https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

Quantitative rainfall analysis of the 2021 mid-July flood event in Belgium

Michel Journée[1], Edouard Goudenhoofdt[1], Stéphane Vannitsem[1], and Laurent Delobbe[1]
[1]Royal Meteorological Institute of Belgium, Avenue Circulaire 3, 1180 Brussels, Belgium
Correspondence: Michel Journée (michel.journee@meteo.be)

Abstract. The exceptional flood of July 2021 in central Europe impacted severely Belgium. As rainfall was the triggering factor of this event, this study aims at characterizing rainfall amounts in Belgium from July 13th to July 16th 2021 based on 2 types of observational data. First, observations recorded by high-quality rain gauges operated by weather and hydrological services in Belgium have been compiled and quality checked. Second, a radar-based rainfall product has been improved to provide a reliable estimation of quantitative precipitation at high spatial and temporal resolutions over Belgium. Several analyses of these data are here performed to describe the spatial and temporal distribution of rainfall during the event. These analyses indicate that the rainfall accumulations during the event have reached unprecedented levels over large areas. Accumulations over durations from 1 to 3 days significantly exceeded the 200-year return level at several places, with up to 90% of exceedance over the 200-year return level for 2-day and 3-day values locally in the Vesdre basin. Such a record-breaking event needs to be documented as much as possible and available observational data must be shared with the scientific community for further studies in hydrology, in urban planning and, more generally, in all multi-disciplinary studies aiming at identifying and understanding factors leading to such disaster. The corresponding rainfall data are therefore provided freely as a supplement (Journée et al., 2023; Goudenhoofdt et al., 2023).

1 Introduction

From July 13 to 16 2021, a long period of sustained and heavy rainfall affected Central Europe producing extreme rainfall amounts in western Germany, eastern Belgium, Luxembourg and The Netherlands. This meteorological event combined with already near-saturation soil conditions and steep slopes of several river valleys in the region caused disastrous flooding (Kreienkamp et al., 2021; Mohr et al., 2022). This event was one of the most severe natural catastrophes in Europe in the last half century and was responsible for at least 220 fatalities and loss amounts estimated to EUR 46 billion (MunichRe, 2022; Mohr et al., 2022). In the Walloon region of Belgium alone, 39 people lost their lives and the total cost for this region is estimated to EUR 2.8 billion (Gouvernement Wallon, 2022). According to Assuralia (2022), the total amount of compensations paid in Belgium by the insurance companies is EUR 2 billion. This exceptional event occurred within a low-pressure system, called Bernd, very slowly moving over Central Europe, feeding the flooded region with very moist air. The associated rainfall was characterized by a sustained large-scale stratiform component combined with locally embedded convective precipitation. In eastern Belgium, this led to extreme rainfall amounts breaking many historical rainfall records at several locations. Excep-

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

tional floods were observed in several tributaries of the Meuse catchment, in particular along the Vesdre where the most severe consequences have been deplored in terms of casualties and damage (Dewals et al., 2021). In the Vesdre catchment, an increase of the seismic noise and an increasing saturation of the weathered zone have been observed during this event by a seismometer and a gravimeter, respectively (Van Camp et al., 2022).

After such a disaster, questions arise on the role of climate change on the occurrence of this type of event. An attribution study has been rapidly performed (Kreienkamp et al., 2021) and concludes that the likelihood of such an event to occur today at any place over Western Europe compared to a 1.2 °C cooler climate has increased by a factor between 1.2 and 9. According to the Clausius-Clapeyron relation, a warmer atmosphere can contain more water vapour which is expected to increase the intensity and frequency of precipitation events. This type of relationship is however difficult to ascertain as the occurrence of such an event is both dynamically and thermodynamically driven (Ludwig et al., 2022; IPCC, 2021; Vergara-Temprado et al., 2021). This makes the attribution of such events to climate change a real challenge, and further analyses are required to address this issue as for instance in Meyer et al. (2022). Such type of analyses strongly rely on the quality of the available observational data.

As described in Mohr et al. (2022), these devastating floods are the result of complex interactions between meteorological, hydrological and hydro-morphological phenomena. An additional aspect that needs to be considered is the presence of dams upstream of some affected valleys in eastern Belgium (Dewals et al., 2021). Multi-disciplinary analysis are required for an in-depth understanding of the course of the events. Investigating the complex dynamics in the affected catchments is necessary to understand the relation between extreme rainfall and the resulting impact. Such analysis require a detailed knowledge of rainfall, as the triggering factor of these events. Extremely rare event need to be documented as much as possible and data must be made available for further studies in hydrology, in urban planning and, more generally, in all multi-disciplinary studies aiming at identifying and understanding factors leading to such disaster.

Several extreme rainfall and flood events that occured in the last decades have been documented in the literature. For example, the extraordinary rainfall and flash flood event which affected the eastern part of The Netherlands in August 2010 is described in Brauer et al. (2011). In Germany, the hydrometeorology of the extreme flood in the Starzel catchment on 2 June 2008 is analysed in Ruiz-Villanueva et al. (2012). On 8-9 September 2002, a catastrophic flash flood affected the Gard region (France) with maximum 24h rainfall values of 600-700 mm. This event is extensively documented and analysed in Delrieu et al. (2005). In Marchi et al. (2010), hydrometeorological data from 25 extreme flash floods across Europe are presented and a high-resolution dataset of rainfall and discharge observations for 49 events in Europe and the Mediteranean region from 1991 to 2015 is described in Amponsah et al. (2018). All these studies underline the importance of rainfall observations at high spatial and temporal resolutions for the post-analysis of extraordinary flood events.

In the present paper, the observational rainfall data for Belgium that can be used for such analyses are exposed and made available to the scientific community (Journée et al., 2023; Goudenhoofdt et al., 2023). These data are twofold: (1) in-situ observations from high-quality rain gauges and (2) rainfall products based on weather radar observations. The radar data are carefully processed and combined with rain gauges measurements to provide a quantitative estimation of precipitation at high

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

spatial (i.e., 1 km) and temporal (i.e., 5 min and hourly) resolutions. Several analyses of these data are performed to describe the spatial and temporal distribution of rainfall during the event and to illustrate its exceptional character.

The paper is organized as follows. In Section 2, the rain gauges data and the radar product available for Belgium during the period from July 13th to July 16th 2021 are detailed. Section 3 includes several analyses of these data to discuss the event regarding its spatial and temporal extent together with its exceptional character. Some conclusions are drawn in Section 4.

2 Precipitation data

2.1 Rain gauges data

Several rain gauges networks are deployed in Belgium to continuously monitor rainfall at ground level. These networks are operated by the Royal Meteorological Institute of Belgium (RMI) and by the Belgian regional hydrological services, i.e., Service Public de Wallonie - Mobilité et Infrastructures (https://hydrometrie.wallonie.be), Vlaamse Milieumaatschappij and Hydrologisch Informatiecentrum (https://www.waterinfo.be/). In total, 323 rain gauges located in 308 different sites have recorded precipitation quantities during the severe rainfall event of mid-July 2021. The spatial distribution of these 308 sites is illustrated in Figure 3 by the grey dots. Among the 323 available devices, 168 weighing rain gauges monitored rainfall in real-time with a 5-min time resolution and 155 manual rain gauges recorded daily precipitation amounts. Manual precipitation observations are made each day at 08:00 a.m. local time.

All the recorded rain gauges data were checked manually for possible errors and inconsistencies. This data quality control (QC) analysis is made with the help of maps and time series plots that allow to compare data from neighboring rain gauges. Comparisons are also made with meteorological radar products. This validation process highlighted that 2 weighing rain gauges provided zero precipitation data continuously during several hours, which was inconsistent with the heavy rainfall observed by neighboring rain gauges as well as by the radars during that period. In addition, a few gaps of short duration (usually 1 or 2 successive timestamps) were reported in the 5-min time series of some other weighing rain gauges. Estimations derived from neighboring weighing rain gauges data have been considered to correct the 2 periods of erroneous data as well as to fill the gaps. Regarding the manual rain gauges, some observers measured the daily precipitation amount later than 08:00 am. As a consequence, the total accumulation during the event was correctly recorded but not properly distributed on each day, needing adjustments of the daily values. Globally, the QC analysis led to few interventions on all available rain gauges data.

2.2 Radar-based rainfall product (RADFLOOD21)

The radar product is obtained after a careful processing of the weather radar measurements and a merging with validated rain gauge measurements. Some of the main challenges of radar-based precipitation estimation are discussed in details in Goudenhoofdt and Delobbe (2016). The method is under a continuous improvement process based on research and quality control. In particular, the significant underestimation of the operational product during the flood event, has led to several

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

90 improvements. Additionally, some of the parameters have been optimized for this particular case. The processing steps are explained below.

2.2.1 Weather radar measurements

Radars transmit electromagnetic pulses, typically within a beam width of 1 degree. Part of the transmitted power is reflected back to the radar by precipitation. Most radars in Europe perform a full volume scan of the atmosphere at different elevation
95 angles (3D) in about 5 minutes. Estimating rainfall from radar measurements is a challenge because of the many sources of errors and uncertainties (Figure 1).

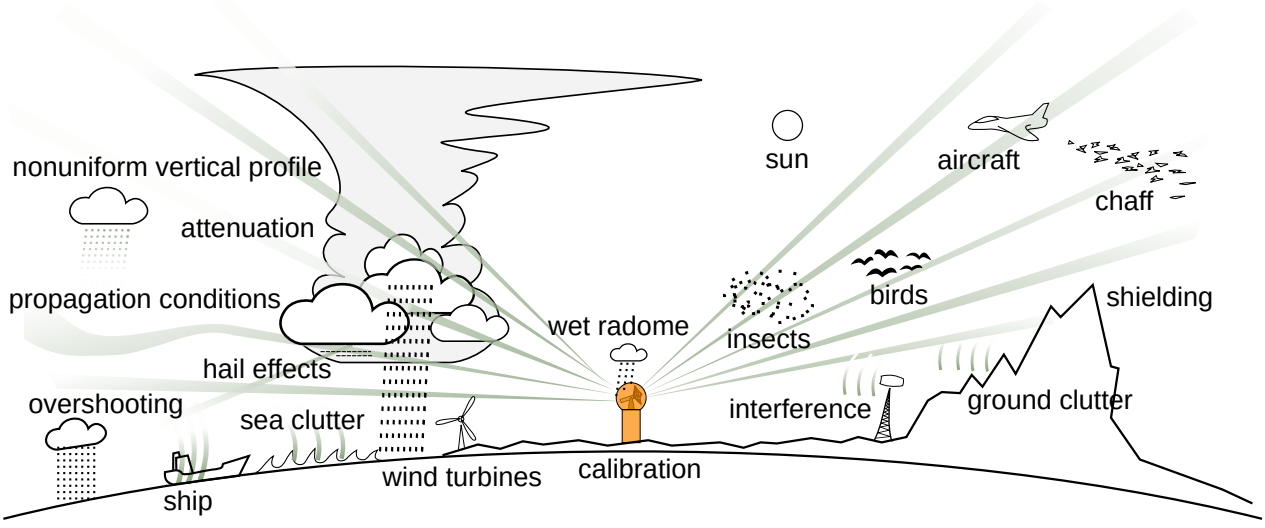

Figure 1. Phenomena affecting the radar data quality. Courtesy of Markus Peura (Finnish Meteorological Institute).

