

## Replies to reviewer comment #1

### RC C241:

**Title:** Temporal and spatial scaling impacts on extreme precipitation

**Authors:** Eggert et al.

We thank the anonymous reviewer for the insightful comments, which we feel have helped improve the clarity of the manuscript! Our point-by-point replies (blue) to the reviewer comments (black) are given below.

### Reviewer #1

My only technical concern relates to the use of the “pdf overlap” metric which comes from Perkins et al (2007). Statisticians have been concerned with this problem for over six decades, with the Kullback-Leibler divergence commonly used to measure the difference between two probability distributions, and the two-sample Kolmogorov-Smirnov test being commonly used to test whether two pdfs differ. I wonder why the pdf overlap is used in place of these more traditional approaches, and whether it matters (for example, is there any potential for the pdf overlap to give misleading results, for example due to different sample sizes between pdfs etc)?

We thank the reviewer for this comment.

The PDF-overlap is calculated using normalized PDFs that use identical bins. All boxes with sample sizes less than 500 are excluded (this only occurred using the seasonal data in the supplement). Due to the normalization procedure the sample size should not make a difference.

The PDF overlap after Perkins is a very intuitive way of comparing PDFs and we find that it nicely mirrors the changes found in the 99th percentile as seen in Fig 3, and is probably better known and easier to comprehend by the climate impact community which we believe will be the main target group of the paper.

We included the following sentence to connect with earlier studies that use alternative measures:

“We find that the PDF overlap mirrors the changes found in the 99th percentile (Fig. 3a). Using cumulative PDF measures as the Kolmogorov-Smirnov statistics is an alternative way of comparing PDFs.”

Abstract [general]: Most of the abstract is focusing on method (what was done) and a bit more emphasis needs to be placed on results (what was found/discovered). The implications are also a bit vague for example the sentence “The resulting curve is relevant when deciding on data resolutions where statistical information in space and time is balanced” is very vague and should be made more precise.

We rephrased the abstract with a focus on the results.

Abstract, line 2: “Risk” is commonly defined as “probability \* consequences”. I think here it is only the probability that is of concern?

Yes it should have been “probability”. The word “Risk” is not used anymore in the new abstract.

Abstract, line 3: “qualitatively” – why not quantitatively?

We meant to emphasize the change from stratiform to convective type extreme events by using the word “qualitatively”. But of course there is also a quantitative change. Due to the focus change of the abstract this wording is not used anymore.

The new abstract:

Convective and stratiform precipitation events have fundamentally different physical causes. Using a radar composite over Germany, this study separates these precipitation types and compares extremes at different spatial and temporal scales, ranging from 1 km to 50 km and 5 min to 6 h, respectively. Four main objectives are addressed: First, we investigate extreme precipitation intensities for convective and stratiform precipitation events at different spatial and temporal resolutions, to identify type-dependent space and time reduction factors and to analyze regional and seasonal differences over Germany. We find strong differences between the types; with up to 30% higher reduction factors for convective extremes, exceeding all other observed seasonal and regional differences within one type. Second, we investigate how the differences in reduction factors affect the contribution of each type to extreme events as a whole, again dependent on the scale and the threshold chosen. A clear shift occurs towards more convective extremes at higher resolution or higher percentiles. For horizontal resolutions of current climate model simulations, i.e.  $\sim 10$  km, the temporal resolution of the data as well as the chosen threshold have profound influence on which type of extreme will be statistically dominant. Third, we compare the ratio of area to duration reduction factor for convective and stratiform events and find that convective events have lower effective advection velocities than stratiform events, and are therefore more strongly affected by spatial than by temporal aggregation. Finally, we discuss the entire precipitation distribution regarding data aggregation, and identify matching pairs of temporal and spatial resolutions where similar distributions are observed. The information is useful for planning observational networks or storing model data at different temporal and spatial scales.

## Introduction

I am finding the introduction a bit underwhelming. There are a lot of great threads of ideas, and the authors have succeeded in capturing the relevant literature, but the ideas could be brought together much better and the relevance of ideas to the paper made more explicit.

For example, how is the “alarming” finding that statistical downscaling procedures assume that the empirical relationships between large and small scales hold in the future relate to the research proposed here?

