample size. One notes that the scale is very high around the Brussels region where the Uccle station is located. The highest values might be affected by overestimation due to hail. The circles of 10 km with very high values (e.g. at the German border) are probably an artefact caused by exceptional clutter.

These results suggest that considering constant extreme statistics over small regions is valid for the 1-hour duration.

## 5 Conclusions

The potential of a radar-based precipitation dataset to study extreme precipitation at a given location is evaluated. The quantitative precipitation estimate (QPE) is obtained by a careful processing of the volumetric reflectivity measurements from a single weather radar in Belgium. The radar dataset covers the period 2005–2016, has a resolution of 1 km, and is available every 5 minutes.

The first evaluation is based on a comparison of the extreme statistics between the radar dataset and two automatic rain gauge networks with 10 min and 1 h resolution, respectively. For each network, two locations are chosen to study sliding 1 h and 24 h extremes using the collocated radar estimation. A regression method in Q-Q plots is used to fit an exponential distribution to independent peaks. An optimal threshold rank is selected by minimising the MSE of the regression.



The 10 highest 1 h extremes occurred in summer and are well captured by both the radar and the gauge. A few exceptions are caused by wind drift or severe radar signal attenuation. There are some differences up to 30 % between the gauge and radar values which can be explained by sampling and estimation errors. In particular, tipping bucket gauges underestimate heavy rainfall rate and can be blocked by accumulated snow. The radar underestimates due to signal attenuation and overestimates in case of hail. Additional radar uncertainties come from time sampling and the Z-R relationship. Nonetheless, the fitting of the exponential distribution to the QPE product is within the large uncertainty bound of the AWS one.

For 24 h accumulation there is a mix of summer and winter events, with more of the latter for stations with higher altitude. There is a clear benefit of bias correction for the highest station, making the distribution fits similar for both stations. For both 1 h and 24 h accumulations, the basic radar product exhibits unrealistic high extremes, which results in an overestimated scale parameter. Such product is therefore not suitable for an extreme value analysis.

In the second evaluation a regional frequency analysis is applied to 1 h radar data at the location of 4 pluviographs with recordings from 1965 to 2010. Spatially independent extremes within a circle of 20 km are selected and fitted with a maximum threshold rank extended from 30 to 100 thanks to the increased sample size. There is a good agreement between the radar and the gauge for the two closest stations. The extremes are slightly higher when a decorrelation distance of 50 km is used instead of 10 km. In Uccle, the radar extremes and therefore the scale parameter are significantly higher. This can be attributed partially to radar overestimation due to hail and gauge underestimations, but the increasing urban heat island effect should not be ruled out. The decreasing tail of the radar extremes is at least partially caused by hail thresholding but a physical limit for the Belgium climate could play a role. Based on 1 h radar data, the location parameter remains relatively constant over Belgium with a slight effect of the topography (a similar result has been obtained by Van de Vyver (2012)). The scale parameter exhibits higher variations between regions of about 40 km size.

There is still some room to improve the quality of the radar and gauge datasets. The recently installed weighting gauges are able to cope with intense rainfall and snowfall. One will have to wait a few decades before it can produce reliable statistics. Radar calibration errors can be mitigated by computing a monthly bias using rain gauges. The attenuation can be solved easily by using other radars when available. An advection correction can be used for the time sampling error. Dual-polarization radars can potentially provide better estimation for high rainfall rates (Figueroa i Ventura and Tabary, 2013). However uncertainties related to relation between the radar measurements and the rainfall rate remain, especially in case of hail. In this study, all kind of precipitation including hail is considered. For some applications, it could be necessary to remove the precipitation associated with hail. Identification of hail at ground level is a challenging problem using radar and ground station networks (Lukach et al., 2017).

For each of the rain gauge networks, only a few stations have been selected and presented in this paper. The results from these stations are representative of the variability of the results obtained from the other stations.

Since the paper focuses on comparison against rain gauges, an at-site extreme value analysis has been conducted, assuming an EXP distribution. In future works one should consider the generalised Pareto distribution and perform the necessary bias correction related to the asymptotic behavior of the distribution. The extreme value theory was applied to the radar datasets by



removing the spatially dependent extremes in the region of analysis. The recent theory of spatial extremes can offer a more elegant approach to this problem (Buishand et al., 2008).

The radar-based regional frequency analysis can be extended to other durations to derive IDF curves. Note that the hypothesis of constant parameter over the region might not be valid for longer durations. In many applications in hydrology, it is the averaged rainfall over a given area which is relevant. A popular technique is to apply areal reduction factors to point-based statistics. The radar dataset can be used directly to derive areal rainfall statistics.

## 6 Code availability

The code used in this study is part of the RMIB radar library.

## 7 Data availability

10 The rain gauge and radar precipitation estimation data are archived at the RMIB.

*Acknowledgements.* The authors would like to thank the hydrological service of the Walloon region (DGO2) for providing their rain gauge data. The comments of Michel Journee and Hans Van de Vyver are highly appreciated.