The product is based on the 3D reflectivity measurements of the following radars:

- Helchteren, Vlaamse Milieumaatschappij, Belgium (BEHEL)
- Jabbeke, Royal Meteorological Institute of Belgium, Belgium (BEJAB)
- 100 Wideumont, Royal Meteorological Institute of Belgium, Belgium (BEWID)
- Neuheilenbach, Deutsche Wetterdienst, Germany (DENHB)
- Essen, Deutsche Wetterdienst, Germany (DEESS)
- Avesnois, Meteo France, France (FRAVE)

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

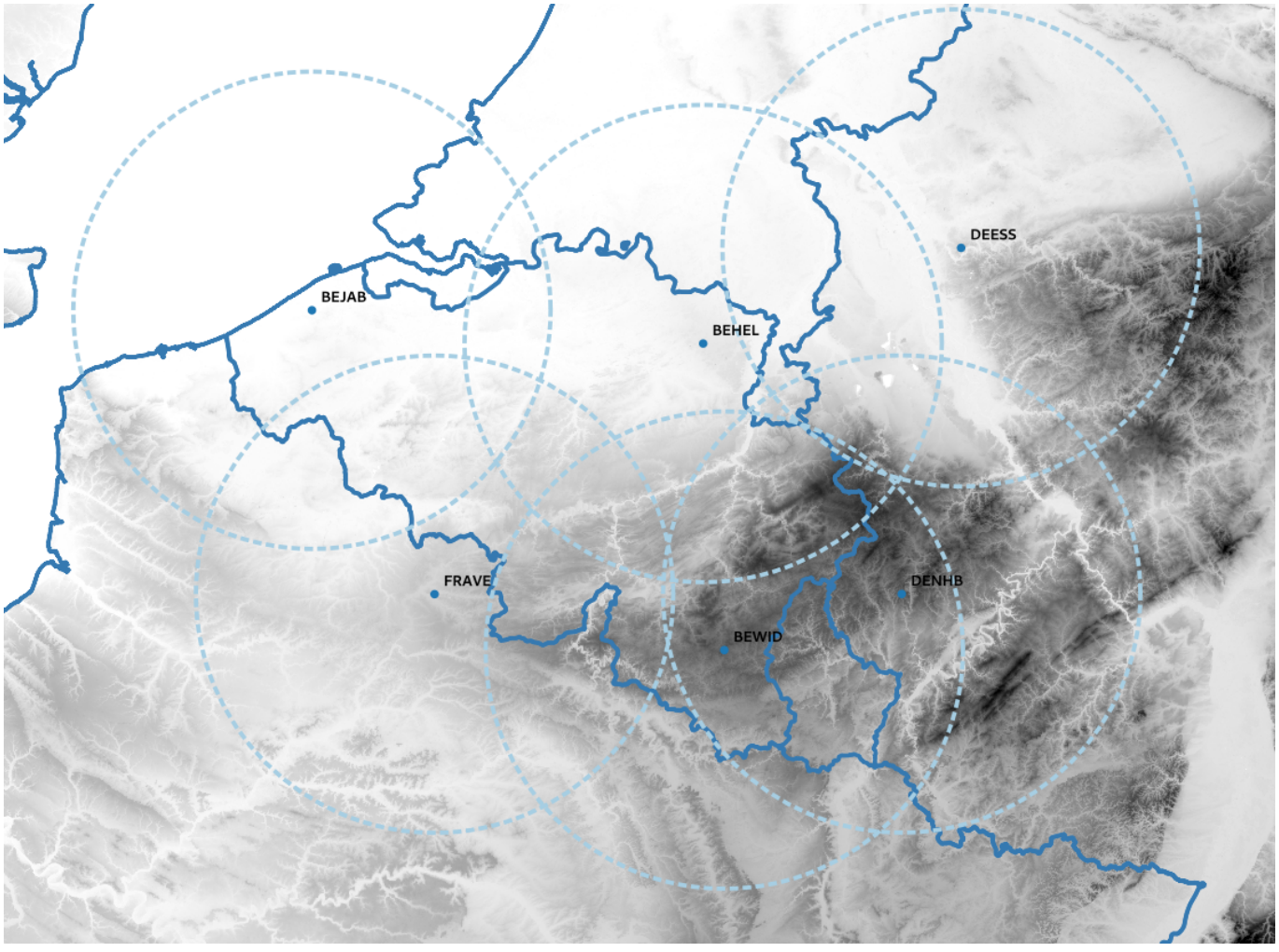

Figure 2. Radar coverage with 100 km radius, country borders and height above see level (in gray scale from 0 to 1000 meter).

These radars use different kind of technology, scanning strategy and data processing, which can evolve over time. A particular challenge is to mitigate the in-homogeneity between the radars and loss of information in the data processing chain.

A radar has a typical range of 250 km but the quality of the measurements tends to decrease with the distance. Data within a 100 km radius are generally considered as appropriate for rainfall estimates of good quality (Figure 2).

2.2.2 Quality control of the radar measurements

The quality control starts by checking the long-term calibration bias of the radar. Using a basic radar estimation method, an average of the bias is computed based on comparison with gauge measurements over the past two months. The method is

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

tuned to remove all uncertainties and errors not related to calibration errors. The calibration bias is removed before further processing.

Radar measurements over a given area can be permanently or regularly affected by clutter (i.e. non-meteorological echoes) coming from hills or wind farms. This can be solved by removing measurements with an abnormally high frequency of echoes. A new method has been developed to keep values as actual precipitation if they are significantly larger than the maximum expected clutter level. The goal is to include radar observations as close to the ground as possible without contamination by ground echoes. This is important to mitigate underestimation due to orographic enhancement of precipitation taking place in the lowest layers of the atmosphere. We suspect that this effect was particularly strong during the flood event.

The radar beam can be partially blocked by elevated areas. The percentage of energy lost is therefore computed. The reflectivity measurements is then corrected accordingly.

The radar measurements can be contaminated by residual non meteorological echo from planes, wireless devices or the ground in case of abnormal propagation. Such clutter are identified based on three methods: (1) comparison with the satellite cloudiness product, (2) detection of abnormal changes between measurements at different altitudes and (3) detection of unrealistic spatial texture.

2.2.3 Rainfall rate estimation at the ground

The processing starts by the identification of convective precipitation based on reflectivity gradients. Non convective precipitation is extrapolated to the ground by using an averaged vertical profile of reflectivity. This allows in particular to mitigate the overestimation caused by melting snow. Missing data after the quality control can be replaced by data from higher radar beams or by data in a close neighborhood. The radar reflectivity (Z) is then converted into rain rates (R) based on the identified precipitation type (snow, hail, convective rainfall, widespread rainfall). A special relation is used for non-convective precipitation in areas with orography, where some enhancement is expected ($Z = 200R^2$, following the US Radar Operations Center).

2.2.4 Compositing, accumulation and merging

In order to mitigate the underestimation produced by radar signal attenuation by rainfall along the path, the observations from several radars are combined. Only data from the 3 closest radars and within a range of 180 km are here used. Data at very long range with low spatial resolution are not included.

Instantaneous rain rates are obtained every 5 minutes corresponding to the full 3D radar scan. The rainfall accumulation over 5 minutes is obtained by computing the movement of precipitation using optical flow techniques. Accumulations over longer durations are obtained by taking the sum of the 5-min accumulations.

The accumulated rainfall over 1 hour is combined with rain gauge measurements from the automatic stations using Kriging with external drift (KED). The KED method interpolates the gauge values in a given neighborhood while taking the radar estimation as a linear combination of the expected value of the (Gaussian) process. The correlation of the residual errors is assumed to decrease up to 20 km. Note that the gauge measurements and sampling errors are taken into account. The spatial correction factor is then applied to the 5 minutes accumulation.

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

3 Quantitative rainfall analysis

3.1 Total rainfall distribution in space

Precipitation accumulations over 3 days, from July 13th 06:00 UTC to July 16th 06:00 UTC, vary largely over Belgium, as illustrated on Figure 3. While the weather was completely dry during this period in the North-West of Belgium, rainfall accumulations reached almost 300mm in the eastern part of the country (i.e, the rain gauges recorded up to 291.6mm). Accumulations over the 3-day period exceeding 200mm have been recorded by 4 weighing rain gauges (Jalhay, Spa, Mont Rigi and Neu-Hattlich) and by one manual rain gauge (Hockai). These 5 rain gauges are located in the Province of Liège.

In Figure 3, the 3-day precipitation total is mapped over Belgium in two ways. On one hand, all available 3-day totals recorded by weighing and manual rain gauges are spatially interpolated by an inverse distance weighted (IDW) approach (Figure 3, panel a.). On the other hand, the hourly radar product data are temporally aggregated on the 3-day period (Figure 3, panel b.). Both estimations provide on a large-scale perspective comparable rainfall distributions over the country. Differences are however noticeable at finer scales. First, the IDW approach tends to give artificial circular patterns around some rain gauges, while radar observations are expected to better capture fine-scale rainfall patterns. Second, the IDW approach is blind is areas with a poor density in rain gauges. The difference between both estimations (Figure 3, panel c.) shows that the largest discrepancies are located in an area without any rain gauges in the North of Belgium. Significant differences can also be noted in eastern part of the Province of Liège where the precipitation accumulations are the largest and vary strongly on short distances. Such rainfall patterns with large gradients need a very dense network of rain gauges to be accurately captured solely by ground observations. The radar product is thus essential for the spatial analysis of such precipitation events. In particular, the radar product shows that the 3-day precipitation total has exceeded 200mm for a narrow but elongated area of 265 km^2 oriented from southwest to northeast.

