

# RESPONSE LETTER

---

Review of Manuscript No.: **essd-2017-35**

Title : **SPREAD: A high-resolution daily gridded precipitation dataset for Spain**

Author(s): **Serrano-Notivol R., Beguería S., Saz M.A., Longares, L.A. and De Luis M.**

---

## POINT BY POINT: REVIEWER #2 COMMENTS:

### GENERAL COMMENTS

This manuscript is devoted to the development of a new high- resolution daily gridded precipitation dataset for Spain using a large number of stations (12 858) and over the period from 1950 to 2012 (peninsular Spain) or from 1971 to 2012 (Balearic and Canary Archipelagos). Data is available over a 5 km grid. The dataset is publically available to users and the authors provide not only precipitation estimates, but also their corresponding uncertainties. The dataset is properly validated with observational data. The methodology is adequate and based on previous studies from the authors, namely on their R-package 'reddPrec'. Furthermore, a number of precipitation indices is also analysed, including indices of extremes. The text is generally clear and well written. The high-resolution dataset produced by this study is of major relevance for impact assessments over a wide range of socioeconomic sectors and for decision-making. Therefore, I recommend the publication of this manuscript after some very minor revisions detailed below.

Thank you for your kind and positive comments. Indeed, the SPREAD dataset is not only a final product in itself, it could help in a wide range of decision-making policies.

### SPECIFIC COMMENTS

**1. The title should explicitly mention that a climatological analysis is undertaken, including an analysis of precipitation extremes.**

As indicated in previous comments in the response to reviewer#1, a short statement in the title has been added as suggested to indicate the analysis of the extreme precipitation developed in the manuscript. The title is now: "*SPREAD: A high-resolution daily gridded precipitation dataset for Spain. An extreme events frequency and intensity overview.*".

**2. Section 2: I would like to see here more discussion regarding the implications of the data gaps on the results. As the authors mention, only 17 stations actually cover the full period. Although the station density remains reasonably high throughout the whole period (please revise Y-axis labels in the bottom panels of Fig. 1), some important limitations/uncertainties are expected to arise from this lack of data. Please enhance this discussion.**

Indeed, only a few stations cover the complete period. However, the use of the complete stations network helped to a more reliable daily estimate of precipitation in longest data series. Only reconstructed stations with more than 10 years of original data were used to build the grid. The use of short series introduces more missing data in the whole dataset, but estimating precipitation day by day with the 10 nearest neighbours there are more probabilities of finding near data, which improves the final estimate.

A new paragraph has been added at the end of the discussion section:

*"The use of the complete information of the precipitation network in Spain provided a more detailed precipitation distribution over time and space. Although only a few stations covered the complete period, the use of short data series helped to estimate the missing precipitation values in longer ones, which were used to build the whole grid. A high number of grid points (higher spatial resolution) in combination with a low-density*

*stations network could lead into higher uncertainties. This work aimed to set a compromise between both factors by using a high number of stations and a mid-high spatial resolution. In addition, the magnitude of the uncertainty informed about the reliability of each estimate. A higher uncertainty means more differences between the data used to estimate precipitation and these differences can be increased with a lower number of stations.”*

**3. Page 5, Lines 9-11: the definitions of suspect wet and dry days seem to be reversed.**

Right, “wet” and “dry” words were switched. Sorry for mistake.

**4. Section 3: from my viewpoint, this section should provide further details concerning the followed methodology. I understand that there are limitations in the paper extent, but a deeper description of the methods should improve the readability of the text and may prevent readers from reading preceding papers.**

As you indicate in your comment, there are limitations in the paper extent. The basics of the methodology were tried to be explained in the text: The reference values (RV) used for quality control and reconstruction (3.1) and the gridding (3.2). Both the reconstruction of the original series and the new estimates for the grid points are based on the computation of the RV, which are calculated with GLM (Generalized Linear Models) based on the 10 nearest neighbors. We tried to synthesize the whole method in section 3, and we thought that, as this work is not only methodological but used to present and validate a data product, the methods basics were enough to the understanding of the grid creation process. The cited previous work, Serrano-Notivoli et al., 2017 (now, 2017a), was referred to the R package used to build the grid. However, we added a new reference (Serrano-Notivoli et al., 2017b) that widely explains the details of the method that we think are not essential for this product presentation.

Serrano-Notivoli, R., de Luis, M., Saz, M.A., Beguería, S.: Spatially-based reconstruction of daily precipitation instrumental data series, *Clim. Res.*, doi: 10.3354/cr01476, In press, 2017b.

We made an effort to summarize the methodological part because, as it is already published, we tried to avoid duplicities. Anyway, we can extend the explanation of any specific point if requested. Please, let us know and we will be pleased to do that.