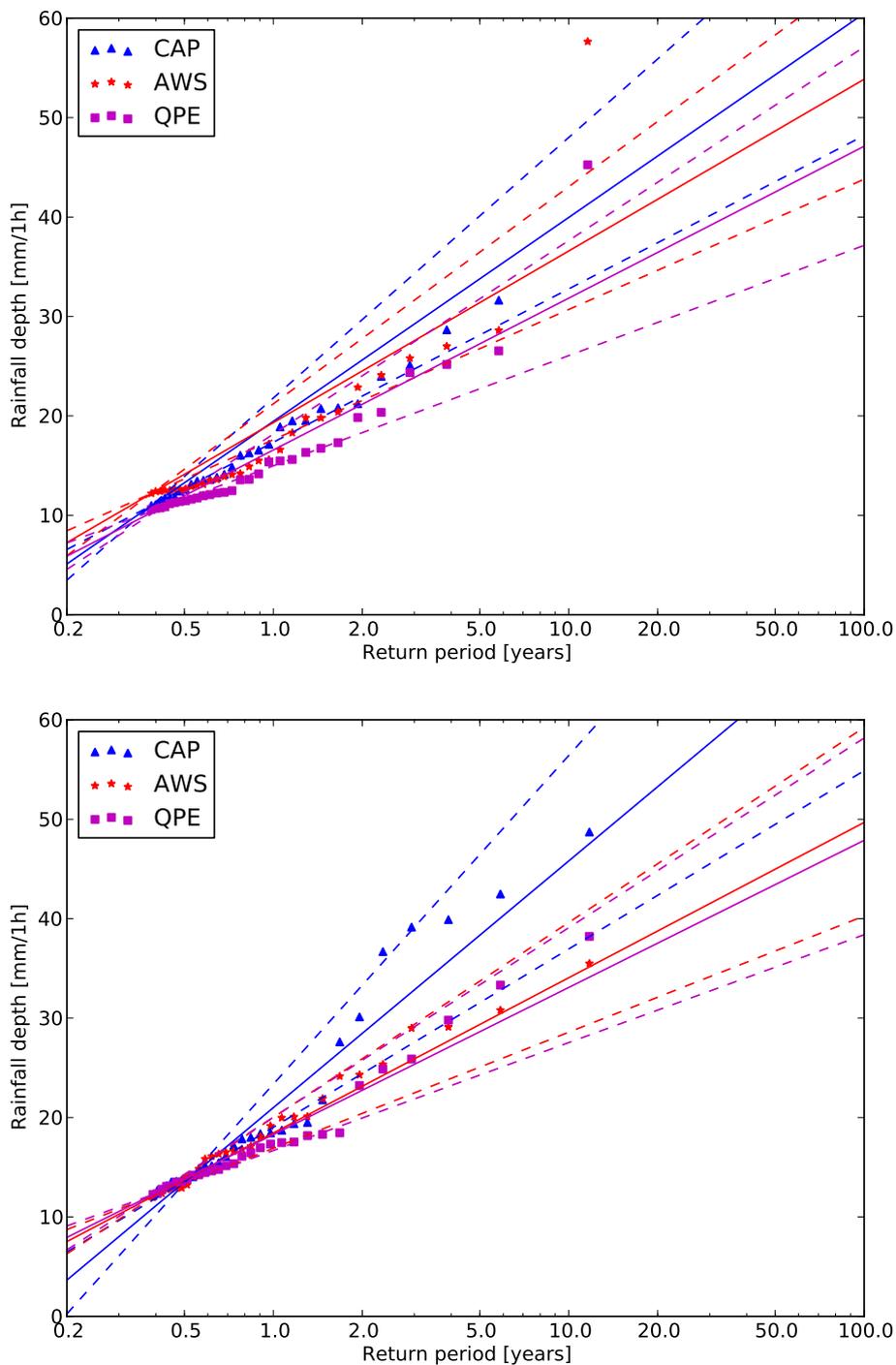


## References

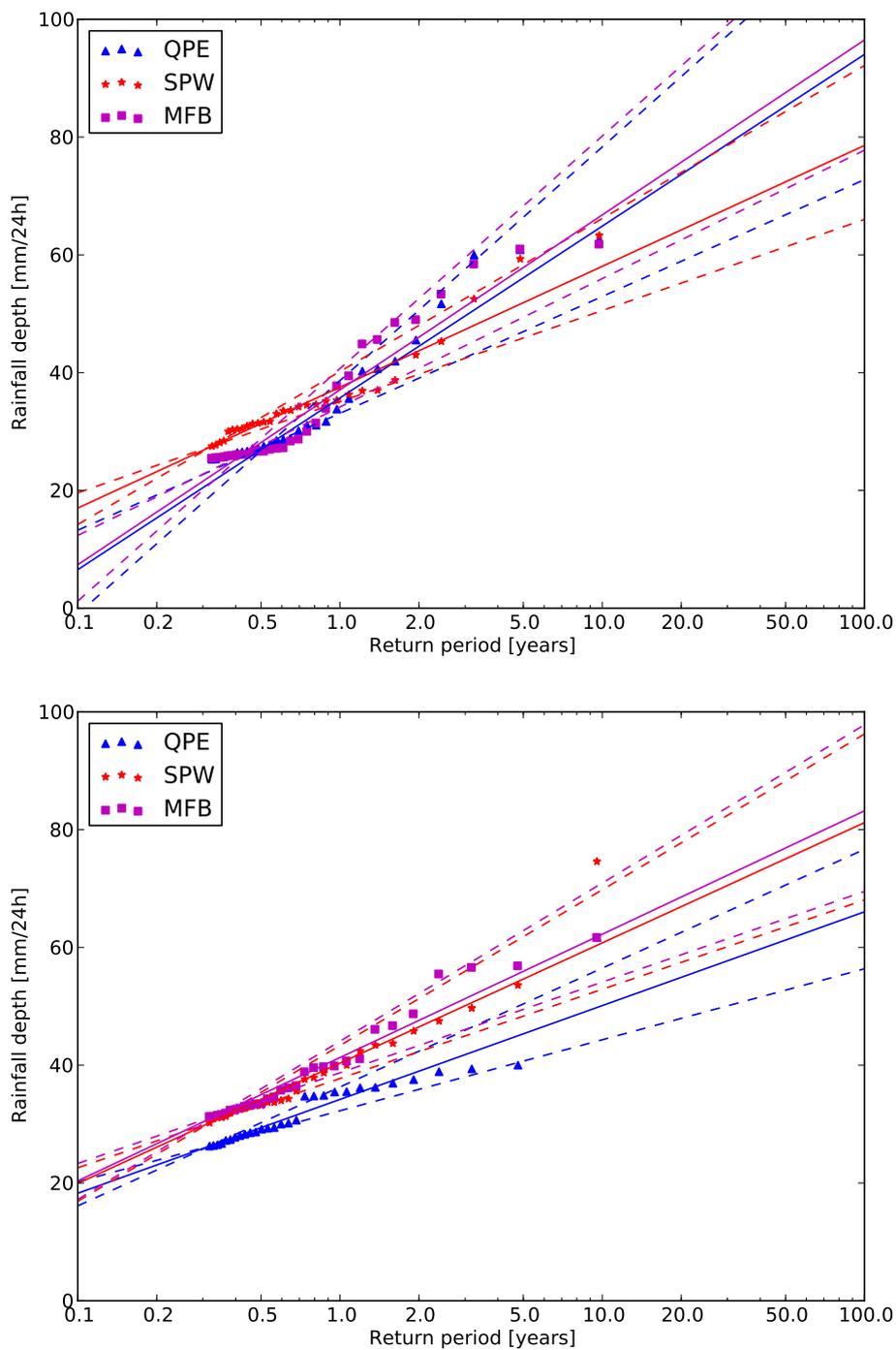
- Buishand, T.: Extreme rainfall estimation by combining data from several sites, *Hydrological Sciences Journal*, 36, 345–365, 1991.
- Buishand, T. A., de Haan, L., and Zhou, C.: On spatial extremes: With application to a rainfall problem, *Ann. Appl. Stat.*, 2, 624–642, doi:10.1214/08-AOAS159, <http://dx.doi.org/10.1214/08-AOAS159>, 2008.
- 5 Colli, M., Lanza, L. G., and La Barbera, P.: An evaluation of the uncertainty of extreme events statistics at the WMO/CIMO Lead Centre on precipitation intensity, *AGU Fall Meeting Abstracts*, 2012.
- Demarée, Gaston: Le pluviographe centenaire du plateau d’Uccle: son histoire, ses données et ses applications, *La Houille Blanche*, pp. 95–102, doi:10.1051/lhb/2003082, <http://dx.doi.org/10.1051/lhb/2003082>, 2003.
- Duchon, C. E. and Biddle, C. J.: Undercatch of tipping-bucket gauges in high rain rate events, *Advances in Geosciences*, 25, 11–15, doi:10.5194/adgeo-25-11-2010, <http://www.adv-geosci.net/25/11/2010/>, 2010.
- 10 Eldardiry, H., Habib, E., and Zhang, Y.: On the use of radar-based quantitative precipitation estimates for precipitation frequency analysis, *Journal of Hydrology*, 531, Part 2, 441 – 453, doi:<http://dx.doi.org/10.1016/j.jhydrol.2015.05.016>, hydrologic Applications of Weather Radar, 2015.
- Figuera i Ventura, J. and Tabary, P.: The new French operational polarimetric radar rainfall rate product, *Journal of Applied Meteorology and Climatology*, 52, 1817–1835, 2013.
- 15 Gellens, D.: Trend and Correlation Analysis of k-Day Extreme Precipitation over Belgium, *Theoretical and Applied Climatology*, 66, 117–129, doi:10.1007/s007040070037, <http://dx.doi.org/10.1007/s007040070037>, 2000.
- Goudenhoofd, E. and Delobbe, L.: Statistical characteristics of convective storms in Belgium derived from volumetric weather radar observations, *Journal of Applied Meteorology and Climatology*, 52, 918–934, 2013.