In order to estimate the uncertainty associated to both spatial distributions of the 3-day precipitation accumulation, a validation analysis is conducted as follows. As the radar product do not integrate observations from manual rain gauges, the 3-day totals from 155 manual rain gauges serve as validation data. For the IDW-derived spatial distribution that relies on both weighing and manual rain gauges data, a leave-one-out cross-validation is applied, i.e., 155 IDW estimations are computed by leaving systematically one different manual rain gauge data out which is used as reference to evaluate the estimation error. The results are illustrated by scatter plots in Figure 4. Concerning IDW, these results indicate an overestimation of the smallest values and an underestimation of the largest ones. IDW tends to smooth the spatial distribution of rainfall and attenuates extremes. Regarding the radar product, a rather good match with the manual observations can be noticed for 3-day accumulations till 100mm. The largest values are however generally underestimated. The derived Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) indicate that the radar product provide nevertheless a more accurate spatial distribution of the 3-day accumulation than IDW.

Based on the extreme precipitation statistics derived for Belgium by Van de Vyver (2012) and Van de Vyver (2013), return periods can be associated to the 3-day precipitation accumulations, as illustrated in Figure 5 for the rainfall distribution estimated by the radar product (i.e., Figure 3b). Large return periods can be noted for extended areas within Belgium. For instance,

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

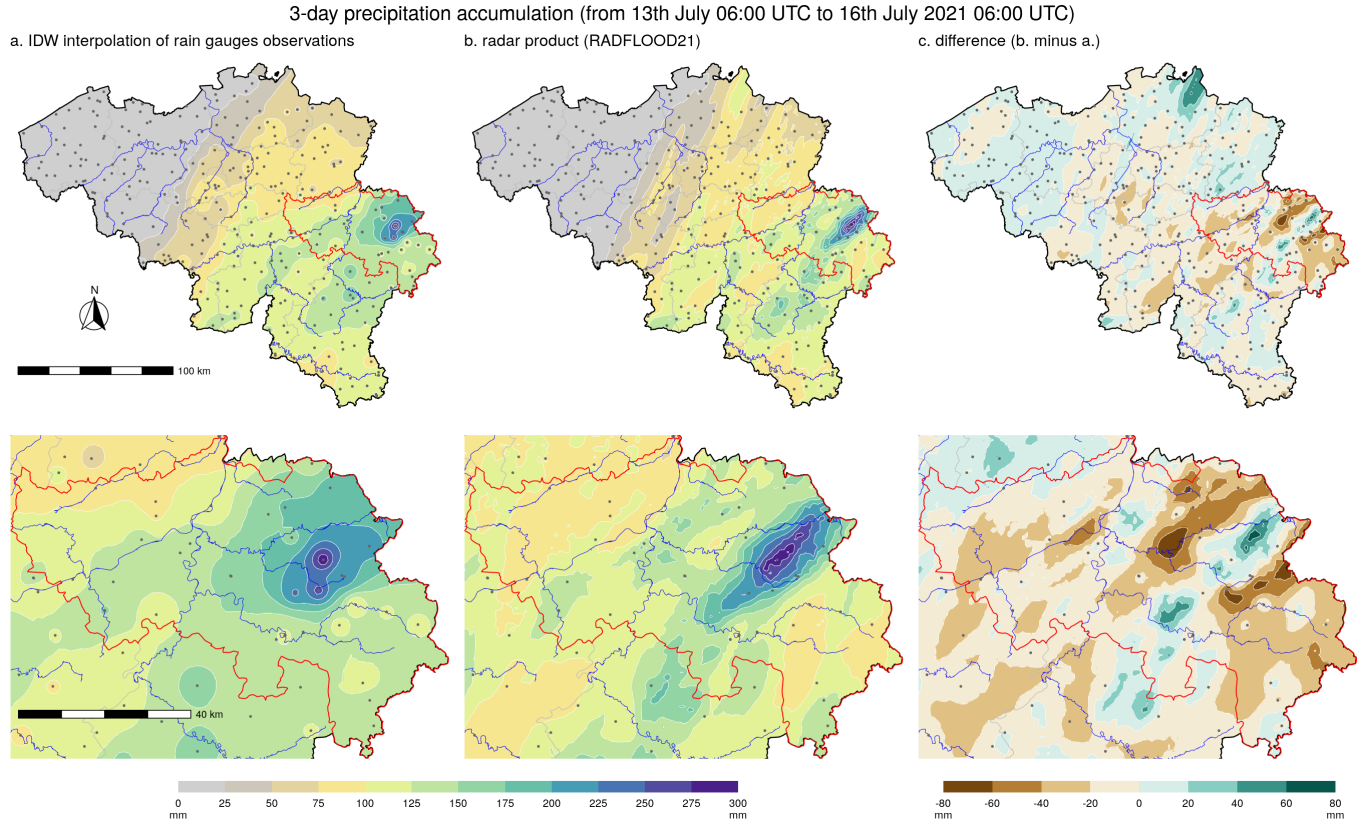

Figure 3. Spatial distribution in Belgium (top panels) and in the Province of Liège (bottom panels) of the precipitation accumulation over 3 days (from July 13th 06:00 UTC to July 16th 06:00 UTC). The spatial distribution is derived by inverse distance weighted (IDW) interpolation of all available automatic and manual rain gauges data (left panels) as well as by the radar product (center panels). The difference between both estimations is provided in the right panels. Rain gauges locations are displayed by the grey dots and the red delimited area corresponds to the Province of Liège.

the return period estimated for the 3-day total exceeds 50 years for 24.5% of the Belgian territory (i.e., 7500 km^2). The area with a return period estimated larger than 100 years covers more than 13% of the country (i.e., 4000 km^2). It should be noted that the extreme precipitation statistics provided by Van de Vyver (2012) and Van de Vyver (2013) are here limited to return periods of 200 years, as the uncertainty increases considerably when considering return periods significantly larger than the length of the time series. This upper level is exceeded for an area corresponding to 6.5% of Belgium (i.e., 2000 km^2). At some places, the 3-day total exceeds considerably the 200-year return level.

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

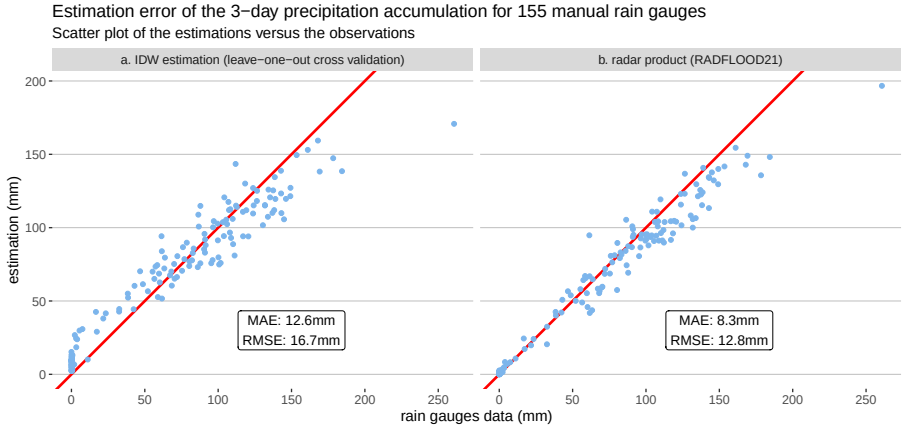

Figure 4. Scatter plots of the IDW and radar product estimations versus the observations of the 3-day accumulation for 155 manual rain gauges. The average discrepancies between estimations and observations are summarized by the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) for both scatter plots.

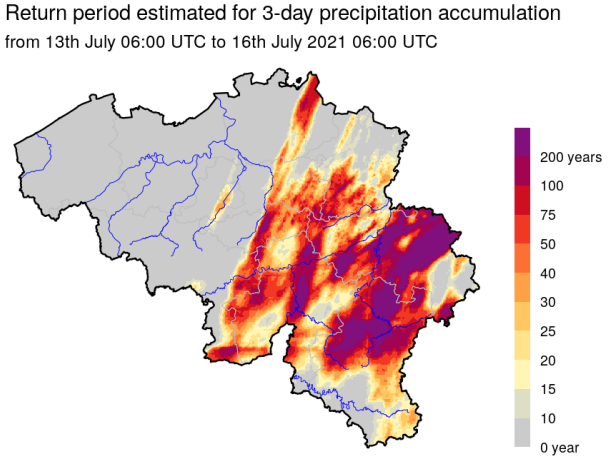

Figure 5. Return period estimated for the 3-day precipitation total in Belgium (corresponding to the rainfall distribution estimated by the radar product, i.e., Figure 3b).

3.2 Rainfall distribution in time

185 Figure 6 provides a closer look to the rainfall distribution in time for the 4 weighing rain gauges with a 3-day accumulation over 200mm. These time series indicate that, for these locations, most of the total rainfall accumulation occurred in a period of

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

approximately 36 hours, thus much smaller than 3 days. It appears also that, within this 36-hour period, the hourly precipitation totals are highly variable with several peaks and some hours with almost no precipitation.

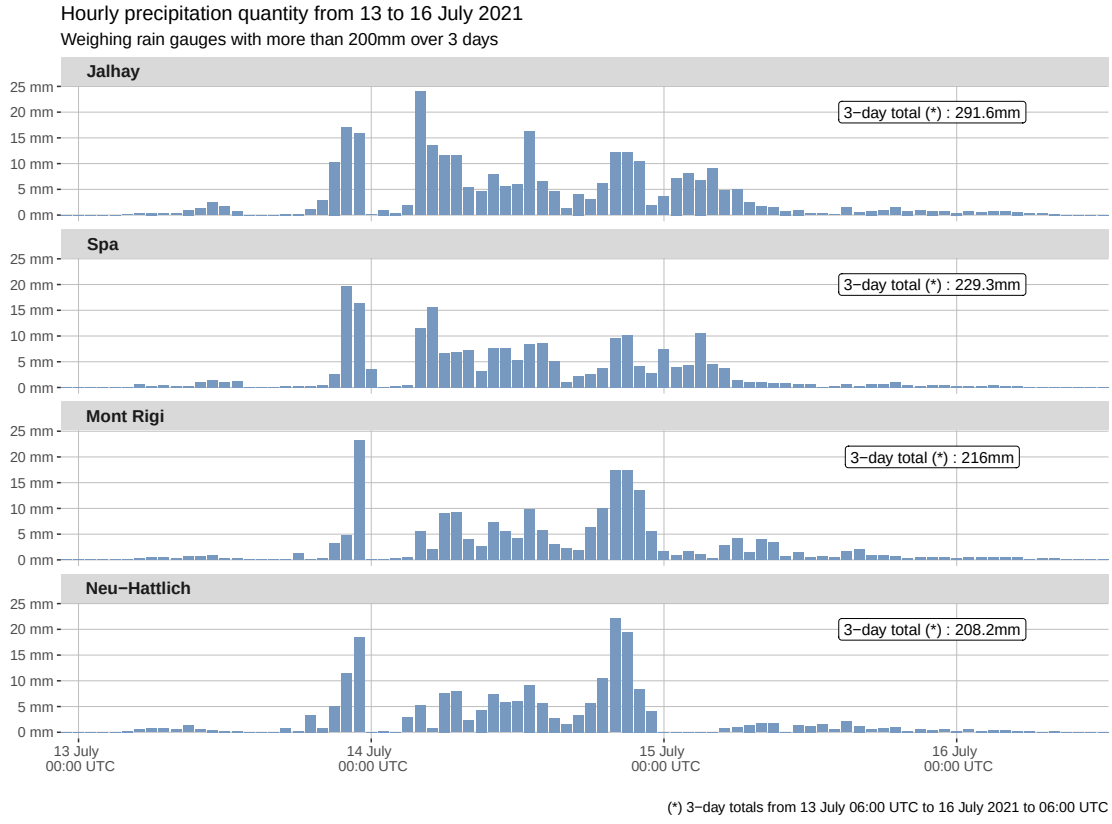

Figure 6. Time series of hourly precipitation totals for the 4 weighing rain gauges with a 3-day accumulation exceeding 200mm. These 4 rain gauges are located in the Province of Liège.

