Review 1

Review for „Hourly precipitation fields at 1 km resolution over Belgium from 1940 to 2016 based on the analog technique" submitted to Earth System Science Data by Elke Debrie et al.

General comments:
This publication presents an analogue method to determine hourly rainfall patterns over Belgium for a 80-year period based on radar observation.

This method shows good potential for gridded precipitation fields over this period, but my main concern is that the training dataset is quite short. In addition, it is at the end of the time period (1940-now). Within this time period, climate change caused an increase of mean temperatures by about 1.5°C in central and Western Europe which also caused a change in circulation patterns. Please discuss this in the paper.

Thank you very for drawing our attention to these important problems. Indeed climate change may affect the analog selection. On one hand, climate change may affect the intensity of precipitation, and on the other hand, the circulation patterns (their frequency or the emergence of new patterns) that lead to specific precipitation fields. Our analysis indeed assumes that the precipitation intensity is not strongly modified and that no new emerging circulation patterns are formed yet. Concerning the former, our time series suggests that there is no substantial trend in the intensity at short time scale (1 hour, 1 day). Concerning the latter, we did not see any report of the emergence of new patterns, but well of the possible change of frequency which is naturally taken into account by the selection of analogs, i.e. only analogs of the specific circulation pattern are in principle selected.

As you will see below in the text we underline the absence of substantial trends in extreme precipitation due to climate change, and we discuss the question of the modification of circulation patterns that could affect the frequency of events. Both points are now discussed in the text as follows in Section 5:

"We acknowledge that the training dataset, which covers only a six-year period, is relatively short. It is a sophisticated merging of the radar datasets and the rain gauges data \citep{journee2023}. Moreover, a coherent coverage with the same radars started only in 2017. Efforts are currently ongoing at the Royal Meteorological Institute to extend this dataset for the following years and also to extract the best information from previous years.

The risk of misrepresenting climate change is an important problem that could affect the results. The question is twofold: On one hand the change of precipitation intensity and on the other hand, the change of circulation patterns. For the precipitation intensity, climate change has currently a very limited impact over Belgium as reflected by the long time series covering Belgium \citep{RMI1}. In the Uccle time series of 1-hour or 24-hour extremes, there is no obvious trend (as for temperature for instance), except some long term modulation probably associated with low-frequency variability that still needs to be uncovered. In the recent attribution study of the extreme July 2021 case which affected the East of Belgium it was *shown that there is no local trends of one-day or two-day precipitation accumulations, except the extreme July event, which appears like a black swan (Tradowsky et al, 2023). We therefore suspect that climate change does not affect much (currently) the extreme intensities. Therefore if there are changes in precipitation it should mainly be related to changes in circulation as illustrated on the website of the Royal Meteorological Institute (\cite{RMI2}, \cite{RMI3}). The analog approach should partly take that into account as the reconstruction of precipitation is done using analogs of weather situations (regimes) that were experienced in the past. So if the circulation regimes (but not necessarily their frequency) were the same in the training dataset and in the past, the method should appropriately take the change of frequency into account. As far as we can experience, there is no major disruption on the weather patterns (large scale pressure fields) yet that would imply that the analog approach cannot provide useful information on past weather situations. This question of change of circulation regimes is however important to address in the future due to the rapid changes experienced by our climate. Such an analysis will then be used to support further developments of the analog dataset."*

Why did you use 2006 as a target year for evaluation? It was a particular year with very different weather conditions (cold winter, hot and dry June/July, cold and wet August, very mild autumn). It would be interesting to compare the analogues for more than one year in section 3.2.1/3.2.2. Do the results (esp. about differences) change?

Thank you for raising this point. The results are in fact robust among the different years, and we selected 2006 as an example for displaying the results as the number of automatic stations have increased at that moment. Moreover this year was relatively wet. The following paragraphs are added to Section 3.2:

"The year 2006 was selected as the case study year because from that moment onward, a considerable number of automatic weather stations became available, and the year was relatively wet. While more 'average' years such as 2007, 2012 or 2014 could have been considered, our focus on evaluating precipitation fields led us to prefer a relatively wet year.

We also examined in details additional years after 2006. Overview figures for evolution of geopotential height for 2011 (drier year) and 2012 (wetter year) are added in the supplement to show the robustness of the results (see Figures S1, S2 and S3, and also the figures attached below). We found that the evolution of geopotential height remains consistent between the real days and the analog days identified using the 4TWS predictor set. The overall statistical properties of weather patterns are preserved across the different years, suggesting that the method reliably captures the key dynamical features and precipitation patterns."

Specific comments:

Section 2.1.2: The RADCLIM product is based on a combination of radar data merged with ground-based precipitation observations? This is not entirely clearly written here.