- 20 Goudenhoofd, E. and Delobbe, L.: Generation and Verification of Rainfall Estimates from 10-Yr Volumetric Weather Radar Measurements, *Journal of Hydrometeorology*, 17, 1223–1242, doi:10.1175/JHM-D-15-0166.1, <http://dx.doi.org/10.1175/JHM-D-15-0166.1>, 2016.
- Haberlandt, U. and Berndt, C.: The value of weather radar data for the estimation of design storms—an analysis for the Hannover region, *Proceedings of the International Association of Hydrological Sciences*, 373, 81–85, 2016.
- Hamdi, R. and Van de Vyver, H.: Estimating urban heat island effects on near-surface air temperature records of Uccle (Brussels, Belgium): an observational and modeling study, *Advances in Science and Research*, 6, 27–34, doi:10.5194/asr-6-27-2011, <http://www.adv-sci-res.net/6/27/2011/>, 2011.
- 25 Journée, M., Delvaux, C., and Bertrand, C.: Precipitation climate maps of Belgium, *Advances in Science and Research*, 12, 73–78, 2015.
- Lukach, M., Foresti, L., Giot, O., and Delobbe, L.: Estimating the occurrence and severity of hail based on 10 years of observations from weather radar in Belgium, *Meteorological Applications*, 2017.
- 30 Marra, F. and Morin, E.: Use of radar {QPE} for the derivation of Intensity–Duration–Frequency curves in a range of climatic regimes, *Journal of Hydrology*, 531, Part 2, 427 – 440, doi:<http://dx.doi.org/10.1016/j.jhydrol.2015.08.064>, hydrologic Applications of Weather Radar, 2015.
- Martins, E. S. and Stedinger, J. R.: Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data, *Water Resources Research*, 36, 737, doi:10.1029/1999WR900330, 2000.
- 35 Ntegeka, V. and Willems, P.: Trends and multidecadal oscillations in rainfall extremes, based on a more than 100-year time series of 10 min rainfall intensities at Uccle, Belgium, *Water Resources Research*, 44, n/a, doi:10.1029/2007WR006471, 2008.