For the 5 manual and weighing rain gauges with the largest 3-day total (i.e., exceeding 200mm), the maximum precipitation accumulation for durations from 1 hour to 3 days is provided in Figure 7 and compared against precipitation amounts corresponding to events with return periods from 10 to 200 years, i.e., 10-year to 200-year return levels. For these stations which are all located in the same area of the Province of Liège, precipitation accumulations on short durations till 3 hours do not exceed a return period of 30 years. The precipitation quantities tend however toward extreme levels when increasing the accumulation duration: 6-hour accumulations present a return period above 75 years at 3 places and 12-hour accumulations exceed the 200-year return level at the 4 weighing rain gauges. For the 5 considered locations, the exceedance over the 200-year return level is even larger for precipitation accumulations over 1 day and especially over 2 and 3 days: these 2-day and 3-day values exceed the 200-year return level by 30% to 90% depending on the location. In Jalhay, where the largest 2-day and 3-day amounts have

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

been recorded, these 2-day and 3-day values correspond to the 200-year return level of rainfall events with a duration of 13 days and 15 days respectively, as it can be noted in Dewals et al. (2021).

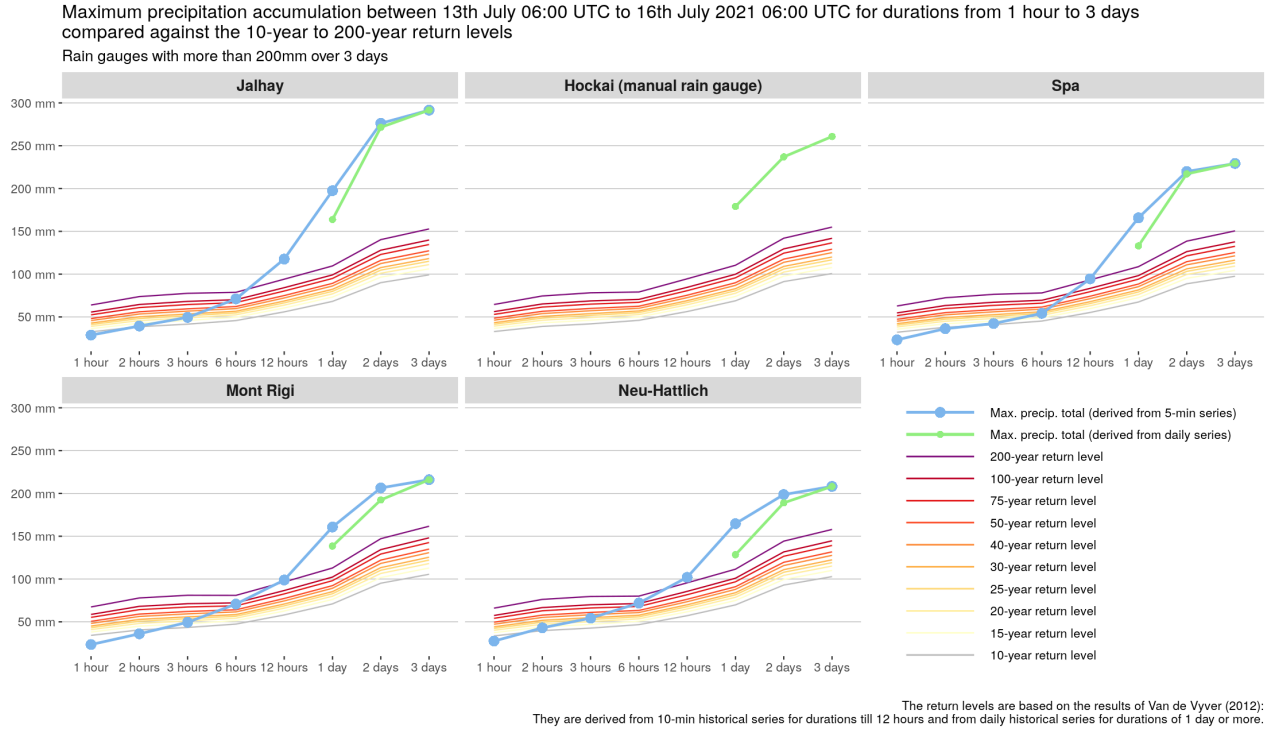

Figure 7. Maximum precipitation accumulation for durations from 1 hour to 3 days for the 5 rain gauges with a 3-day accumulation exceeding 200mm compared against the 10-year to 200-year return levels. These maximum values are derived from 5-min time series (weighing rain gauges) as well as from daily time series (weighing and manual rain gauges). These 5 rain gauges are located in the Province of Liège. The return levels are derived from 10-min historical series for durations till 12-hours and from daily historical series for durations of 1 day or more (Van de Vyver, 2012, 2013).

The analysis of Figure 7 can be extended to any location in Belgium based on the hourly radar product. For each 1-km pixel of the radar product, the maximum rainfall accumulation is derived for durations from 1 hour to 3 days during the period July 13th 06:00 UTC to July 16th 06:00 UTC. Return periods can then be associated to these precipitation maxima based on the results of Van de Vyver (2012) and Van de Vyver (2013). Figure 8 summarizes the frequency distribution of these 1-km sampled return periods for the 8 considered accumulation durations. The spatial distribution of the estimated return periods for each accumulation duration can be found in Figure A1. These results indicate that the precipitation accumulations for durations up to 3 hours did not reach exceptionally high values anywhere in Belgium (i.e., return period below 30 years). For longer accumulation durations (from 6 hours to 3 days), a clear trend towards extreme values can be noticed with, in parallel, an increase in the size of the severely affected areas. The apparent discontinuity in Figure 8 from 12-hour to 1-day accumulations is related to the temporal granularity of the considered data: maximum accumulations are derived from hourly radar product

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

for duration till 12-hours and from fixed daily accumulations (from 08:00am to 08:00am local time) for durations of 1, 2 or 3 days in order to be consistent with the extreme value statistics (Van de Vyver, 2012). The Hershfield factor (van Montfort, 1990) aimed to adjust the impact of the time series granularity when deriving extreme value statistics is not used to stay as close as possible to the actual statistics.

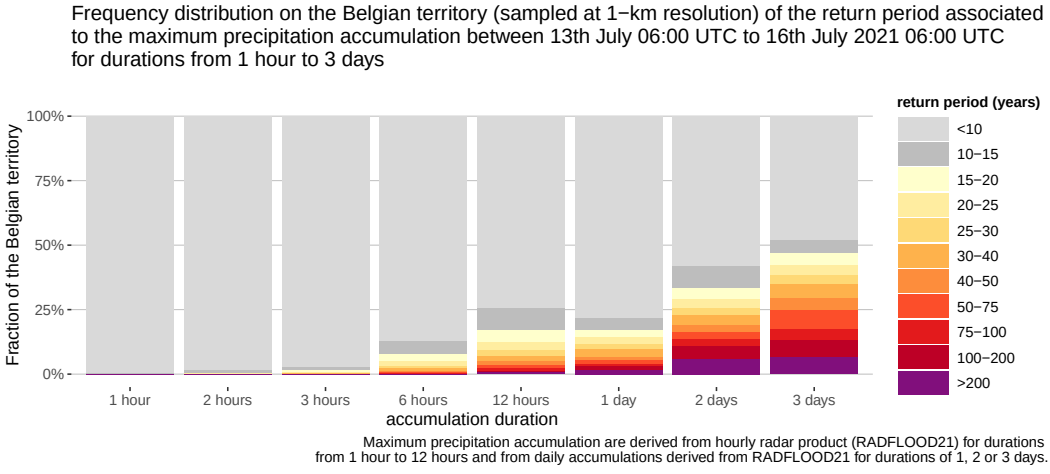

Figure 8. Frequency distribution on the Belgian territory of the return period associated to the maximum precipitation accumulation between 13th July 06:00 UTC to 16th July 2021 06:00 UTC for durations from 1 hour to 3 days. The maximum accumulations are derived from hourly (resp. daily) radar product for durations till 12 hours (resp. of 1, 2 or 3 days). The area of Belgium is 30688 km^2.

3.3 Spatio-temporal analysis of the heavy rainfall event

The rainfall distribution in time illustrated in Figure 6 is limited to the rain gauges with the largest 3-day accumulation, which are all located in a same area (i.e., East of the Province of Liège). Rainfall is however distributed in time differently in other areas of Belgium. The sequence of 5-min radar product (Goudenhoofdt et al., 2023) allows to grasp the spatial-temporal dynamic of the 3-day event. In complement, maps of 3-hourly precipitation totals derived from the 5-min radar product are provided in Figures A2 and A3. It remains however difficult to get an overall insight on the spatial and temporal evolutions of the rainfall patterns during the event from these 5-min or 3-hourly sequences.

Dimensionality reduction methods (Carreira-Perpiñán, 1997; Fodor, 2003; Cunningham, 2008) can be helpful in this context. The goal of dimensionality reduction is to provide a low-dimensional approximation of large data while minimizing the loss of relevant information contained in the data. These approaches aim thus at summarizing the essence of the data in a few representative variables. They provide an approximate but easier to interpret representation of the data. One of the most popular technique for dimensionality reduction is principal component analysis (Jolliffe and Cadima, 2016; Wilks, 2011). In this analysis, because of the non-negative character of precipitation data, we will compute a non-negative matrix factorization (Gillis, 2020) of the 5-min radar product.