Our study shows that large and small scales emphasize different events. Assuming that the empirical relationships between the scales will hold in the future would hence mean to assume that both types of events will behave similarly in the future. The different response of the different precipitation types to temperature increase is however largely discussed and we cite 4 papers to this topic.

We strongly shortened this part of the introduction since it is only indirectly related to the research presented here.

Similarly, while it is obvious that convective and stratiform rainfall would require different climate model resolution (page 2162, line 5), it's not clear whether the space-time resolutions

identified in this paper would necessarily be equivalent to the “minimal climate model resolution”.

We here only talk about model output resolution. The actual calculation time step has to be much shorter and is not subject of this research. Still the space-time resolutions identified in this paper are obtained using observational data and cannot directly be translated to model resolutions. But knowing the space time relation of precipitation events will lead to the knowledge about what is statistically possible to be captured at a certain resolution. This will be a great help in order to validate and set up model and observation studies.

What is the issue with the simple power law dependence not holding generally, and is the “regime-distinction” related to the classification of convective/stratiform rainfall or is this a different issue?

We could not find a direct connection. To avoid confusion we left this issue out of the introduction.

Page 2162, paragraph 2: the need for the study could be made stronger; the importance of this question is not stressed enough. Please include more literature or other examples to explain why it is importance.

We made major changes in the introduction by shortening and emphasizing stronger on ideas directly relating to the results of the paper. The need for the study should now be made stronger.

New Introduction:

The IPCC’s fifth assessment report highlights an intensification of heavy precipitation events in North America and Europe (Hartmann et al., 2013), and projects further increase of extremes as global temperatures rise (Collins et al., 2013). The study of extreme events is complex due to a strong inhomogeneity of precipitation intensities in space and time. Assessment of precipitation extremes, e.g. as defined by an intensity threshold, therefore always requires specification of the relevant spatial and temporal resolution.

Even though spatial and temporal scales are far from independent (Taylor, 1938), it is often unclear how to compare datasets directly, when their data is measured at differing resolutions. The data resolution needed by users, e.g. hydrologists or crop modelers, often differs from that at which observed or modeled data is recorded (Willems et al., 2012).

The primary societal interest in extreme precipitation lies in its hydrological implications, typically requiring statistics of precipitation extremes for the area of a given catchment or drainage system, which is not identical to that of model grid boxes or the observations.

Moreover, temporal scales relevant to flood risk vary enormously with area (Blöschl and Sivapalan, 1995; Westra et al., 2014): For catchments, hours to days are relevant (Mueller and Pfister, 2011), whereas urban drainage systems of  $\sim 10$  km (Arnbjerg-Nielsen et al., 2013) are impacted at timescales from minutes to hours (De Toffol et al., 2009), and soil erosion can occur at even smaller scales (Mueller and Pfister, 2011).

Areal Reduction Factors (ARF) and Intensity Duration Functions (IDF) have previously been used to describe the decrease of average precipitation intensity due to spatial and

temporal aggregation (Bacchi and Ranzi, 1996; Smith et al., 1994). The capability of radar data to capture the spatial structure of storms was identified as a key factor in deriving the ARFs (Bacchi and Ranzi, 1996; Arnbjerg-Nielsen et al., 2013). A general outcome was that ARFs exhibit a decay with respect to the return period (Bacchi and Ranzi, 1996; Sivapalan and Blöschl, 1998) and a dependency on the observed region, resulting from different governing rainfall generation mechanisms (Sivapalan and Blöschl, 1998).

In the current study we separate the physically different processes leading to convective and stratiform type precipitation events. Using synoptic observation data, we classify precipitation events into these two types, allowing us to analyze their aggregated statistics individually across scales.

The two types physically differ in that convection is often initiated by local radiative surface heating, resulting in a buoyantly unstable atmosphere (Houze, 1997), whereas stratiform precipitation stems from large-scale frontal systems and relatively weak and uniform up-lifting. Analyzing these two types separately regarding their intensities at different scales can e.g. be important when considering temperature changes, such as anthropogenic warming: Over large scales, the changes were found to be moderate, whereas for very small scales, it has been argued that the two processes may increase with warming (Trenberth, 1999; Trenberth et al., 2003; Trenberth, 2011; Lenderink and van Meijgaard, 2008), albeit at very differing rates (Berg et al., 2013). Using high-resolution model simulations, heavy precipitation at high temporal resolutions was suggested to increase strongly in a future climate, and a dominant contribution to extreme events to stem from convective events (Kendon et al., 2014; Muller et al., 2011; Attema et al., 2014).