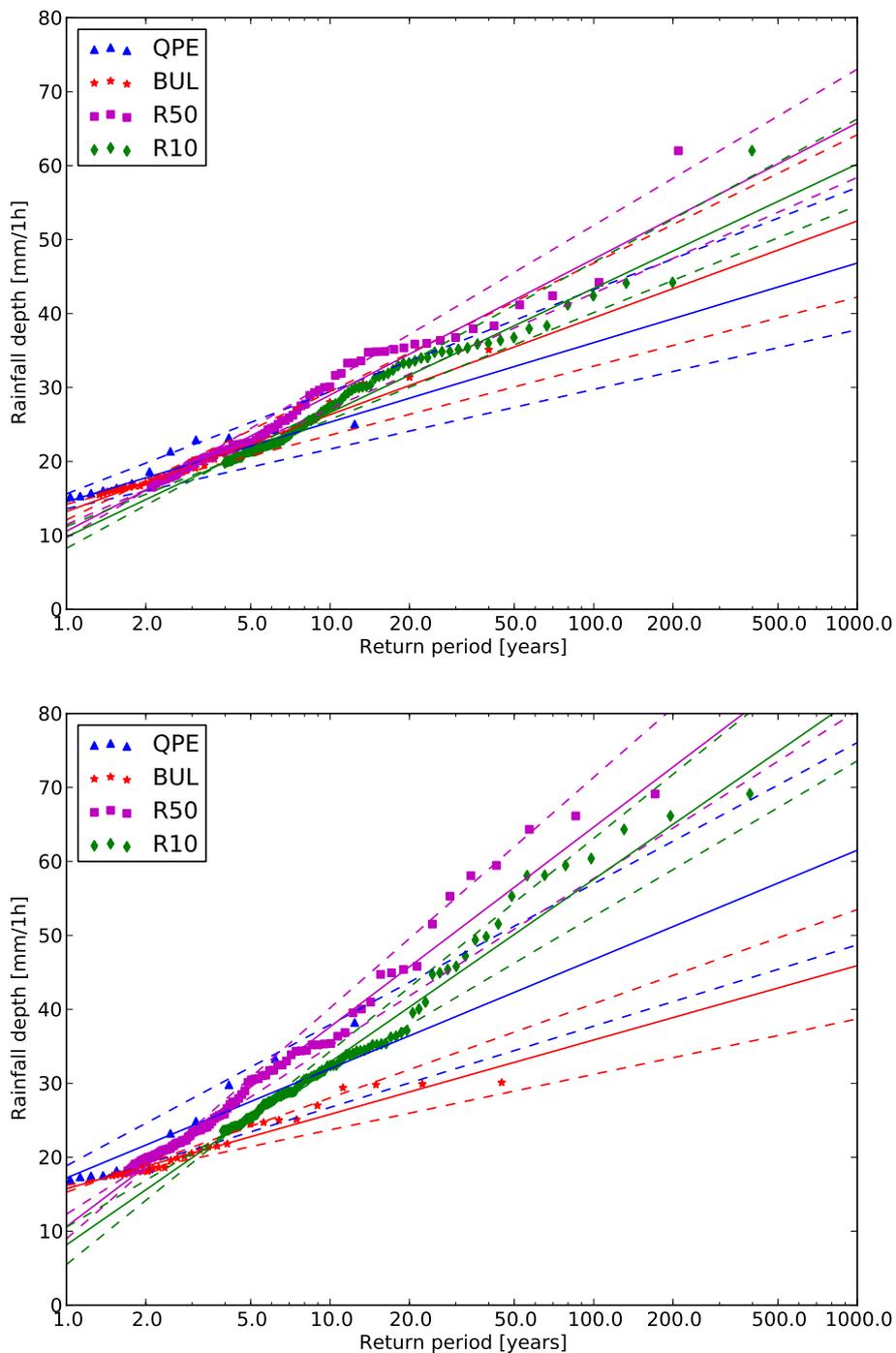
- Nystuen, J. A.: Relative Performance of Automatic Rain Gauges under Different Rainfall Conditions, *Journal of Atmospheric and Oceanic Technology*, 16, 1025–1043, doi:10.1175/1520-0426(1999)016<1025:RPOARG>2.0.CO;2, [http://dx.doi.org/10.1175/1520-0426\(1999\)016<1025:RPOARG>2.0.CO;2](http://dx.doi.org/10.1175/1520-0426(1999)016<1025:RPOARG>2.0.CO;2), 1999.
- Overeem, A., Buishand, T. A., and Holleman, I.: Extreme rainfall analysis and estimation of depth-duration-frequency curves using weather radar, *Water Resources Research*, 45, n/a, doi:10.1029/2009WR007869, 2009.
- 5 Paixao, E., Mirza, M. M. Q., Shephard, M. W., Auld, H., Klaassen, J., and Smith, G.: An integrated approach for identifying homogeneous regions of extreme rainfall events and estimating {IDF} curves in Southern Ontario, Canada: Incorporating radar observations, *Journal of Hydrology*, 528, 734 – 750, doi:<http://dx.doi.org/10.1016/j.jhydrol.2015.06.015>, 2015.
- Pickands III, J.: Statistical inference using extreme order statistics, *the Annals of Statistics*, pp. 119–131, 1975.
- 10 Reed, D. W., Faulkner, D. S., and Stewart, E. J.: The FORGEX method of rainfall growth estimation II: Description, *Hydrology and Earth System Sciences*, 3, 197–203, doi:10.5194/hess-3-197-1999, <http://www.hydrol-earth-syst-sci.net/3/197/1999/>, 1999.
- Rulfova, Z., Buishand, A., Kysely, J., and Roth, M.: Two-Component Extreme Value Distributions for Convective and Stratiform Precipitation, *AGU Fall Meeting Abstracts*, 2014.
- Saito, H. and Matsuyama, H.: Probable Hourly Precipitation and Soil Water Index for 50-yr Recurrence Interval over the Japanese Archipelago, *SOLA*, 11, 118–123, doi:10.2151/sola.2015-028, 2015.
- 15 Svensson, C. and Jones, D.: Review of rainfall frequency estimation methods, *Journal of Flood Risk Management*, 3, 296, doi:10.1111/j.1753-318X.2010.01079.x, 2010.
- Uboldi, F., Sulis, A., Lussana, C., Cislighi, M., and Russo, M.: A spatial bootstrap technique for parameter estimation of rainfall annual maxima distribution, *Hydrology and Earth System Sciences*, 18, 981–995, 2014.
- 20 Uijlenhoet, R. and Pomeroy, J. H.: Raindrop size distributions and radar reflectivity-rain rate relationships for radar hydrology, *Hydrology and Earth System Sciences*, 5, 615–628, 2001.
- Van de Vyver, H.: Spatial regression models for extreme precipitation in Belgium, *Water Resources Research*, 48, n/a, doi:10.1029/2011WR011707, 2012.
- Vannitsem, S. and Naveau, P.: Spatial dependences among precipitation maxima over Belgium, *Nonlinear Processes in Geophysics*, 14, 25 621–630, 2007.
- Willems, P.: Compound intensity/duration/frequency-relationships of extreme precipitation for two seasons and two storm types, *Journal of Hydrology*, 233, 189, doi:10.1016/S0022-1694(00)00233-X, 2000.
- Willems, P., Guillou, A., and Beirlant, J.: Bias correction in hydrologic GPD based extreme value analysis by means of a slowly varying function, *Journal of Hydrology*, 338, 221, doi:10.1016/j.jhydrol.2007.02.035, 2007.