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

Given a data matrix V with m rows and n columns (i.e., a matrix of dimension $m \times n$) and with non-negative elements $v_{ij} \geq 0$ (i.e., V is a non-negative matrix), the non-negative matrix factorization (NMF) approximates V as the product of two non-negative matrices W and H of dimensions $m \times k$ and $k \times n$ respectively,

$$V^{(m\times n)} \approx W^{(m\times k)}\, H^{(k\times n)}. \tag{1}$$

The approximation (1) of V can equivalently be written with vectors instead of matrices by

$$V^{(m\times n)} \approx \sum_{i=1}^{k} \mathbf{w}_i^{(m\times 1)}\, \mathbf{h}_i^{(1\times n)}, \tag{2}$$

where $\mathbf{w}_i$ (resp. $\mathbf{h}_i$) of dimension $m \times 1$ (resp. $1 \times n$) denotes the ith column of W (resp. the ith row of H). Equation (2) formulates the approximation of V as the sum of k matrices of rank 1 (i.e., the matrix product of 2 vectors), which are called *components* or *features*. In practice, k is set much smaller than the dimensions m and n in order to decompose V into a limited number of components expected to represent meaningful characteristics of the original data. NMF is generally solved as an optimization problem, where an objective function characterizing the difference between V and matrix product WH has to be minimized. Various algorithms for NMF have been proposed in the literature. In this analysis, we used the algorithm rNMF (Sun et al., 2013) available as an R package.

In order to analyze the 5-min rainfall by NMF, the data need first to be structured in a non-negative $m \times n$ matrix V where each element v_{ij} contains the radar value of the pixel i for the timestamp j. This analysis is performed for the 1-km radar pixels located in Belgium (i.e., $m = 30669$) and the period July 13 00:00 UTC to July 16 12:00 UTC (i.e., $n = 1009$). The result of the NMF approximation of this data into 3 components is illustrated in Figure 9. Each component i is composed of a vector $\mathbf{w}_i$ with values for the m pixels (i.e., spatial pattern of the component i) and a vector $\mathbf{h}_i$ with values for the n timestamps (i.e., temporal pattern of the component i).

These results show that NMF provides a decomposition of the rainfall data into 3 factors with distinct spatio-temporal characteristics. The first factor corresponds to stratiform precipitation in the South East part of Belgium that started already in the morning of July 13 and lasted rather continuously till the end of July 14. The second factor concerns the central part of Belgium with precipitation in the evening of July 13, followed by a rather dry situation on July 14 and again precipitation the whole July 15. Finally, the third factor extracts the heaviest precipitation pattern located in the Province of Liège. This pattern is characterised by successive periods with very intense convective precipitation that started in the evening of July 13 and lasted till the morning of July 15. The NMF analysis clearly shows that rainfall was distributed in time differently in various areas of the country. In view of temporal distributions that are overlapping for the three factors, it is hard to figure out what is the dynamical source of these patterns. This constitutes an interesting future research, going beyond the scope of the current data description.

The information given in Figure 10 indicates that the rank-3 NMF approximation provides a meaningful summary of the rainfall data. First, Figure 10A quantifies how the 3-day accumulation is approximated by the rank-3 NMF. The approximation error is rather low in most places: between -8% and 16% where the 3-day accumulation exceeds 100mm. There is however a region in central Belgium with 3-day accumulation (around 80mm in the radar product) that is largely underestimated by

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

almost 50mm in the rank-3 NMF. This is a local pattern that is not captured by the rank-3 NMF. In Figure 10B, the correlation between the radar product and the 3 temporal components $\mathbf{h}_i$ is assessed for each pixel. This analysis confirms that each of these temporal components are mostly correlated with the radar product in distinct geographical areas, which corresponds the spatial components $\mathbf{w}_i$. To conclude, the information provided by the 3 spatial components $\mathbf{w}_i$ and 3 temporal components $\mathbf{h}_i$ (i.e., the 3 maps and 3 time series of Figure 9) summarizes well in an easily interpretable way the rainfall data for areas with a 3-day accumulation above 100mm.

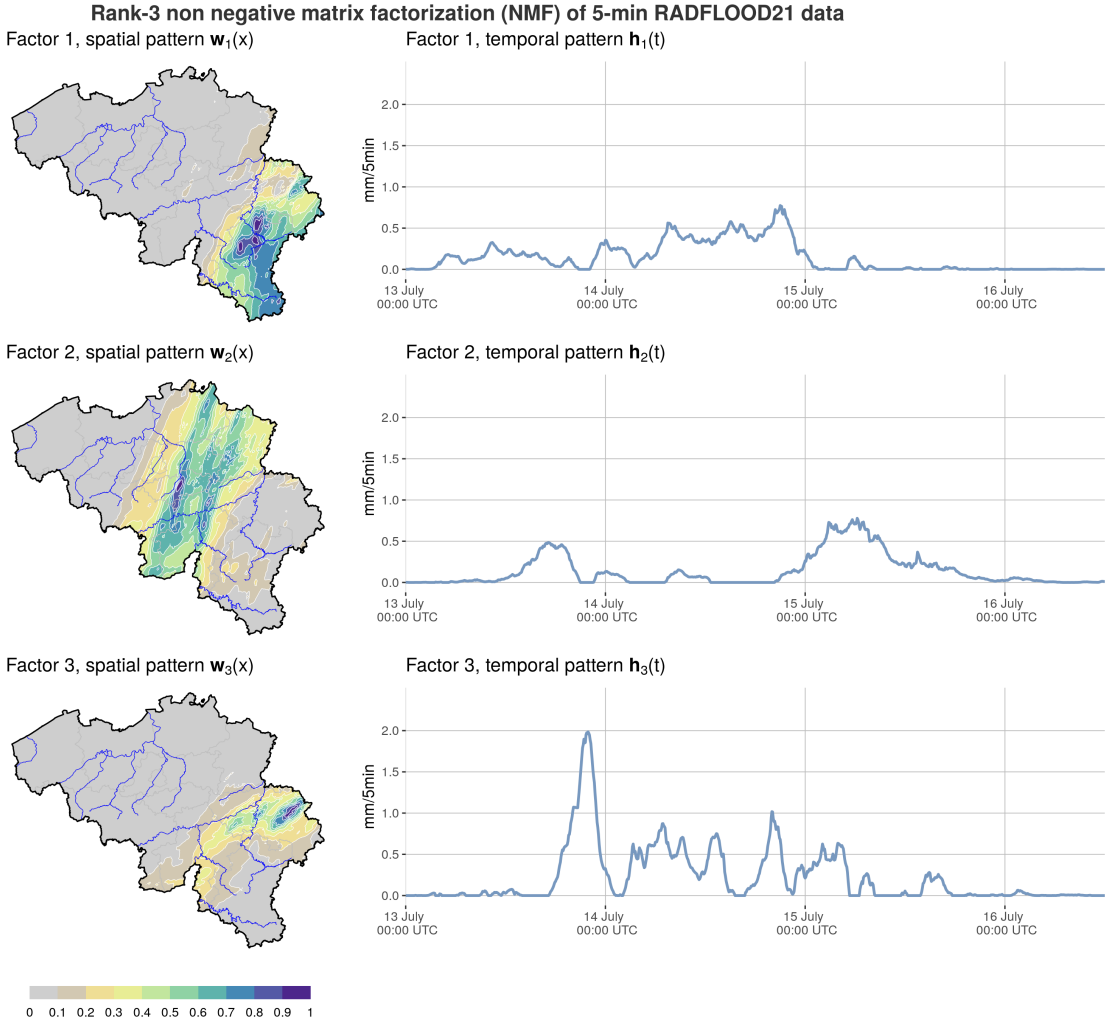

Figure 9. Spatial and temporal patterns of the 3 components resulting from a rank-3 NMF of the 5-min rainfall data.

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

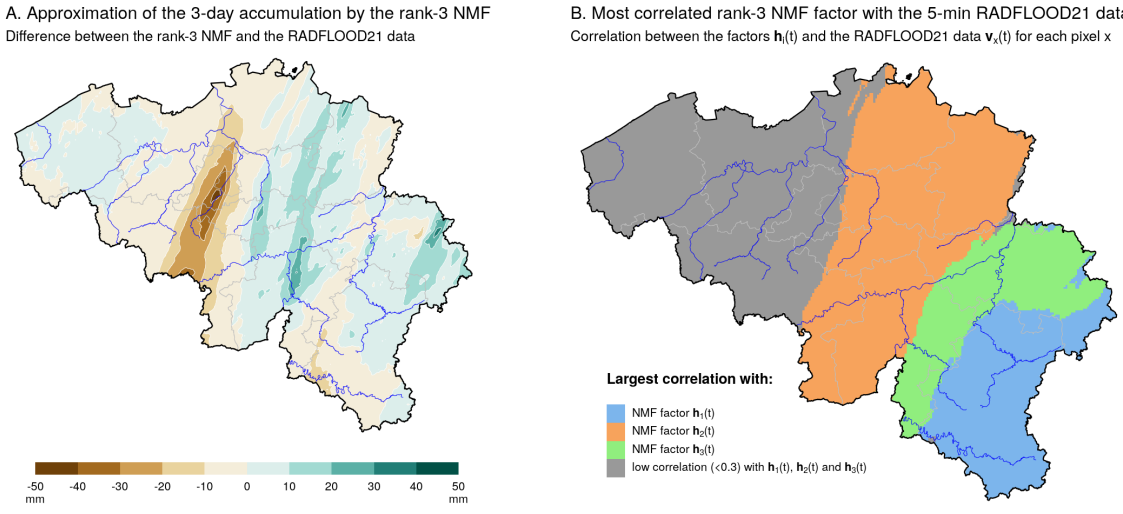

Figure 10. A. Approximation of the 3-day accumulation by the rank-3 NMF of the 5-min RADFLOOD21 data.
B. Result of the correlation analysis between the rainfall data and the 3 temporal components $\mathbf{h}_i$ for each pixel: the map illustrates which component $\mathbf{h}_i$ is the most correlated with each pixel data, provided that the correlation is larger than 0.3.

3.4 Analysis of areal averages

Estimation of areal rainfall averages for catchments may be useful of the hydrological analysis of the event. Such areal averages can easily be derived from the radar product by considering the mean value of all pixels included in the domain of interest. Figure 11 illustrates hourly timeseries of areal rainfall averages for 6 catchments of different size (defined in Figure 12). Areal averages derived from the radar product are compared against an alternative estimation solely based on rain gauges data, i.e., an IDW interpolation of hourly rain gauges data at 1km resolution followed by a spatial aggregation of the interpolation points included in the respective domains. The comparison between both estimations indicates that their difference increases for smaller catchments. Figure 11 indicates also a very large variability in time of hourly rainfall for the Vesdre, Hoëgne and Wayai catchments. This variability in time is significantly smoothed for averages over the large-sized Meuse catchment. By spatially and temporally integrating the radar product, the total precipitation amount that fell over the Belgian part of the Meuse catchment during the 3 days is estimated to 1.527 km^3. Similarly, the total precipitation over the Ourthe and Vesdre catchments is estimated to 0.240 km^3 and 0.125 km^3, respectively.

