Spatial Econometrics in R - Paper Review

Comparing spatial metrics of extreme precipitation between data from rain gauges, weather radar and high-resolution climate model re-analyses

Author: Michele Guderzo

Warsaw, Academic Year 2021/22

Contents

1 Introduction

The research paper named *"Comparing spatial metrics of extreme precipitation between data from rain gauges, weather radar and high-resolution climate model re-analyses"*, conducted by Emma Dybro Thomassen et al. in 2021, focuses on the representation of extreme precipitation in south-eastern part of Denmark. The authors compares point and radar observations in order to reanalyse climate model output data for a period of 14 years where there is full spatial and temporal overlap between datasets. To this end, they used several methods such as radar data, spatial correlation in point observations, convective-permitting climate model and coarser convective parameterized model, highlight the strengths and weaknesses of each.

In this paper, my aim is to deepen and analyse the article. I am going to briefly discuss the methods the authors used, giving more attention to the spatial modelling techniques that have been applied, suggest, if required, improvements to the study and debate the results they found. But before doing so, I am going to provide the reader with some background information, in order to give a better understanding about the topic we are talking about.

2 Review

2.1 Background

Hydrology is the science that studies the distribution, movement, biology and chemistry of the planet's water masses. Although knowledge of hydrological cycles is important for human life and activities, hydrology is a relatively young science and no complete theory exists to date. An important characteristic of many hydrological cycle flows is their spatial and temporal variability, which makes them even more difficult to model mathematically. In general, there are two techniques related to probability concepts that have great application in this field: the branch of statistics called geostatistics and the theory of random variables.

More specifically, hydrology is closely linked to the place (city) where it is studied and can vary considerably depending on the location of the city on Earth. To address this relationship, the science of urban hydrology is born. Urban hydrology is the scientific application of hydrologic principles and knowledge to the planning and management of urban areas and their surroundings. It embraces all aspects of the interactions of man and water in occupancy of land. It includes the special hydrologic studies needed to accomplish these ends and deals with minimizing the adverse effects of man's use of land and water and with maximizing the effective use of the available water resources.

In the next section I will look at the literature review the authors dealt with, since they were based on many articles.

2.2 Literature review

The literature review provided many insights and useful information for the authors. It was seen that, according to Berne et al. (2004) and Ochoa-Rodriguez et al. (2015), urban hydrology applications, in general, require a higher spatio-temporal description of extreme precipitation statistics. Arnbjerg-Nielsen et al., 2013; Thorndahl et al., 2017, on the other hand, posed the problem of the precipitation-generating mechanism, which changes significantly whether large or urban scale is considered. The authors went on to emphasise the importance of spatial metrics, in areas such as statistical downscaling of regional climate models (RCMs) (Willems and Vrac, 2011) or in the design of urban hydrological systems (Ochoa-Rodriguez et al., 2015; Thorndahl et al., 2016). Moreover, within the representation of high-resolution extreme precipitation, several spatial correlation metrics have been developed as a function of the temporal resolution (Gregersen et al., 2013; Mayer et al., 2015; Sørup et al., 2016; Sunyer et al., 2013) using climate models constructed with a parametrization scheme for convective rainfall. On the other hand, high-resolution convective permitting climate models, i.e. CPMs, have only recently gained popularity have only recently gained popularity (Kendon et al., 2014). Finally, as reported by Schleiss et al., 2020; Thorndahl et al., 2017, radar data is often compared to point observations for verification and bias correction purposes.

2.3 Study and aim

The authors then proceed to set out the structure of their research, highlighting its objectives. Mainly they calculate different spatial and temporal metrics for extreme precipitation and compare point observations, radar precipitation and high-resolution RCM re-analysis data for the same region and period. Their purpose is to identify if and how much radar data provide extra information compared to point observations and to assess the quality of calculated RCMs models.

2.4 Data and methods

In this section, E.D. Thomassen et al. described their dataset, which mainly consists of four types of spatial measurements: observational point measurements from tipping bucket rain (POINT), bias-corrected records from a weather radar (RADAR) and data from RCM and CPM models. The analysed dataset refers to the years 2005-2018, and was subsequently represented graphically.

As can be seen from Figure 1, the CPM and RCM graphs do not give very significant spatial information, whereas POINT and RADAR show that the data tend to thicken at certain points, thus creating clusters, i.e. spatial correlation. It is also noticeable that the RADAR and CPM models show a significantly higher number of data points. As for the RADAR data, it is likely that the visible clusters are due to the oversampling in azimuth of the same data, denoted by the authors, in ranges from 30 to 100 km.