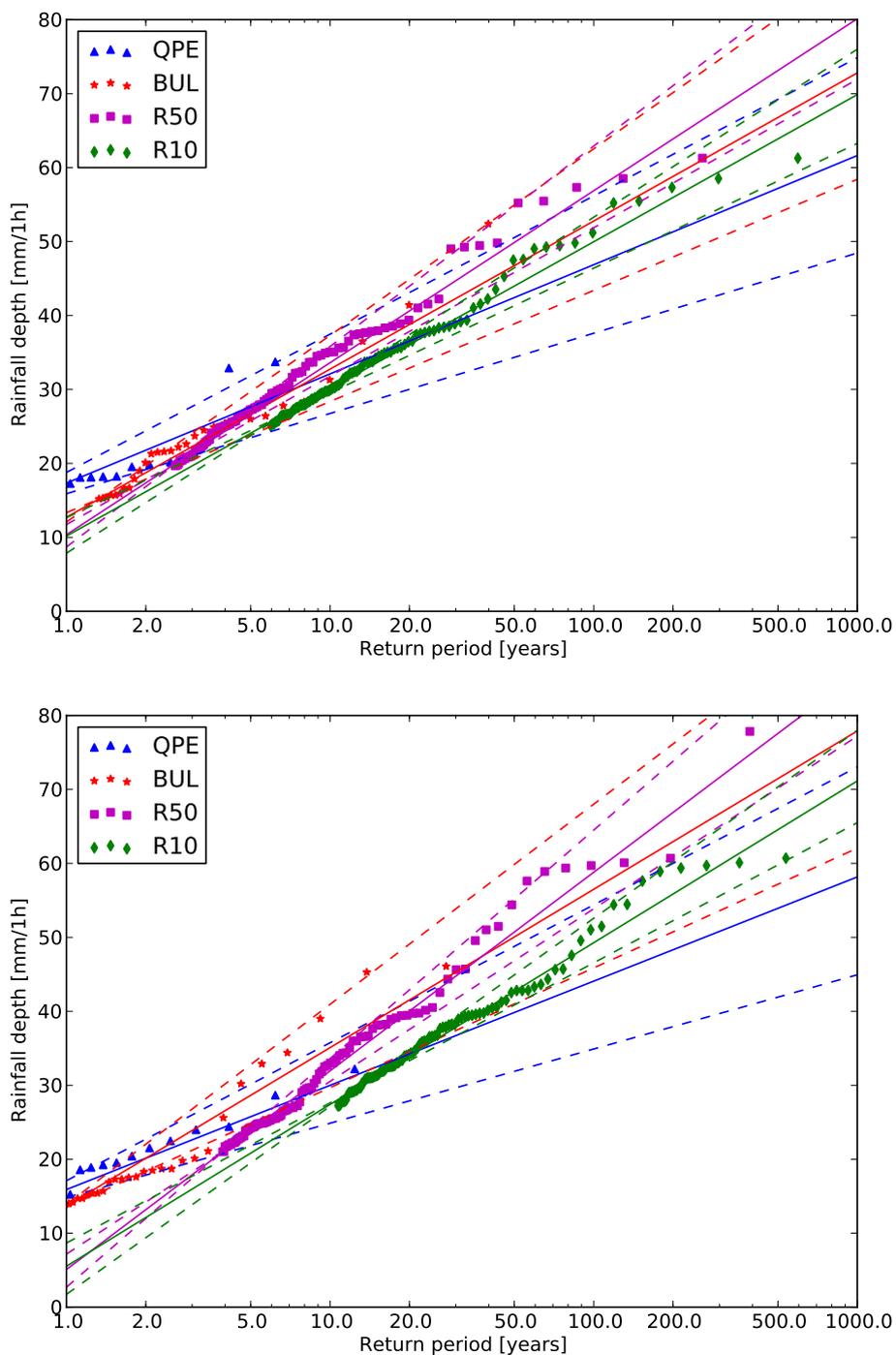
**Figure 2.** Extreme 1-hour precipitation quantiles at location Humain (top) and Uccle (bottom) of the AWS gauge (red stars) compared to CAP (blue triangles) and QPE (magenta squares) radar products. The extreme value distribution (solid line) fitted to the extremes and its confidence intervals (dashed line) are also displayed.