The maximum accumulation for durations from 1 hour to 72 hours derived from these hourly areal time series is provided in Figure 13 for the 6 considered catchments. These values increase significantly between 6 hours and 48 hours of accumulation duration. The 48-hour totals are estimated to 100mm on average over the large-sized catchment of the Meuse and reach almost 200mm on average for the 2 smallest catchments of the Hoëgne and Wayai.

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

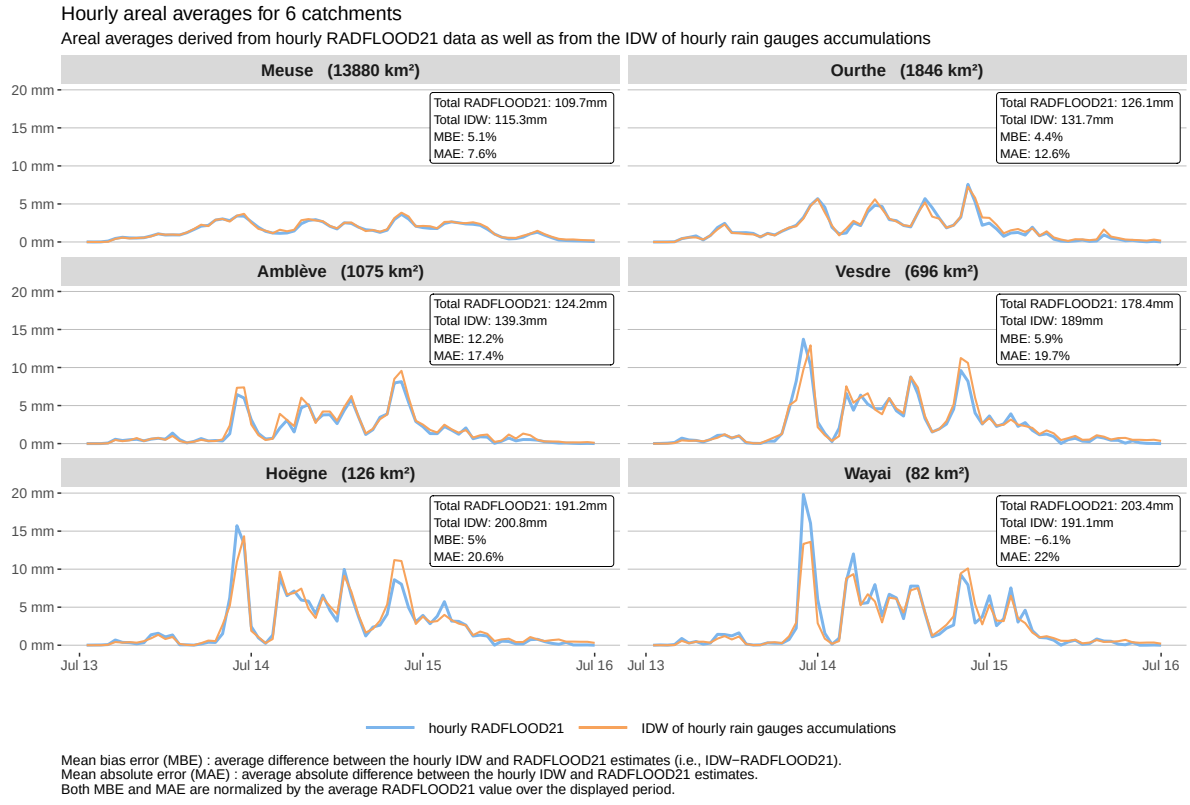

Figure 11. Hourly timeseries of areal rainfall averages for 6 catchments of different size which are mapped in Figure 12. The hourly areal averages are derived from the hourly radar product as well as from hourly rain gauges data (i.e., IDW interpolation of rain gauges data at 1km resolution followed by a spatial aggregation of the interpolation points included in the respective domains). The difference between both estimations is summarized by the mean bias error (MBE) and the mean absolute error (MAE).

4 Conclusions

The rainfall event of July 13 to 16, 2021 over Central Europe affected considerably the eastern part of Belgium and caused disastrous floods in several river catchments. In order to understand the course of the events and the various factors relating rainfall to flood response, it is essential to obtain the best picture of the spatio-temporal distribution of rainfall. In the present work, we present and we analyse quality controlled rain gauge observations and radar-based rainfall products. The radar products include a merging with automatic gauge measurements. These data are freely available and provided as a supplement on Zenodo (Journée et al., 2023; Goudenhoofdt et al., 2023).

The most extreme rainfall were observed in areas with high orography and estimating rainfall from radar observations is particularly challenging for this event. Radar ground echoes are more frequent and intense in areas with high orography which affects the quality of the raw radar data. Orographic enhancement is suspected to have played an important role, which makes

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

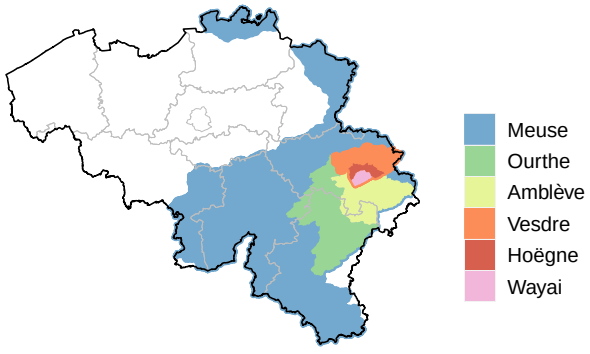

Figure 12. Definition of the 6 catchments considered in Figure 11. Some of these domains are overlapping: the Wayai and Hoëgne catchments are included in the Vesdre catchment; the Vesdre, Ourthe and Amblève catchments are included in the Meuse catchment. The Meuse catchment is here limited to its Belgian part.

very important to use radar observations as close to the ground as possible. A proper treatment of ground echoes was therefore essential for this case.

In some areas, spatial structures at fine scale present in the radar-based products differ substantially from those obtained by classical interpolation of rain gauges. Verification shows that errors made by the radar-based product are considerably smaller than those obtained from rain gauge interpolation. For small catchments, incorporating radar observations is required for capturing the temporal evolution of rainfall.

A generic statistical analysis of these data is also provided allowing for characterizing the return period of such an event and the spatial structure of the precipitation field. This analysis shows that the event of mid-July 2021 is a record-breaking event over most of the eastern part of Belgium with return periods largely exceeding 50 years. The 200-year return level is exceeded over 2000 km^2. The most extreme rainfall over 2 days is almost twice the 200-year return level. The analysis of the spatial structure of the rainfall field suggests that there are three dominant spatial structures over Belgium. As these are not well separated in time, additional analyses are needed in order to relate these features to physical processes. Regarding total precipitation amounts integrated over a region during the event, quantities of 0.484 km^3 and 1.527 km^3 are estimated for the Province of Liège and for the Belgian part of the Meuse catchment, respectively.

It is also interesting to note that the return periods are relatively moderate for shorter durations than 6 hours. Rainfall temporal evolution can not be considered as typical of a flash flood event. In contrast, the response of the system appears, according to several testimonies, as very close to flash flood with surprise effects at various levels (Van Camp et al., 2022; Dewals et al., 2021). This points out the high non-linearity of the system and the fact that various other catchment-specific parameters like topography, land use or soil properties played an important role. Additional processes like the transport of debris by fast-flowing water need to be considered to understand the relation between precipitation and flood response. The presence of dams in some valleys affected by floods should be also taken into account.

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

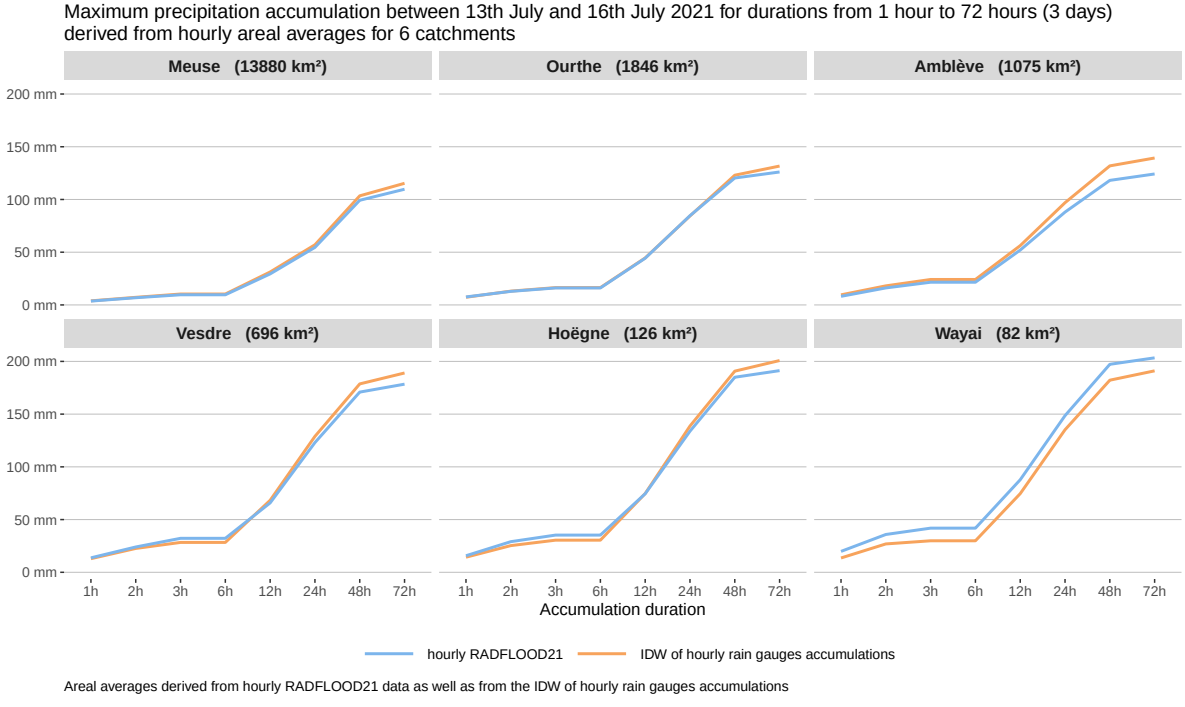

Figure 13. Maximum precipitation accumulation for durations from 1 hour to 72 hours (3 days) derived from hourly timeseries of areal rainfall averages for 6 catchments of different size which are mapped in Figure 12. The hourly areal averages are derived from the hourly radar product as well as from hourly rain gauges data (i.e., IDW interpolation of rain gauges data at 1km resolution followed by a spatial aggregation of the interpolation points included in the respective domains).