Thank you for raising this point. The text about RADCLIM has now been expanded as follows in Section 2.1.2:

“The RADCLIM product is an offline counterpart to the real-time RADQPE (RADar Quantitative Precipitation Estimation) system \citep{ GenerationandVerificationofRainfallEstimatesfrom10YrVolumetricWeatherRadarMeasurements}. It reconstructs historical precipitation datasets using the same core methodology as RADQPE, but adapts it for retrospective analysis by incorporating missing radar or rain gauge measurements that were unavailable during real-time operations. This approach addresses the gaps and uncertainties inherent to instantaneous processing, resulting in a more complete and reliable dataset suited for climatological analyses.

RADCLIM provides high resolution radar-based quantitative precipitation estimation for Belgium and its surroundings. The RADCLIM product is obtained after a careful processing of the weather radar measurements and a merging with rain gauge measurements. This combination provides a detailed and accurate picture of precipitation distribution in time and space.

Data is collected through a network of weather radars and rain gauges. RADCLIM incorporates 3D reflectivity data from a network of four C-band dual-polarization weather radars: Jabbeke and Wideumont, owned and operated by RMI; Helchteren, managed by the Flanders Environment Agency (VMM); and Avesnois, operated by Meteo-France. The rain gauge data come from automatic networks of the Flemish Environment Agency (VMM), the Service Public de Wallonie (SPW), the Hydraulic Laboratory (WL), the Royal Meteorological Institute (RMI) and from manual rain gauges network CLIM of RMI. Both automatic and manual quality checks are carried out to ensure data reliability.

Radar reflectivity measurements are converted into rain rates. To get from raw radar reflections to reliable precipitation estimates, several processing steps are carried out. To mitigate non-meteorological echoes (e.g., ground clutter, wind farms), satellite cloudiness comparisons, vertical reflectivity profile analyses, and spatial texture detection are employed. The data from the different radars are then merged into a precipitation composite. Precipitation accumulations are calculated over 5-minute and 1-hour intervals using optical flow techniques. These data are combined with rain gauge observations every 5 minutes via a statistical interpolation method based on the hypothesis of Gaussian process. This creates a correction factor, which is then also applied to the 5-minute accumulations.

The final composite product has a spatial resolution of 1 km and uses the Belgian Lambert 2008 projection. Coverage extends from 0.3°W to 9.7°E and from 47.4°N to 53.7°N. Precipitation accumulations for 5 minutes and 1 hour are available based on UTC time and are provided in standard formats such as GeoTIFF and HDF5.“

Section 3, lines 108-112: How did you choose these parameters, especially concerning (4) – why did you use the geopotential at each level at different times of the day? Do you have a physical reasoning for this? Did you try other combinations?

Thank you for your feedback. To clarify this aspect, we changed the text in the beginning of Section 3 as follows:

“The analog technique used in this study is based on the principle that local weather conditions are influenced by large-scale atmospheric patterns. According to the weather analog hypothesis, if two synoptic-scale situations are similar, they are likely to produce similar local weather outcomes. By comparing a given target day with past days that had similar atmospheric conditions, we aim to estimate local variables, such as precipitation. This approach assumes that similar large-scale atmospheric conditions lead to similar local precipitation patterns. While this assumption may not always hold perfectly, it provides a solid foundation for developing a high-resolution gridded precipitation dataset.

To identify analog days, we tested various combinations of atmospheric variables and time steps. The selection of predictors and time steps was guided by insights from previous research on analog methods. These previous studies have consistently shown that geopotential height is among the most effective predictors for identifying analog days, which is why it was prioritized from the outset \citep{horton2017, horton2019}. The results confirmed that sets including geopotential height consistently outperformed those without it.

Due to the computational burden of selecting analogs and validation when testing multiple variables across many pressure levels and time steps, only a limited selection of predictor sets was addressed in line with the previous literature. Therefore, we made informed, targeted selections to ensure feasibility, guided by a combination of trial-and-error and the findings from previous studies. Several combinations of fields and times of day were tested, and we ultimately arrived at a scheme that provides good results. However, we acknowledge that other combinations might perform even better, and further work is needed to optimize the method. These improvements are planned for the next version of the dataset.”

Figure 3: Which unit does the “mean RMS distance” have? Precipitation (mm per day/hour)? Please explain the meaning of these numbers!

The unit – mm/24h – is added to the plots and in the text. A more clear explanation about the meaning of the mean RMSE is now added to Section 3.1:

“For every day, the analog precipitation field and the target precipitation field are compared and for every pair of fields the following measure was calculated: \sqrt{\frac{1}{n_s}\sum_s(p_{rs} - p_{as})^2}, with p_r the precipitation amount in mm/24h on the real day, p_a the precipitation amount in mm/24h on the analog day, s the AWS and n_s the number of stations. This measure is calculated for every single day and was then averaged over months, seasons and the whole year to make the comparison between the different methods more convenient. This average RMSE with mm/24h as unit is displayed on the y-axis.”