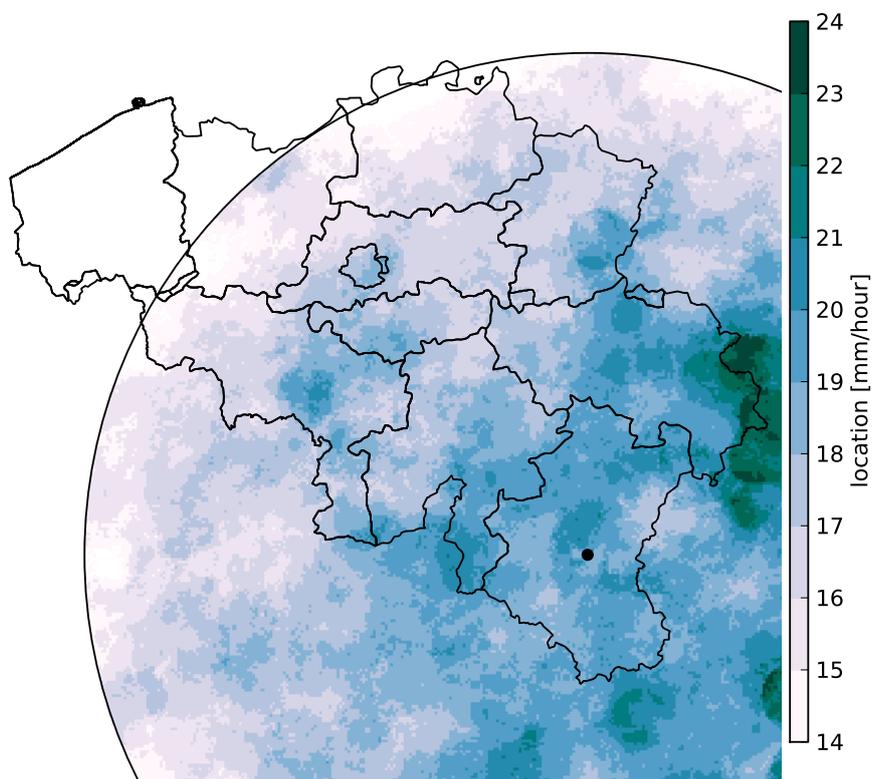
**Figure 3.** Extreme 24-hour precipitation quantiles at location Uccle (top) and Saint-Vith (bottom) of the SPW gauge (red stars) compared to QPE (blue triangles) and MFB (magenta squares) radar products. The extreme value distribution (solid line) fitted to the extremes and its confidence intervals (dashed line) are also displayed.



