Nat. Hazards Earth Syst. Sci. Discuss.,
doi:10.5194/nhess-2016-198-AC2, 2016

Natural Hazards and Earth System Sciences
Discussions
Open Access
EGU

Interactive comment on "Spatio-temporal smoothing of lightning climatologies" *by* Thorsten Simon et al.

Thorsten Simon et al.

thorsten.simon@uibk.ac.at

Received and published: 23 August 2016

Answers to Referee 2

———————

Thank you very much for your informative comments. They will clearly help improving the quality of the manuscript.

General Comments:

———————

"This paper is basically comprehensible, well structured and written in good English. Moreover, the general idea of the paper is interesting and the given approach is straight forward and certainly viable."

Printer-friendly version

Discussion paper

Answer: Thanks for acknowledging this. We agree that the approach is straightforward for someone with experience in GAMs. However, we feel that this is not necessarily the case for all readers of this manuscript at the intersection of lightning science, climatology, and applied statistics. Hence one objective of the manuscript is to bridge some of the gaps and make GAMs more accessible to researchers in the field of lightning science. Both your comments and those of Referee 1 show that we haven't fully accomplished this goal and hence we are grateful for your suggestions for improvements.

"Since I got the impression that a major asset is the modeling for a complex terrain, I would like to know what is the benefit of adding an altitude effect to the statistical model, whereas the lon/lat part seems to be the most influential effect? Moreover, I am not sure whether spatial function and altitude function are really distinct. Isn't it just sufficient to take the location into account because it implicitly contains the altitude?"

Answer: It is true, the altitude is a function of longitude and latitude. In general the presented method would be capable to model the influence of the altitude within the spatial effect implicitly. However, the shape of the altitude in the region of interest is very complex. Thus, a spatial effect with a large degree of freedom would be required in order to account for the complex altitude shape. As we know the shape of the altitude we can pass it to the GAM as an isolated effect. The altitude effect contains only information associated with the altitude while the remaining effects are captured by the lon/lat term. We will mention this aspect in the results section.

"Finally, in terms of verification, it is not clear what kind of scores were calculated or used and what their results are."

Answer: The log-likelihood is applied, also called logarithmic score in the literature on proper scoring rules (see Gneiting 2007). We tried to avoid showing the results of the scores in detail, which would mean showing a longish table with proposed values of the smoothing parameter and associated scores from which the best is selected. Instead we wanted to put more emphasis on the results. We will add a paragraph in

the methods section to discuss the verification score.

A table summarizing the verification scores would look like this:

λ	Q. 2.5%	Median	Q. 97.5%	d.o.f.
–	762455	765275	768364	0.00
5e+09	754052	756881	759698	1.07
1e+09	751785	754598	757377	1.27
5e+08	749880	752880	755764	1.45
1e+08	746750	749558	752554	2.11
5e+07	**746356**	**749251**	**752243**	**2.51**
1e+07	746754	749571	752341	3.77
1e+06	748266	751320	754269	6.73
100000	753277	756236	759019	11.78
10000	764341	767496	770352	19.38
0	786475	789802	792789	29.00

Table 1. 6-fold cross-validated negative log-likelihood for different smoothness parameters of the temporal effect. The dash in the λ-column indicates that no temporal effect was included into the model. Median, 2.5% quantile and 97.5% quantile was generated by bootstrapping 1000 times.

Specific Comments:

Title "I am afraid that the title 'Spatio-temporal smoothing of lightning climatologies' is misleading, because spatio-temporal smoothing implies some kind of grid-wise and time-wise moving average or filter, while the main idea of your study is to decompose the signal into a seasonal, spatial and also altitude effect by a statistical model. Reading the paper, I would have entitled it something like 'Statistical modeling of lightning climatologies for complex terrains' or 'Spatio-temporal smoothing of lightning climatologies for complex terrains'..."

Answer: Thanks for pointing this out and for suggesting the alternatives. We will change the title to "Spatio-temporal modeling of lightning climatologies for complex terrain".

Introduction "Reading your introduction, I got the impression that thunderstorms/lightning tends to occur at regions with moderate or lower altitude (page 2, lines 4-8). But your figures 3 and 4, top-left implying a positive and linear relationship between altitude and occurrence/intensity. Why doesn't the GAM fits a function with maxima for lower/moderate altitudes?"

Answer: The observed maxima at moderate or lower altitude are not the general case, but special cases associated with local effects. Thus it is not visible in the altitude effect. E.g., the maximum in the Gurktal Alps cannot be explained by its altitude, but the maximum has to be a consequence of local attributes of the terrain in that region. We will add a sentence in the results section where the effect is introduced.

Data "Page 3, line 1: Reading this, with little experience on this scientific field, I would like to know the distinction between lightning, flash and stroke?"

Answer: Lightning is defined as a transient, high-current (typically tens of kiloamperes) electric discharge in the air whose length is measured in kilometers. The lightning discharge in its entirety is usually termed a 'lightning flash' or just a 'flash'. Each flash typically contains several 'strokes' which is the basic element of a lightning discharge. We will add a sentence in the data section to clarify the nomenclature.

"Maybe, it would be interesting to show a figure with the spatial climatologies of the number of flashes in Carinthia for the raw data."