Figure 4, panel c: You write: “differences in geopotential height on analog days, real days and random days”. What is the reference of the differences here? Is it compared to the smoothed yellow line?

The term ‘differences’ was explained more clearly in Section 3.2.1:

“The term 'difference' refers to the daily consecutive changes in geopotential height. Specifically, for the entire year, the change in geopotential height between 12:00 UTC on

day x and 12:00 UTC on day x+1 is calculated. This provides an indication of the temporal consistency of the geopotential height on consecutive analog days."

Technical comments

Line 201: I couldn't find the reference "Wil, 2019"

Thank you for pointing that out. The reference was not made well in the LaTeX document. It should be fine now.

Figures 9-11: To me, it would make sense to merge these plots (9a+10a+11a in one plot with three lines, and correspondingly for the others) Also Figures 12 a-c should be merged into one plot

Thank you very much for the suggestion. Now, panels a, b and c from Figures 9, 10, 11 and 12 were merged into one plot in Figure 9 till Figure 12 in Section 3.2.3

Additional figures added to the supplement:

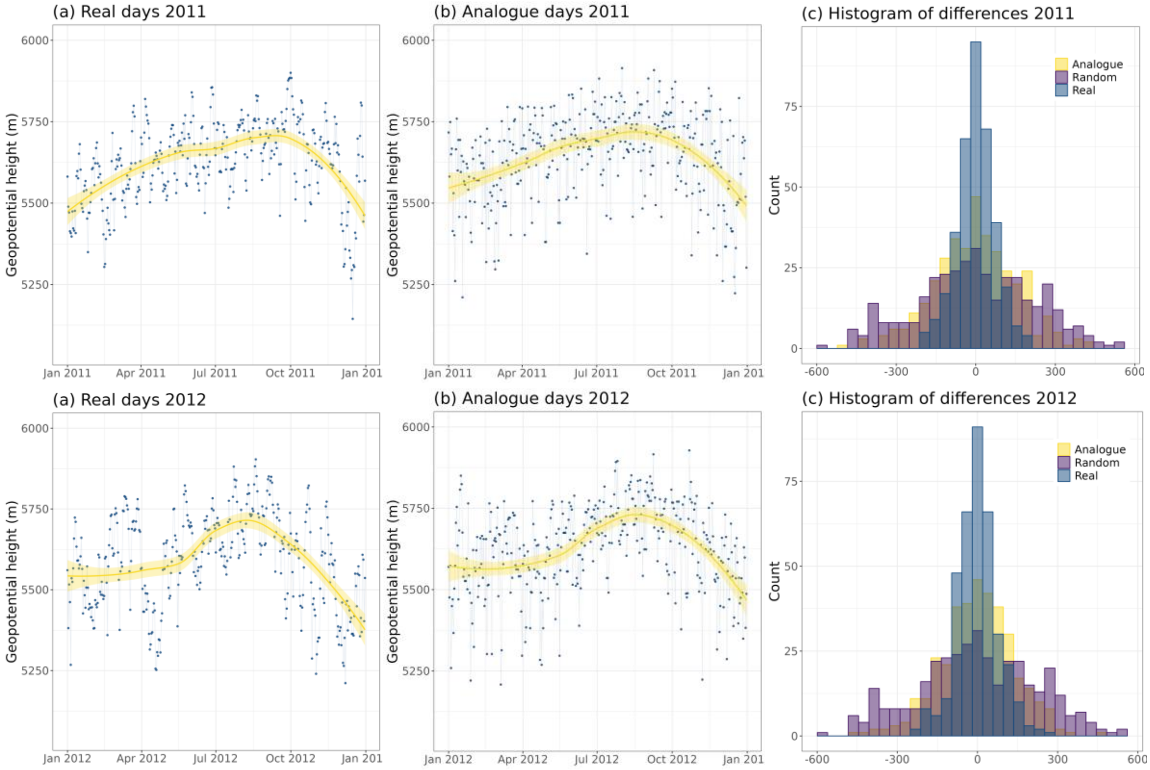

Figure S1. The temporal evolution in geopotential height at the 500 hPa pressure level at 12 UT in Uccle in 2011 resp. 2012 (panel a); the corresponding data on the best analog days in 2011 resp. 2012 (panel b); differences in geopotential height on analog days, real days and random days in 2011 resp. 2012 (panel c). The yellow smoothed curves in the first two plots aid the eye in seeing patterns.