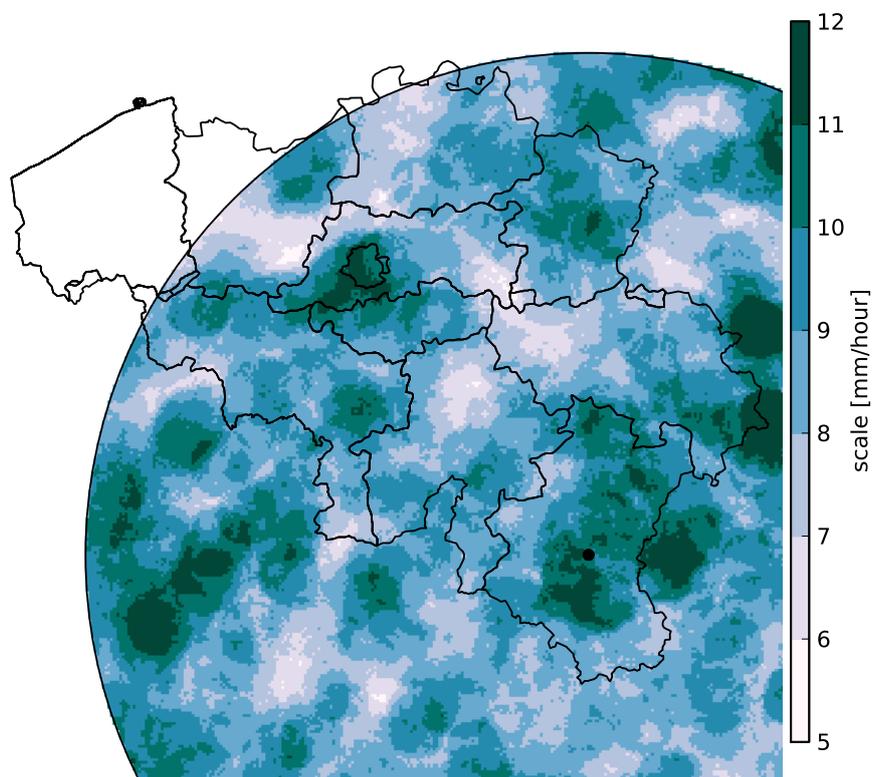
**Figure 4.** Extreme 1-hour precipitation quantiles at location Deurne (top) and Uccle (bottom) from the BUL gauge data (red stars) compared to the at-site QPE (blue triangle), R10 (purple square) and R50 (green diamond) radar data. The extreme value distribution (solid line) fitted to the extremes and its confidence intervals (dashed line) are also displayed.



**Figure 5.** Extreme 1-hour precipitation quantiles at location Gosselies (top) and Nadrin (bottom) from the BUL gauge data (red stars) compared to the QPE (blue triangle), R10 (purple square) and R50 (green diamond) radar data. The extreme value distribution (solid line) fitted to the extremes and its confidence intervals (dashed line) are also displayed.



**Figure 6.** QQR estimator of the location parameter of the Exponential distribution fitted for each pixel to 2005-2016 QPE data in a radius of 10 km.



**Figure 7.** QQR estimator of the scale parameter of the Exponential distribution fitted for each pixel to 2005-2016 QPE data in a radius of 10 km.



**Table 1.** Rain gauge stations used for comparison

Station	Altitude (DNG)	Distance to radar (km)	Duration	Avail. Gauge (%)	Avail. Radar (%)	Avail. All (%)
Humain (AWS)	296	36	1h	98.5	94.8	93.5
Uccle (AWS)	100	128	1h	99.9	94.8	94.7
Uccle (SPW)	100	128	24h	90.6	86.0	78.2
St-Vith (SPW)	456	61	24h	89.2	86.0	76.7
Deurne (BUL)	12	161	1h	86.0	–	–
Uccle (BUL)	100	128	1h	96.3	–	–
Gosselies (BUL)	187	97	1h	85.7	–	–
Nadrin (BUL)	403	30	1h	59.3	–	–



**Table 2.** Comparison of the 10 highest 1-hour precipitation extremes from the gauge (AWS) and radar (QPE) at Humain and Uccle stations

Humain				
Event	Date	Time (end)	Gauge [mm/hour]	Radar[mm/hour]
1	2016-06-07	18:50:00	57.65	45.25
2	2005-07-30	00:40:00	28.60	11.62
3	2014-04-24	15:40:00	27.00	20.35
4	2014-06-10	21:40:00	15.60	26.40
5	2007-06-14	01:20:00	25.80	16.32
6	2008-05-14	17:40:00	13.10	24.35
7	2009-05-25	13:10:00	24.10	25.17
8	2015-07-19	01:00:00	22.87	15.47
9	2009-06-27	14:30:00	20.40	19.83
10	2009-07-22	21:20:00	19.80	12.08
11	2010-07-14	15:40:00	19.80	—
12	2012-06-12	22:20:00	18.30	15.61
13	2013-03-23	07:40:00	—	17.30
14	2005-06-28	22:20:00	—	16.74