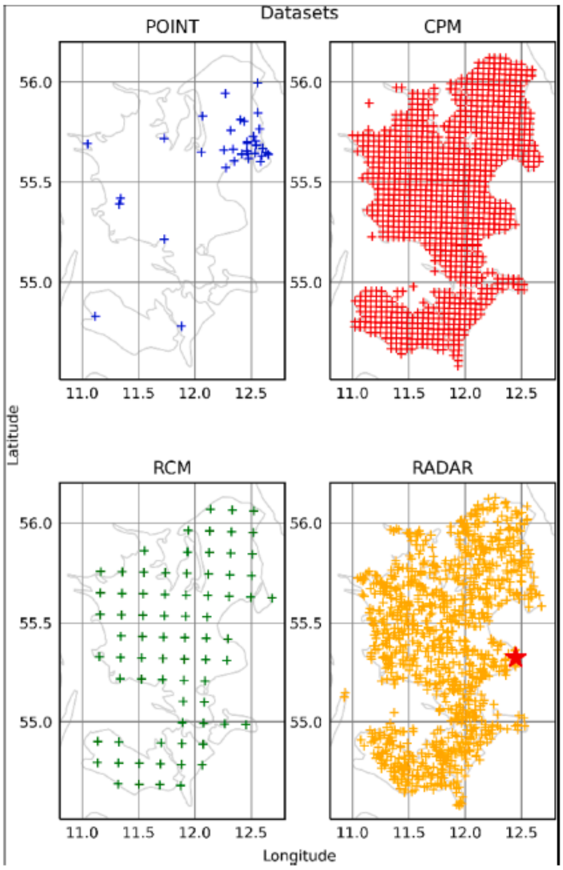

Figure 1: POINT, CPM, RCM, RADAR

Later, the authors talked about how they opted for random sampling over structured grid resampling. I believe they made the correct choice as, by using random sampling, the spatial correlation is not affected by the limitations of the grid.

Interestingly, the authors decided to apply a threshold for the RADAR data and the RCM and PCM model datasets (0.1 mm/hr and 0.2 mm/hr respectively). In this way, some information is cut off, but it is likely that too little data from a very light rainfall was collected. For this reason, it would not have made much sense to keep a layer containing too little information in the dataset, and a threshold was consequently applied, probably to limit the computational burden as well.

3 Methods

In this section, the focus was on extreme events, and how these may have influenced the phenomenon of extreme rainfall, trying to model them. From the sampling of extreme events, we move on to their magnitude and seasonal variation and end with their unconditional spatial correlation.

4 Results and discussion

As a first step, the focus is on the magnitude of extreme rainfall intensities. The authors here studied the intensity distributions for extremes for different return levels by comparing the four datasets, using a plot to better represent the results.