The data sets presented here have been produced and quality-controlled with the utmost care. Nevertheless, some additional work can be performed to further refine rainfall estimates. The radar-based rainfall product used in the present study is based on single-polarization radar information only. The exploitation of polarimetric information available from the dual-polarization radars covering the area of interest could bring some additional benefits. Extending the area of interest, incorporating additional information from radars and rain gauge networks available in neighbouring countries, could be realized within an international effort for producing the best picture of this event over the full area affected by floods.

The usage of high-quality rainfall data of extreme events is of great interest interest in many field. Besides the question of the impact of climate change on the occurrence of such events, many questions arise related to water resource and flood risk management or land use planning. Understanding the full causality chain from rainfall to floods and further to the resulting impact in terms of casualties and economic losses is a challenging multidisciplinary enterprise. The present work is a contribution to this global effort.

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

Data availability. The discussed rain gauges data and radar product RADFLOOD21 can be accessed on Zenodo (Journée et al., 2023; Goudenhoofdt et al., 2023).

Author contributions. MJ extracted and verified the rain gauges data; EG developed the radar product RADFLOOD21; MJ analyzed the data; MJ, EG, SV and LD wrote the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This study is based in part on recordings from rain gauges operated by the following Belgian regional hydrological services: Service Public de Wallonie - Mobilité et Infrastructures (https://hydrometrie.wallonie.be), Vlaamse Milieumaatschappij (https://www.waterinfo.be/) and Hydrologisch Informatiecentrum (https://www.waterinfo.be/).

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

References

Amponsah, W., Ayral, P.-A., Boudevillain, B., Bouvier, C., Braud, I., Brunet, P., Delrieu, G., Didon-Lescot, J.-F., Gaume, E., Lebouc, L., Marchi, L., Marra, F., Morin, E., Nord, G., Payrastre, O., Zoccatelli, D., and Borga, M.: Integrated high-resolution dataset of high-intensity European and Mediterranean flash floods, Earth System Science Data, 10, 1783–1794, https://doi.org/10.5194/essd-10-1783-2018, 2018.

Assuralia: Actualisation relative aux inondations de juillet 2021, https://press.assuralia.be/actualisation-relative-aux-inondations-de-juillet-2021, last access: 3 February 2023, 2022.

Brauer, C. C., Teuling, A. J., Overeem, A., van der Velde, Y., Hazenberg, P., Warmerdam, P. M. M., and Uijlenhoet, R.: Anatomy of extraordinary rainfall and flash flood in a Dutch lowland catchment, Hydrology and Earth System Sciences, 15, 1991–2005, https://doi.org/10.5194/hess-15-1991-2011, 2011.

Carreira-Perpiñán, M.: A Review of Dimension Reduction Techniques, Technical Report CS-96-09, Department of Computer Science, University of Sheffield, Sheffield, UK, 09, 1997.

Cunningham, P.: Dimension Reduction, pp. 91–112, https://doi.org/10.1007/978-3-540-75171-7_4, 2008.

Delrieu, G., Nicol, J., Yates, E., Kirstetter, P.-E., Creutin, J.-D., Anquetin, S., Obled, C., Saulnier, G.-M., Ducrocq, V., Gaume, E., Payrastre, O., Andrieu, H., Ayral, P.-A., Bouvier, C., Neppel, L., Livet, M., Lang, M., du Châtelet, J. P., Walpersdorf, A., and Wobrock, W.: The Catastrophic Flash-Flood Event of 8–9 September 2002 in the Gard Region, France: A First Case Study for the Cévennes–Vivarais Mediterranean Hydrometeorological Observatory, Journal of Hydrometeorology, 6, 34 – 52, https://doi.org/https://doi.org/10.1175/JHM-400.1, 2005.

Dewals, B., Erpicum, S., Pirotton, M., and Archambeau, P.: The July 2021 extreme floods in the Belgian part of the Meuse, Hydrolink Magazine, https://hdl.handle.net/20.500.11970/109527, 2021.

Fodor, I.: A survey of dimension reduction techniques. Tech. Rep. UCRL-ID-148494, https://doi.org/10.2172/15002155, 2003.

Gillis, N.: Nonnegative Matrix Factorization, Society for Industrial and Applied Mathematics, Philadelphia, PA, https://doi.org/10.1137/1.9781611976410, 2020.

Goudenhoofdt, E. and Delobbe, L.: Generation and Verification of Rainfall Estimates from 10-Yr Volumetric Weather Radar Measurements, Journal of Hydrometeorology, 17, 1223–1242, https://doi.org/10.1175/jhm-d-15-0166.1, 2016.

Goudenhoofdt, E., Journée, M., and Delobbe, L.: Observational rainfall data of the 2021 mid-July flood event in Belgium – Part 2. Radar product RADFLOOD21, https://doi.org/10.5281/zenodo.7740059, 2023.

Gouvernement Wallon: Inondations de juillet 2021 : bilan et perspectives, https://www.wallonie.be/fr/actualites/inondations-de-juillet-2021-bilan-et-perspectives, last access: 1 February 2023, 2022.

IPCC: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, vol. In Press, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/9781009157896, 2021.

Jolliffe, I. and Cadima, J.: Principal component analysis: A review and recent developments, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374, 20150 202, https://doi.org/10.1098/rsta.2015.0202, 2016.

Journée, M., Goudenhoofdt, E., and Delobbe, L.: Observational rainfall data of the 2021 mid-July flood event in Belgium – Part 1. Rain gauges observations, https://doi.org/10.5281/zenodo.7739983, 2023.

Kreienkamp, F., Philip, S. Y., Tradowsky, J. S., Kew, S. F., Lorenz, P., Arrighi, J., Belleflamme, A., Bettmann, T., Caluwaerts, S., Chan, S. C., Ciavarella, A., De Cruz, L., de Vries, H., Demuth, N., Ferrone, A., Fischer, r. M., Fowler, H. J., Goergen, K., Heinrich, D., Henrichs, Y.,

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

Lenderink, G., Kaspar, F., Nilson, E., Otto, F. E. L., Ragone, F., Seneviratne, S. I., Singh, R. K., Skålevåg, A., Termonia, P., Thalheimer, L., van Aalst, M., Van den Bergh, J., Van de Vyver, H., Vannitsem, S., van Oldenborgh, G. J., Van Schaeybroeck, B., Vautard, R., Vonk, D., and Wanders, N.: Rapid attribution of heavy rainfall events leading to the severe flooding in Western Europe during July 2021, 2021.

Ludwig, P., Ehmele, F., Franca, M. J., Mohr, S., Caldas-Alvarez, A., Daniell, J. E., Ehret, U., Feldmann, H., Hundhausen, M., Knippertz, P., Küpfer, K., Kunz, M., Mühr, B., Pinto, J. G., Quinting, J., Schäfer, A. M., Seidel, F., and Wisotzky, C.: A multi-disciplinary analysis of the exceptional flood event of July 2021 in central Europe. Part 2: Historical context and relation to climate change, Natural Hazards and Earth System Sciences Discussions, 2022, 1–42, https://doi.org/10.5194/nhess-2022-225, 2022.

Marchi, L., Borga, M., Preciso, E., and Gaume, E.: Characterisation of Selected Extreme Flash Floods in Europe and Implications for Flood Risk Management, Journal of Hydrology, 394, 118–133, https://doi.org/10.1016/j.jhydrol.2010.07.017, 2010.

Meyer, J., Neuper, M., Mathias, L., Zehe, E., and Pfister, L.: Atmospheric conditions favouring extreme precipitation and flash floods in temperate regions of Europe, Hydrology and Earth System Sciences, 26, 6163–6183, https://doi.org/10.5194/hess-26-6163-2022, 2022.

Mohr, S., Ehret, U., Kunz, M., Ludwig, P., Caldas-Alvarez, A., Daniell, J. E., Ehmele, F., Feldmann, H., Franca, M. J., Gattke, C., Hundhausen, M., Knippertz, P., Küpfer, K., Mühr, B., Pinto, J. G., Quinting, J., Schäfer, A. M., Scheibel, M., Seidel, F., and Wisotzky, C.: A multi-disciplinary analysis of the exceptional flood event of July 2021 in central Europe. Part 1: Event description and analysis, Natural Hazards and Earth System Sciences Discussions, 2022, 1–44, https://doi.org/10.5194/nhess-2022-137, 2022.

MunichRe: Hurricanes, cold waves, tornadoes: Weather disasters in USA dominate natural disaster losses in 2021. Europe: Extreme flash floods with record losses, https://www.munichre.com/en/company/media-relations/media-information-and-corporate-news/media-information/2022/natural-disaster-losses-2021.html, last access: 20 December 2022, 2022.

Ruiz-Villanueva, V., Borga, M., Zoccatelli, D., Marchi, L., Gaume, E., and Ehret, U.: Extreme flood response to short-duration convective rainfall in South-West Germany, Hydrology and Earth System Sciences, 16, 1543–1559, https://doi.org/10.5194/hess-16-1543-2012, 2012.

Sun, J., Xu, Y., Lopiano, K., and Young, S.: Robust Non-Negative Matrix Factorization Procedures for Analyzing Tumor and Face Image Data, Proceedings of the Joint Statistical Meetings, 2013.

Van Camp, M., de Viron, O., Dassargues, A., Delobbe, L., Chanard, K., and Gobron, K.: Extreme Hydrometeorological Events, a Challenge for Gravimetric and Seismology Networks, Earth's Future, 10, e2022EF002 737, https://doi.org/https://doi.org/10.1029/2022EF002737, e2022EF002737 2022EF002737, 2022.

Van de Vyver, H.: Spatial regression models for extreme precipitation in Belgium, Water Resources Research, 48, https://doi.org/10.1029/2011wr011707, 2012.

Van de Vyver, H.: Practical return level mapping for extreme precipitation in Belgium, RMI scientific and technical publication, 62, 2013.

van Montfort, M.: Sliding maxima., Journal of Hydrology, 118, 77–85, https://doi.org/10.1016/0022-1694(90)90251-r, 1990.

Vergara-Temprado, J., Ban, N., Ban, N., and Schär, C.: Extreme Sub-Hourly Precipitation Intensities Scale Close to the Clausius-Clapeyron Rate Over Europe, Geophysical Research Letters, https://doi.org/10.1029/2020gl089506, 2021.