Answer: Such a figure will be added with explainations (cf. Fig. 1 in this response).

Methods "page 4, line 6: As mentioned before, are altitude and horizontal space (lon/lat) really distinct. Thus, eq. (1) probably would have the form: $g(\theta) = \beta_0 + f_1(doy) + f_2(lon, lat, logalt)$"

Answer: As pointed out above one could just use $f_2(lon, lat)$ because logalt is a function of lon/lat but that would necessitate a very complex lon/lat term (using many degrees of freedom). The suggestion $f_2(lon, lat, logalt)$ could be interpreted as a spatially varying logalt effect. This is, in principle, also possible but is also more challenging to estimate. The additive decomposition $f_2(lon, lat) + f_3(logalt)$ is "the usual" trick of using an additive decomposition of the effect which leads to relative parsimonious effects f_2 and f_3. Of course, there may be even better parametrizations but this seems to work well and is (relatively) easy to interpret for practitioners.

Results "page 5, line 20: How does the 1000 day-wise block-bootstrapping work?"

Answer: With day-wise block-bootstrapping we mean the following: We draw 738 dates of all available days with repetition and pick all the data observed on these days spatially. If we would relax the day-wise structure we would draw 7309152 samples with repetition from all available data points. An explaination will be added in the manuscript.

"page 6, line 4: Is there any explanation for the maximum in the Gurktal Alps, although this region is quite low elevated?"

Answer: We haven't found an explanation yet. However, in a follow-up study we will set the focus on analysis of single events and associated synoptical situations. Hopefully, this study will provide more insights.

"page 6, line 11: Is there any explanation for the flatter shape of the altitude effect function?"

Answer: We haven't found a sound and strong explanation for that shape.

Discussion "In my opinion the part where the authors explain that cross-validation with day-wise blocks is much smoother and subsequently recommend to explore dependence structure of the data first would be more suitable for the method section."

Answer: Yes, we agree that this could be part of the methods section which would probably be the more natural section for readers with experience in flexible regression

modeling (with GAMs). However, we deferred it to the discussion in order to make the methods section more accessible for readers not so familiar with GAMs. Hence we felt it would be easier for that audience if the the cross-validation is explained along the concrete example rather than abstract formulae. To better accomodate readers with experience in GAMs we have now added a forward reference in the methods section with only some short comments.

Conclusion "Page 8, line 30-32: As far as I understand, in section 4.2 the higher spatial variability of the intensity model is explained due to local constructions, that trigger the number of flashes without affecting the occurrence. However, in the conclusion part one get the impression that higher spatial variability of the intensity model is distinct from local maxima in the vicinity of radio towers. Thus, I would suggest a sentence like: 'In particular the spatial effect of the intensity model varies more strongly than the corresponding effect of the occurrence model, because local intensity maxima are triggered in vicinity of radio towers. Moreover other new features were exhibited like...'"

Answer: We adopt this sentence.

References:

Gneiting, Tilmann, and Adrian E. Raftery. "Strictly proper scoring rules, prediction, and estimation." Journal of the American Statistical Association 102.477 (2007): 359-378.

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., doi:10.5194/nhess-2016-198, 2016.

Printer-friendly version

Discussion paper

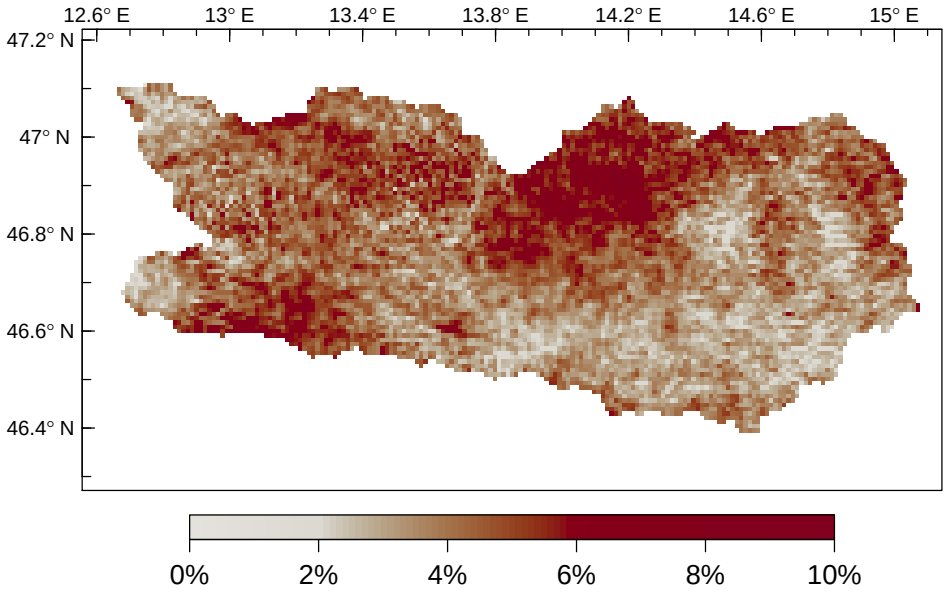

Fig. 1. Empirical climatological probability of lightning for a day in July in Carinthia on the 1km x 1km scale.

Printer-friendly version

Discussion paper