Journal: Climate of the Past Discussions
Title: Increasing cloud cover in the 20th century: review and new findings in Spain
Author(s): A. Sanchez-Lorenzo et al.
MS No.: cp-2012-29

Response to Referee #1

General comment

This work serves as an excellent example of how to use historical weather and climate data to build and analyze a quasi-homogeneous time series of total cloud cover. The authors have found digitized historic climate data from various incarnations of the Spanish weather service (AEMET) and have developed a simple and appropriate routine to homogenize cloud data that was gathered using different formats. Their analysis shows that total cloud cover was increasing over Spain from the late 1800's through the 1960's, and has been decreasing since. They also use principal component analysis to separate out climatologically similar regions. This work could (should) be used as a template for looking farther into past climate records in many nations, as records have only been homogeneous across most national boundaries after the 1970's.

We really appreciate the reviewer's comments and his/her support for the publication of our manuscript in Climate of the Past.

The manuscript has been revised after considering these comments, which are addressed below.

Specific Comments

1) Would it be possible to analyze the truly homogeneous stations in individual groups, and then to compare these sub-analyses to the single time series shown? I envision a plot consisting of the single curve at the top of Figure 4 along with time series of stations that only reported in tenths, oktas, or clear/overcast format. Showing agreement between these time series prior to homogenization would be helpful. This might be similar to Figure 2b, but with the 'true' TCC curve divided into multiple curves, each based on truly homogeneous data.

Following the reviewer's suggestion, we have added a new plot (Fig. 3 in the revised manuscript, see below), as well as a new paragraph in Section 3.3 that reads as: "Finally, Figure 3 shows the mean annual estimated TCC homogenized series obtained by averaging the 39 stations during the whole 1866-2010 period. Figure 3 also shows the mean annual estimated TCC series for the original series, a curve divided into the 1866-1960 and 1961-2010 subperiods that are derived from overcast/cloudless days and "true TCC" observations in oktas, respectively. It is clear from this figure that the estimation of the TCC by using only cloudless and overcast conditions before the 1960s and the derivation of these sky conditions from the "true TCC" after the 1960s do not introduce a bias in the whole estimated TCC series. Equally, it is obvious that the homogenization step does not introduce any artificial trends in the series."

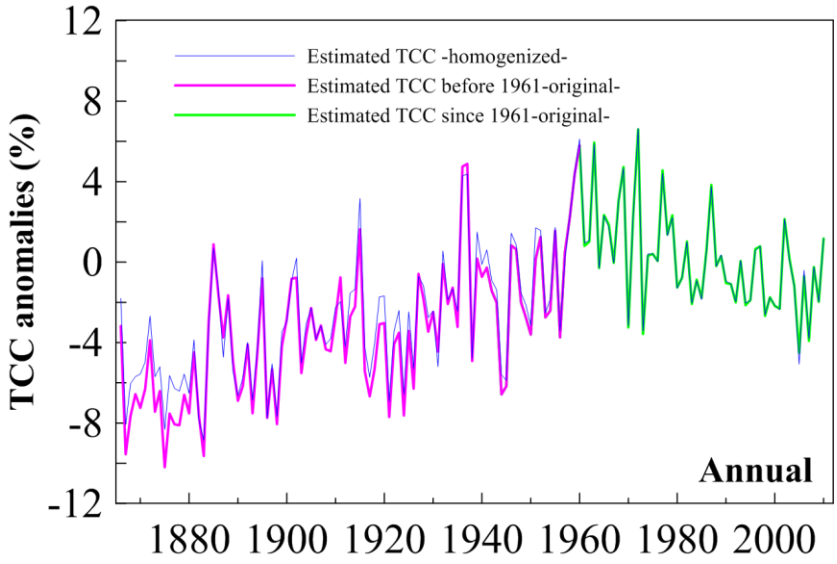

Figure 3. Time evolution of the mean annual series in Spain for the estimated TCC homogenized series over the 1866-2010 period and for the original estimated TCC series divided in the 1866-1960 and 1961-2010 subperiods. The anomalies are expressed as differences from the 1971-2000 mean.

2) Spatial averaging (projecting stations onto a uniform grid, creating grid-box averages, then averaging grid-boxes) may help reduce bias in the time series. This may not be much of a factor, since the chosen stations are fairly well-distributed, but the current time series is slightly biased towards more densely sampled areas.

The reviewer is right, but it is also true that a grid-box average will not produce significant changes in our mean series. As we highlighted in Section 3.4, the subset of 39 series available in this study can capture almost the same interannual and decadal variability during the 1961-2010 period as using the grid data generated by Sanchez-Lorenzo et al. (2009). This paragraph reads as: "Due to the strong spatial autocorrelation in the TCC series, with the subset of 39 series available in this study it is possible to capture the same variability as using the grid data generated by SL09 with 69 series over the Iberian Peninsula, i.e. correlation coefficients between 0.96 and 0.99 for seasonal mean series (not shown). Consequently, the dataset used here can perfectly capture the interannual and decadal variability of the whole Iberian Peninsula."

On the other hand, in the framework of the postdoctoral position of the first author (http://www.iac.ethz.ch/people/arturos/suncloud/), we are planning to generate a gridded version of the total cloud cover variable, but including also series from other countries over Europe. Before reaching this goal, we prefer to analyze and provide only a station-mode of this data set.

3) On page 1136 on line 26, you say that Warren et al uses 54000 stations, however that study actually only uses 5400 stations.

We have corrected the error.