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, vol. 100, Elsevier Academic Press, Amsterdam; Boston, 2011.

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

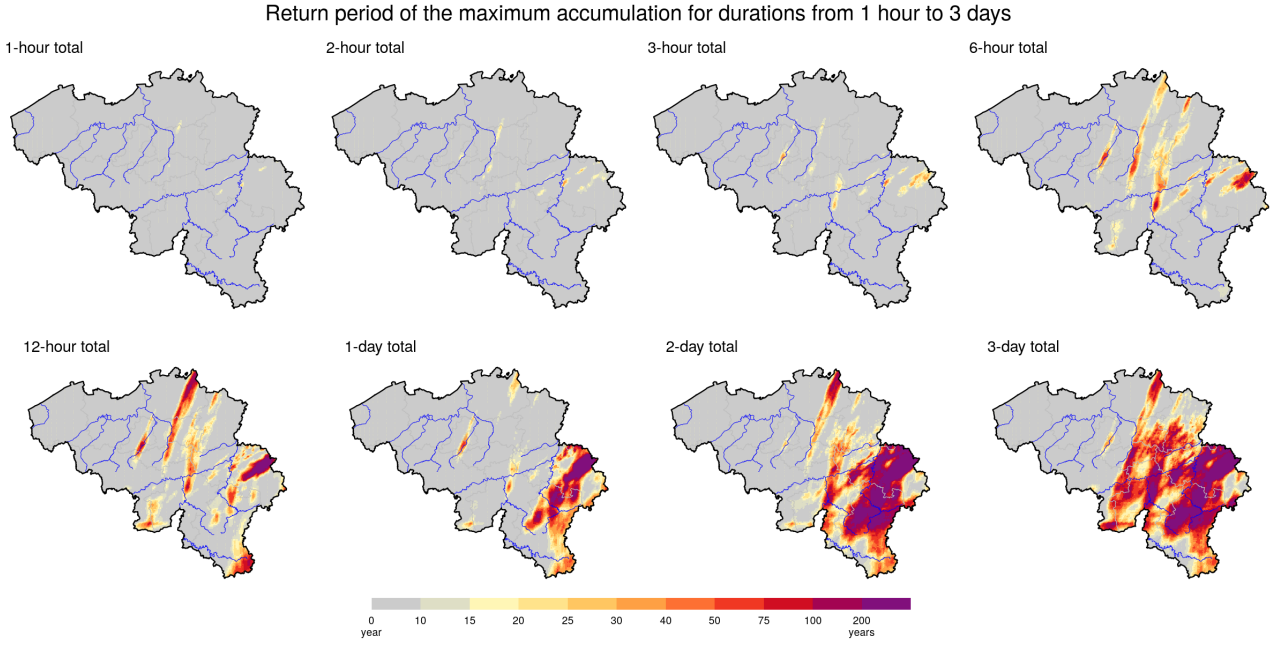

Figure A1. Return period estimated for the maximum precipitation accumulation between 13th July 06:00 UTC to 16th July 2021 06:00 UTC for durations from 1 hour to 3 days. The maximum accumulations are derived from hourly (resp. daily) radar product data for durations till 12 hours (resp. of 1, 2 or 3 days).

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

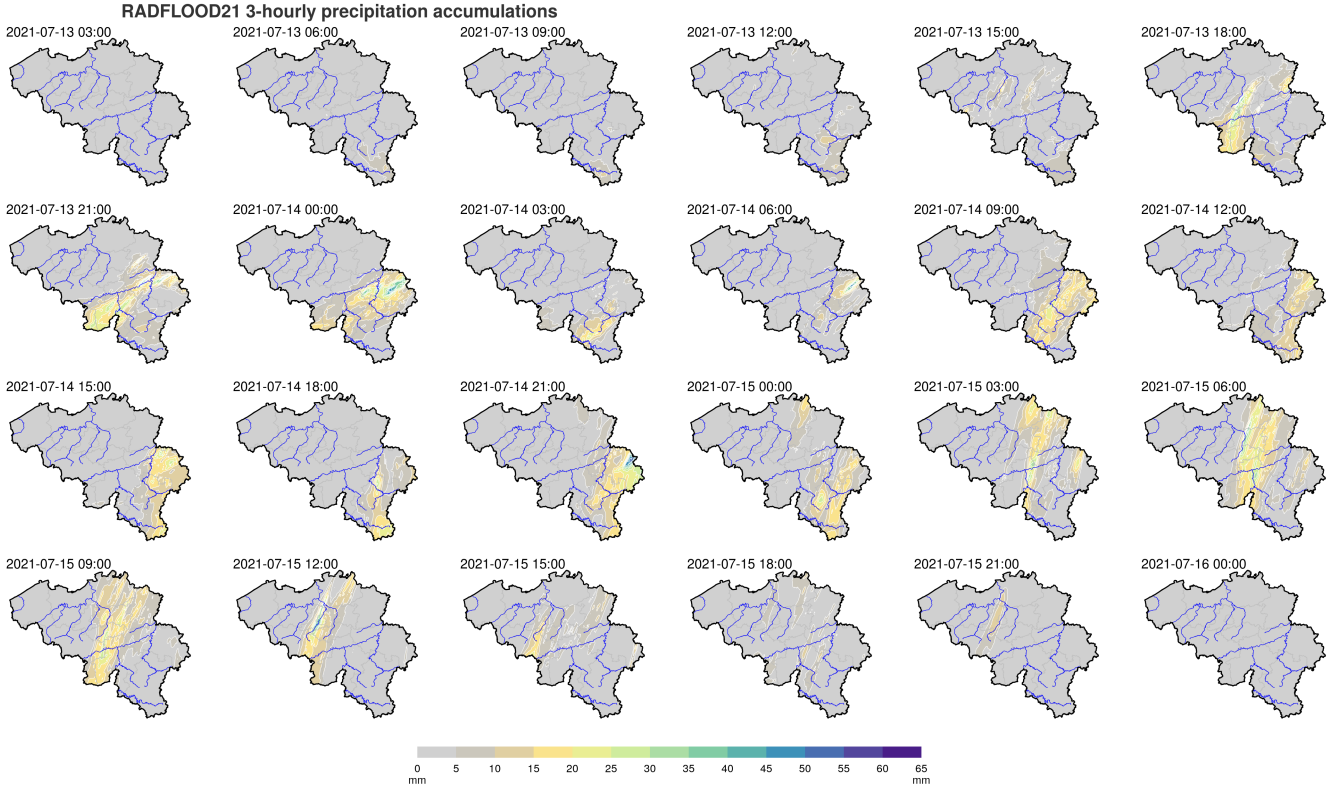

Figure A2. 3-hourly precipitation accumulations in Belgium derived from 5-min radar product data. The timestamp on top of each map indicates the end of the 3-hour period in UTC.

https://doi.org/10.5194/egusphere-2023-542
Preprint. Discussion started: 30 March 2023

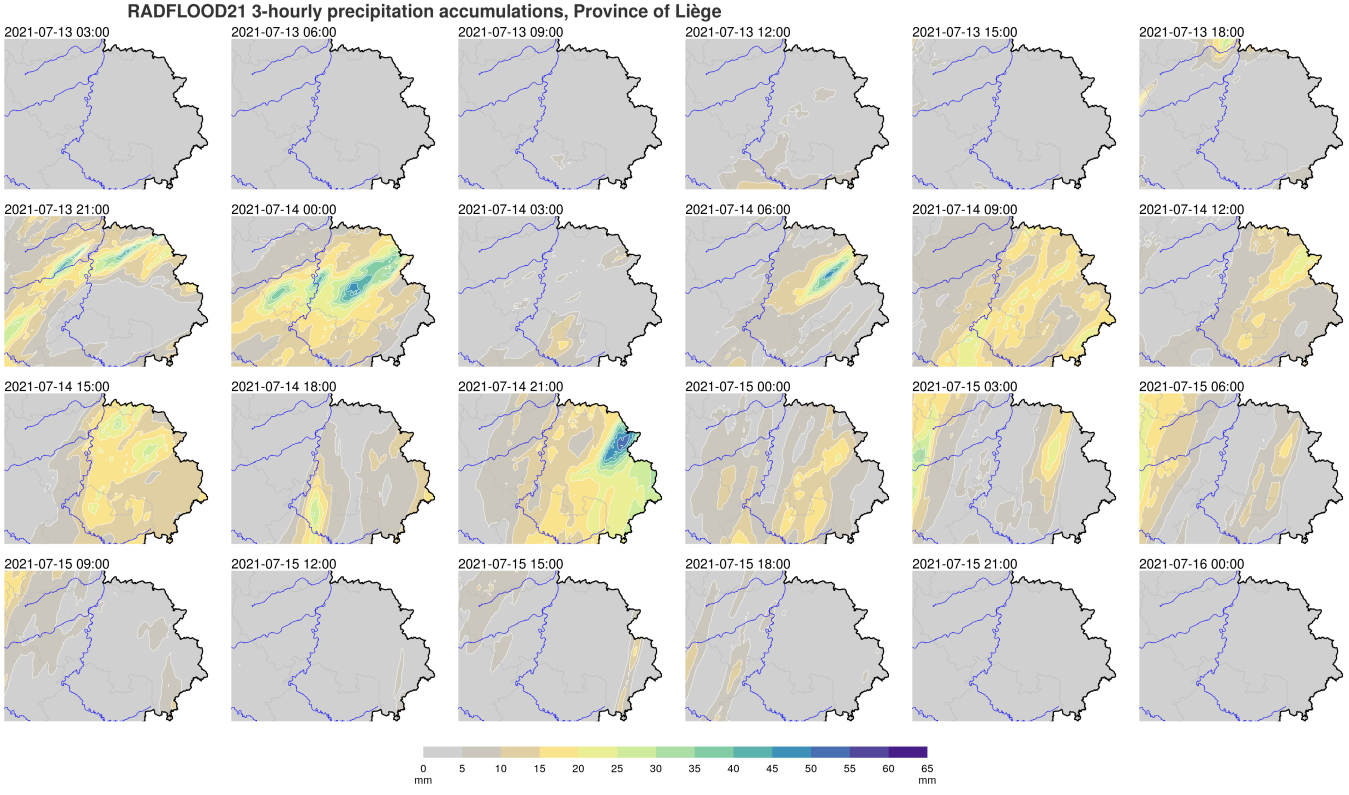

Figure A3. 3-hourly precipitation accumulations in the Province of Liège derived from 5-min radar product data. The timestamp on top of each map indicates the end of the 3-hour period in UTC.