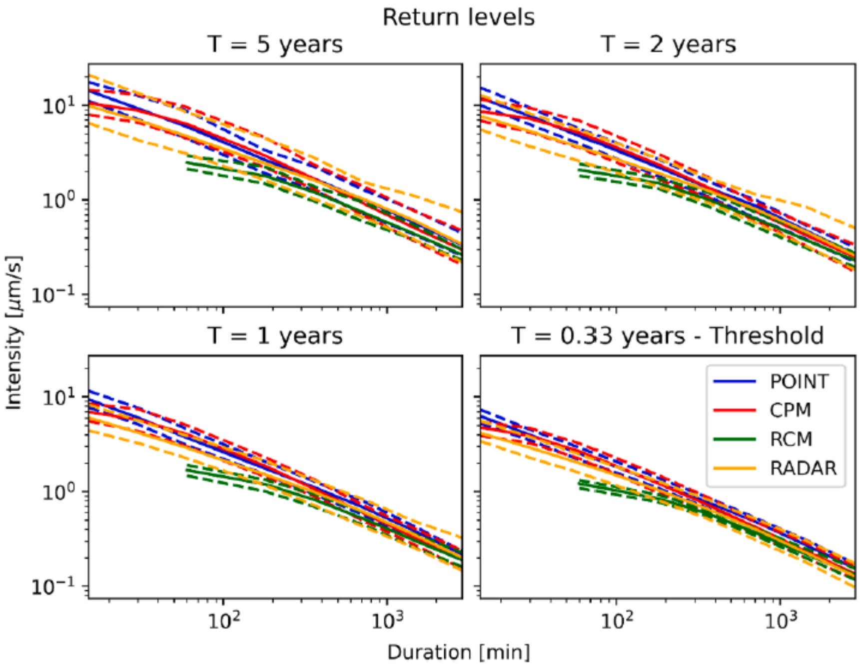

Figure 2: Return levels of datasets POINT, CPM, RCM, RADAR

I do not entirely agree when the authors speak of ”substantial differences” between the datasets. The differences are there, but I do not consider them to be so marked. From Figure 2, one can also see that most of the lines are overlapping, but it is quite evident that the RCM dataset has lower return levels than the other datasets. Moreover, these graphs are not very clear, either because of the overlapping of the lines or because of the chromatic range of the colours: I think they are too strong and not pleasing to the eye.

The seasonality of extremes was then explored. In Figure 3, although one notices a similar colour scheme to that used in the previous graph, the visualisation appears clearer because there is no overlapping problem. Moreover, here the RCM model shows similar behaviour to the other datasets.

Finally, the authors examined the unconditional spatial correlation. It is interesting to note that a solution is implemented to combine the occurrence of several events with the locations where they occur. This decision is similar to what I have already discussed in the previous paragraphs, and I find it sensible. As a conclusion of this analysis, it was found that the RADAR dataset appears to be the most consistent across scales, although the POINT and CPM datasets still perform well but in more specific circumstances. The RCM, on the other hand, is the worst of all, being unsuitable to be applied to an area such as urban hydrology.

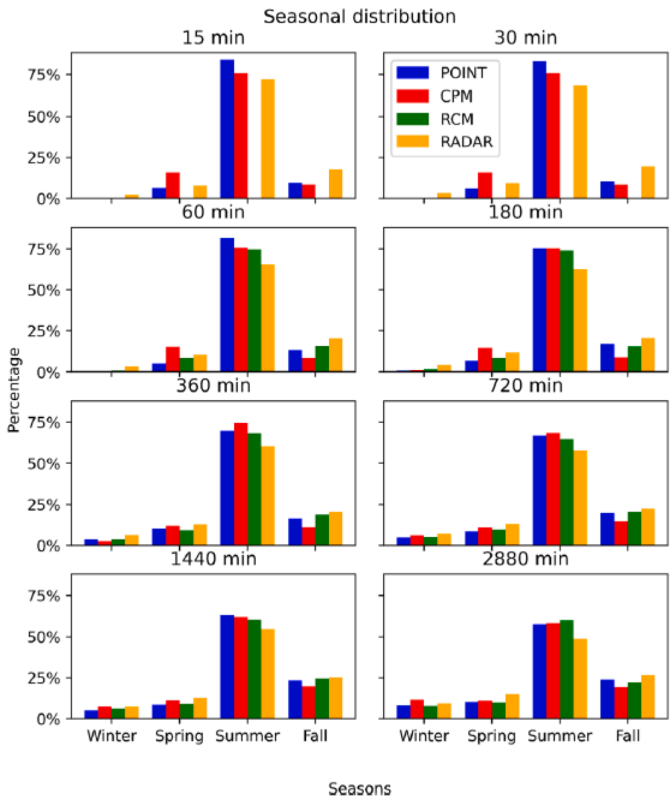

Figure 3: Seasonal distributions of datasets POINT, CPM, RCM, RADAR

As a final step, the authors decided to subdivide the RADAR dataset according to the distance from the radar, using a threshold distance of 63 km as this results in two bins of almost equal sizes (754/746 grid points). It was also considered that this choice also enables a separation into a bin with a constant altitude and a bin with a gradually increasing altitude of the lower scans due to the Earths curvature. The study uncovered a substantially differences between the correlation metrics, which indicates benefits of treating the subdivided RADAR datasets separately.

5 Conclusions

I reviewed the scientific research of E.D. Thomassen et al. regarding the study of extreme precipitation in south-eastern part of Denmark. Overall I can say that this article is quite complete and detailed, it treats each topic in depth and provides adequate explanations for questions that may arise during the reading process. However, I found the research slightly confusing, especially with regard to the different graphical representations. That said, I believe that the research was carried out critically in all its aspects.

Regarding further studies, I recommend adjusting the general clarity of the plots, i.e. the colour range and overlapping problems, perhaps by implementing other types of graphs. In addition, I would have liked the authors to have explained in more detail the reasons why the numerous combinations (merges and thresholds) seen during the review were necessary.