Uccle				
Event	Date	Time (end)	Gauge [mm/hour]	Radar [mm/hour]
1	2016-06-07	15:20:00	18.08	38.21
2	2011-08-23	08:40:00	35.50	23.22
3	2009-10-07	18:40:00	30.79	33.32
4	2012-05-20	16:30:00	12.37	29.79
5	2005-09-10	19:40:00	29.10	17.54
6	2011-08-18	15:50:00	28.98	14.77
7	2007-06-14	14:50:00	21.90	25.88
8	2011-09-03	22:40:00	25.34	18.46
9	2016-06-11	18:50:00	—	24.88
10	2005-07-29	19:10:00	24.29	—
11	2010-07-14	15:20:00	24.15	—
12	2014-07-29	16:10:00	20.10	18.17
13	2013-07-27	22:20:00	20.07	—
14	2008-07-26	10:40:00	16.60	18.30



**Table 3.** Results of the extreme value distribution fitting at two locations of the AWS network. The tables shows successively the temporal independence, optimal rank, the location parameter and the scale parameter.

temporal independence [%]

Station	Gauge	CAP	QPE	MFB
Humain	25.6	20.7	22.6	–
Uccle	20.8	19.4	21.0	–

optimal rank

Station	Gauge	CAP	QPE	MFB
Humain	30	30	28	–
Uccle	29	23	30	–

location parameter [mm/hour]

Station	Gauge	CAP	QPE	MFB
Humain	12.2	11.0	10.7	–
Uccle	12.3	13.9	12.3	–

scale parameter

Station	Gauge	CAP	QPE	MFB
Humain	7.5	8.9	6.6	–
Uccle	6.8	10.8	6.4	–



**Table 4.** Comparison of the 10 highest 24-hour precipitation extremes from the gauge (SPW) and radar (MFB) at Uccle and Saint-Vith stations

Uccle				
Event	Date	Time (end)	Gauge [mm/24h]	Radar [mm/24h]
1	2010-08-16	23:00:00	63.30	48.99
2	2009-10-07	23:00:00	52.50	61.83
3	2011-08-23	15:00:00	59.31	61.00
4	2006-08-03	23:00:00	43.00	58.44
5	2016-05-30	23:00:00	35.30	53.34
6	2014-08-26	15:00:00	45.30	48.51
7	2012-10-04	08:00:00	34.60	45.63
8	2012-06-12	11:00:00	—	44.87
9	2016-06-12	17:00:00	31.30	39.45
10	2011-09-04	21:00:00	38.70	26.10
11	2015-08-16	03:00:00	—	37.75
12	2007-06-15	11:00:00	36.99	33.91
13	2014-07-10	04:00:00	36.90	—
14	2016-01-16	02:00:00	36.30	—

Saint-Vith				
Event	Date	Time (end)	Gauge [mm/24h]	Radar [mm/24h]
1	2007-01-18	16:00:00	74.60	56.88
2	2009-07-03	16:00:00	37.90	61.68
3	2011-12-16	23:00:00	—	56.62
4	2012-07-28	21:00:00	53.60	46.72
5	2012-10-04	12:00:00	49.70	39.86
6	2007-08-22	19:00:00	47.50	48.73
7	2010-08-16	03:00:00	45.80	55.50
8	2006-08-05	06:00:00	43.70	41.10
9	2007-12-03	08:00:00	43.40	46.09
10	2007-09-28	08:00:00	42.40	38.87
11	2014-09-21	14:00:00	34.00	40.71
12	2016-05-31	02:00:00	40.01	33.44
12	2016-07-23	21:00:00	40.00	—



**Table 5.** Results of the extreme value distribution fitting at two locations of the SPW network. The tables shows successively the temporal independence, optimal rank, the location parameter and the scale parameter.

temporal independence [%]

Station	Gauge	CAP	QPE	MFB
Uccle	7.1	6.0	6.6	6.7
St-Vith	7.4	8.4	9.0	8.4

optimal rank

Station	Gauge	CAP	QPE	MFB
Uccle	30	26	19	23
St-Vith	30	30	30	28

location parameter [mm/24h]

Station	Gauge	CAP	QPE	MFB
Uccle	27.2	25.0	27.2	27.5
St-Vith	30.2	25.8	26.3	31.5

scale parameter [mm/24h]

Station	Gauge	CAP	QPE	MFB
Uccle	9.0	13.5	12.7	12.9
St-Vith	8.9	8.2	6.9	9.1



**Table 6.** Results of the extreme value distribution fitting for the regional frequency analysis. The tables shows successively the independence (temporal or spatial), the optimal rank, the location parameter and the scale parameter.

independence [%]

Station	QPE	BUL	R50	R10
Deurne	–	27.5	1.4	2.6
Uccle	–	28.0	1.1	2.6
Gosselies	–	22.2	1.7	3.9
Nadrin	–	19.9	2.6	7.0

optimal rank [%]

Station	QPE	BUL	R50	R10
Deurne	28	22	100	99
Uccle	30	30	70	88
Gosselies	29	30	96	90
Nadrin	23	30	100	91

location parameter [mm/hour]

Station	QPE	BUL	R50	R10
Deurne	10.8	16.7	16.5	20.0
Uccle	11.5	17.5	21.1	24.2
Gosselies	11.9	15.2	20.4	26.5
Nadrin	12.2	12.9	21.0	29.0

scale parameter [mm/hour]

Station	QPE	BUL	R50	R10
Deurne	4.7	5.7	8.0	7.3
Uccle	6.4	4.4	11.7	10.7
Gosselies	6.4	8.7	10.1	8.6
Nadrin	6.1	9.3	11.7	9.5