

***Interactive comment on “Multilinear approach to the precipitation–lightning relationship: a case study of summer local electrical storms in the northern part of Spain during 2002–2009 period” by I. Herrero et al.***

**Anonymous Referee #1**

Received and published: 1 January 2014

I have read this paper dealing with precipitation-lightning relationships in northern Spain. While the paper is interesting, I am not sure WHY the authors decide to use the lightning data to estimate the rainfall for the region? Figs. 2, 3 and 4 show poor correlations between lightning and rainfall, however, the authors still decide to go forward and use such relations in their estimates of the rainfall. What is the justification of using the lightning data in this region for the estimates, when the agreement of the basic raw data is so weak? Stronger justification needs to be given.

C2176

The whole idea of the paper is to give a reasonable "estimate" of rainfall for the region (equation 2), and the verification is done using correlations between rainfall observations and estimates. However, the estimates themselves use the observed rainfall in some way to predict the observed rainfall. Hence, these two parameters (observed and estimated rainfall) are not INDEPENDENT parameters. The estimates depend on the observed values, not only on the lightning data. Hence, I am not sure about the validity of these statistical correlations. You can only correlate estimates with observed rainfall if the estimate is entirely independent of the observations (e.g. from model estimates, satellite estimates, lightning estimates, etc.)

P6472, line 16: The DAILY accumulated rainfall may be from gentle rain throughout the day, or a heavy rainfall in one hour. This would result in very different lightning amounts. So this may be the cause of the bad correlations.

P6477, line 21: Equation 2 includes the lightning data, but the Figs 2-4 show very poor links between rainfall and lightning in this region. So why use this method? How can you justify this? line 26: change "excess/defect" to "over/under estimation"

Table 2: I suggest highlighting the 3 stations discussed here. I guess each line in the table is another station.

P6478, lines 9-24: You mention only 3 stations giving good results. That is 3 out of 22 stations (< 15%). What about the other stations? How can you base your results on less than 15% of the stations?

P6480, line 21-23: You propose using only one station for the estimates, but then state that only 3 stations show high correlations. So only 3 out of 22 stations could be used for the estimate. This is not very reliable statistics.

P6481, line 3: use of an empirical line 6: adjusted in terms

---

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., 1, 6467, 2013.

C2177