In spite of their small horizontal and temporal range, convective events can cause substantial damage (Kunz, 2007; Kunz et al., 2009), e.g. through flash floods (Marchi et al., 2010).

Numerous studies have assessed the temporal and spatial characteristics of precipitation events using a storm centered, or *Lagrangian*, approach (Austin and Houze Jr., 1972; Houze Jr. and Hobbs, 1982; Moseley et al., 2013), which focuses on the storm dynamics, e.g. lifetime or the history of its spatial extent. Moseley et al. (2013) showed that, for Lagrangian event histories of 30 min, the convective type can produce significantly higher intensities than the stratiform type. As we here focus on potential hydrological applications and those addressing possible impact of extremes, e.g. floods, defining events over a *fixed* surface area and time period is more appropriate (Berndtsson and Niemczynowicz, 1988; Onof et al., 1996; Bacchi and Ranzi, 1996; Michele et al., 2001; Marani, 2003, 2005). The statistics thereby constitute averages over a defined space-time window within which both dry and wet sub-intervals may occur.

In this study, we analyze at which fixed temporal and spatial scales convective precipitation dominates precipitation extremes. To this end, we analyze two years of mid-latitude high-resolution radar data (5 min temporally and 1 km spatially), classified by precipitation types and separated into seasons (summer vs. winter) and geographic areas (north vs. south Germany). Analysis of these data over large spatial and temporal periods characterizes the statistical aggregation behavior in space and time. It can quantify the requirements on minimal model resolution sufficient for the proper description of the respective extremes. Revisiting the Taylor-hypothesis (Taylor, 1938), we contrast the two precipitation types, as to how resolutions in space and time can be compared. Using a resulting effective advection velocity, we give a simple means of quantifying effective temporal averaging in models, resulting from a given spatial resolution.

The structure of the article is as follows: In Sec. 2 we describe the data and methods used. Section 3 presents the results for extremes at different resolutions (Sect. 3.1) and suggests a method to compare the corresponding probability density functions (Sect. 3.2). We close with discussions and conclusions (Sect. 4).

Data and methods:

Page 2163: How many synoptic cloud observation stations were used?

222 stations in total, we included this in the text.

Synoptic cloud observations, at 222 stations, obtained from the Met Office Integrated Data Archive System (MIDAS) data base [[http://badc.nerc.ac.uk/view/badc.nerc.ac.uk\\_ATOM\\_dataent\\_ukmo-midas](http://badc.nerc.ac.uk/view/badc.nerc.ac.uk_ATOM_dataent_ukmo-midas)] are used to separate large-scale and convective precipitation following Berg et al. (2013).

Results:

Second paragraph, page 2165 on temporal aggregation: Just a suggestion, but could “the effect of temporal aggregation is to even out spatial variations due to large-scale flow” be illustrated using a conceptual diagram? Similar for the subsequent discussion of Taylor’s hypothesis. Again, this would help expand the appeal of this paper.

We included a diagram showing the concept of the Taylor hypothesis together with the two major assumptions made in order to use this hypothesis for our analyzes (frozen in time, no variability perpendicular to the advection direction). We further rewrote parts of the chapter to better explain the concept.

Page 2159: Can you provide a more formal definition of an Areal Reduction Factor here?

We changed the definition to a more formal one.

Equation 2, page 2167: Is “x” a length scale? Can you confirm whether this is consistent with the standard definition of an ARF? [since an area is a length<sup>2</sup> scale].

x is the grid size hence it is a length<sup>2</sup> scale.

Page 2168-2169: This section would have been much more compelling with some illustrative diagrams of the “frozen turbulence” vs self-affine concepts, and how the choice of interpretation would lead to differences in space-time aggregation. This is the same issue as made in reference to the Taylor’s hypothesis on page 2165, and I think that a conceptual diagram would make the results much easier to interpret.

We have rewritten parts of the chapter that describes the self-affine process to make this clearer. Also a conceptual diagram is added.