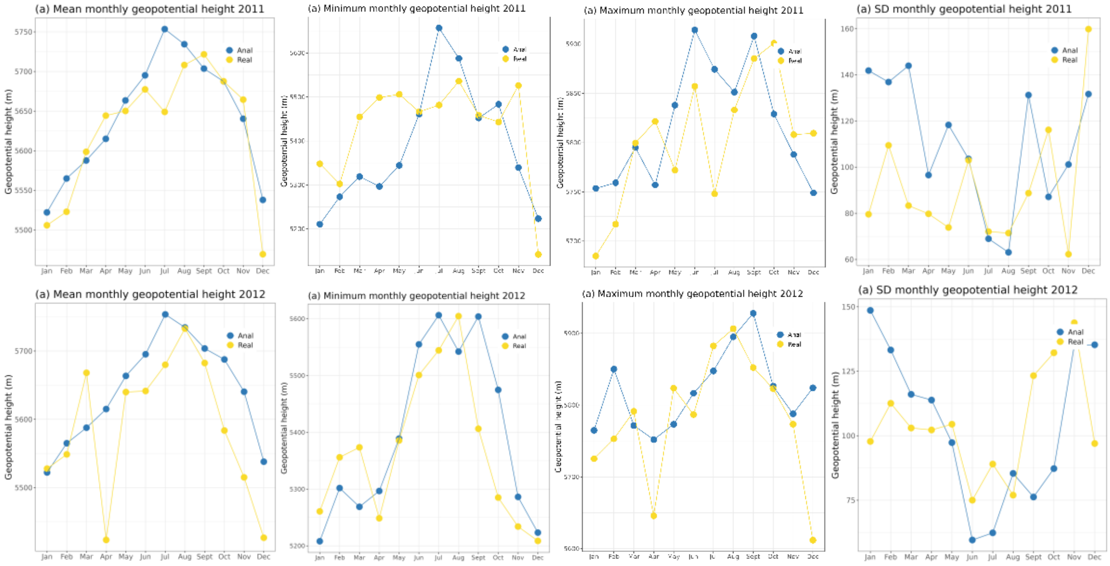

Figure S2. Monthly statistics of geopotential height at 500 hPa pressure level in Uccle at 12 UT in 2011 and 2012. Respectively mean, minimum, maximum and standard deviation are shown in the figures. Consistent with previous observations, the general pattern between real and analog days is comparable across all four summary measures.

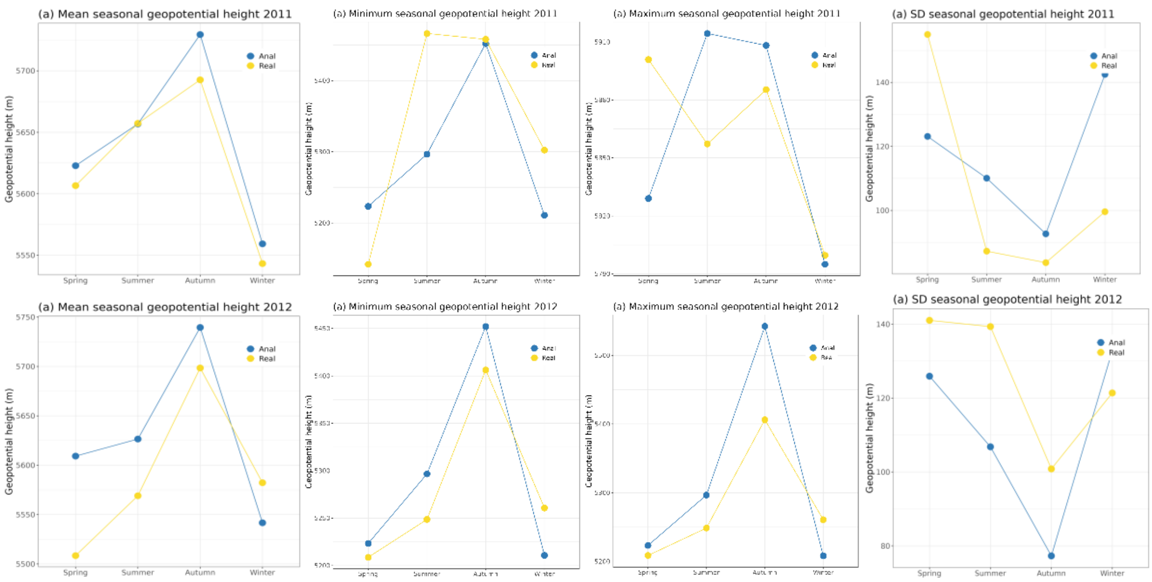

Figure S3. Seasonly statistics of geopotential height at 500 hPa pressure level in Uccle at 12 UT in 2011 and 2012. Respectively mean, minimum, maximum and standard deviation are shown in the figures. The general pattern observed across all four summary measures shows a high degree of similarity between real and analog days, aligning with earlier results.

References

Tradowsky, J. et al, Attribution of the heavy rainfall events leading to the severe flooding in Western Europe during July 2021. *Climatic Change*, 176, 90, https://doi.org/10.1007/s10584-023-03502-7, 2023 .