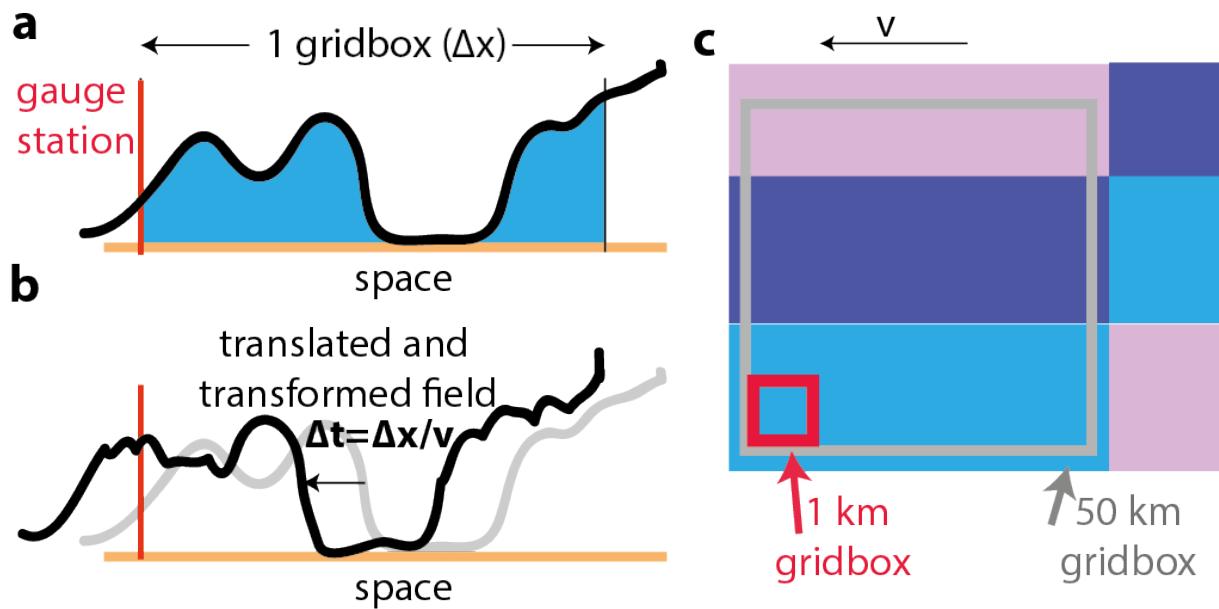


Figure 6. Schematic illustration of the Taylor hypothesis. (a) One-dimensional case, showing space, gridbox width and precipitation intensity (black curve); the location of a gauge station is marked in red. (b) Similar to (a), but illustrating how the curve may change due to small scale dynamics after a time interval  $\Delta t = \Delta x/v$ , with  $v$  the atmospheric advection velocity. (c) Two-dimensional inhomogeneity (different colors indicate different intensities) perpendicular to the advection direction (direction indicated by the thin arrow). Small (red) and large (gray) gridboxes as marked.

Page 2178: "the optimum temporal resolution for state of the art regional climate simulations, performed at a 11 km horizontal resolution, would be approximately 20 to 25 minutes." This to me is an extremely important practical outcome of the paper but is a bit buried here. This is the sort of thing that could be highlighted in the abstract? Similarly, the finding that different meteorological events are considered extreme depending on the threshold is an interesting finding.

We changed the abstract and the conclusions to highlight this point more clearly

Page 2159, line 3-6: this sentence was unclear, please rewrite.

we rephrased the sentence.

Assessment of precipitation extremes, e.g. as defined by an intensity threshold, is strongly scale dependent and therefore requires specification of the analyzed spatial and temporal resolution.

Page 2164: what is the role of the apostrophe in  $I'$ ? This is not defined or used elsewhere.

That is standard math nomenclature to show that it is the variable being integrated over.

Page 2176, line 20: "smoothening" should be "smoothing"

yes we changed this.

Page 2162, line 4-6: should be rephrased as a question

This sentence is not included anymore in the new introduction.

Page 2165, line 4-5: unclear, please rephrase

We rephrased the sentence:

Stratiform precipitation is more uniform in the sense that sampling over small areas yields a good description of the statistics also at larger areas of aggregation.

Figure 7 caption: “larger or equal” should be “greater than or equal to”

We changed the figure caption.

Figure 8 caption: “larger or equal” should be “greater than or equal to”

We changed the figure caption.