Anonymous Referee #1

The manuscript by Yiou et al. provides an interesting example of how unknown meteorological fields in the past can be reconstructed from limited observational data through searching for analogous fields from more recent observations or reanalysis data. The study makes use of historical pressure data to reconstruct an ensemble of potential meteorological conditions (SLP, wind and temperature anomalies) in 1781-1785 that could have occurred at the years around the Laki eruption.

As the methodological design of using "atmospheric flow analogues" has been described already in another publication by the author (Yiou et al., 2012, Clim. Dynam.), the paper focuses on the investigation of the reconstructed ensemble of the fields.

Based on the retrieved analogous fields for the past and their sampling frequency from the period after a more recent volcanic eruption on Iceland 2010, the authors test whether these events share some similar characteristics following the volcanic eruptions.

We thank the referee for this constructive and thoughtful review.

General comments:

Although the analog method was already evaluated for different statistical down-/upscaling purposes since Zorita and von Storch (1999), the number of applications as presented in this study is still quite low. This study is hence an important contribution to further evaluate and potentially establish the analog-method as an alternative approach in the context of climate field reconstructions. In addition, the authors propose a nice way to test whether two events like weather conditions after volcanic eruptions share some similar characteristics based on the sampling frequency of analogs from a specific period. Relying the reconstruction on an ensemble of likely analogs rather than one best analog is an important aspect of this study to deal with the uncertainties of the historical data and the analog-reconstruction.

Overall, the manuscript is mostly well written and provides interesting information on an important climate/weather episode of the late 16th century's European history. In addition to the analog reconstruction, the study introduces new historical temperature data of that period to evaluate potential temperature anomalies after the Laki event.

Thank you.

Some sections of the paper should be explained/discussed in more detail helping the reader to get more familiar with the method and/or the climate impact of volcanoes.

We will strive to clarify the manuscript in the directions suggested by both reviewers.

Also the newly introduced historical temperature data should be presented in a bit more detail. The sections to be enhanced are listed below in the specific comments.

This will be done in the revised manuscript. The datasets will also be made available.

While the points above can be easily improved by the authors, I'm a bit more concerned about the comparison of the Laki eruption with Eyjafjallajökull in 2010. The Laki eruption was a very long-lasting event injecting huge amounts of aerosols into the atmosphere over a long time while the recent

eruption was comparably short, sulfates were only injected to the troposphere and hence got quickly washed out. Based on this difference, I would a priori not expect any similarities due to the occurrence of the eruption. I would suggest that the authors say more clearly that they aim mainly for reconstructing the potential flow patterns transporting the aerosols rather than its absolute cooling impact.

The original rationale of the study was to investigate whether the atmospheric circulation that prevailed in the Spring of 2010 was similar to the one of Spring 1783, all else being equal. This will help us (in a further study) make a first guess estimate of the volcanic plume, from a chemistry transport model and simple hypotheses of the gas and particle emissions of the Laki. In this paper, we just wanted to investigate this hypothesis (of a similar atmospheric circulation), which would bring an interesting constraint to such an experiment. We are not comparing the eruptions themselves, which were quite different.

There is also a large difference of how tropical vs. extra-tropical explosive volcano eruptions might influence atmospheric circulation and regional temperatures. This should be discussed in more detail e.g describing also the results from model simulations (e.g. Kravitz & Robock, 2011: The climate effects of high latitude volcanic eruptions: The role of the time of year; Oman et al. 2005: Climatic response to high latitude volcanic eruptions; Oman et al. 2006: Modeling the distribution of the volcanic aerosol cloud from the 1783-1784 Laki eruption; Schmidt et al. 2012: Climatic impact of the long-lasting 1783 Laki eruption).

We agree that comparing volcanic eruptions in terms of emissions and climate impacts is not covered by this study, as it not our intent to do so. We will put less emphasis on the volcanic impacts, as it seems to confuse the message of the manuscript.

I'm sure that all aspects can be easily addressed by the authors and highly recommend the publication of the manuscript in CP after taking into account the general remarks and the points below.

Specific comments:

Abstract, page 5158, line 10 ff.: As the paper reconstructs temperature anomalies relative to the short mean around the Laki eruption, I would not agree that Laki did or did not make these winters regionally/locally even colder based on this study. The special situation of Laki was the very long-lasting eruption/aerosol injection so that the question about the memory of the atmosphere was less relevant here. In principal you show that at least the flow anomalies are not unprecedented. I would modify or remove this statement about the cold winter related to Laki or better justify it in the paper to be valid. The same applies for p. 5169, line 11.

The paragraphs will be rephrased. The question of atmospheric memory is less substantiated in the paper, and hence will be moderated.

p. 5159, l. 6ff.: To provide an ensemble rather than one realization of a reconstruction or reanalysis is a very good and important argument and so is your ensemble reconstruction. Please mention briefly that the ensemble spread does only tell something about the uncertainty related to small changes in initial conditions at every time step. However, the ensemble spread does not necessarily represent the full uncertainty of the reconstruction if the underlying changes e.g. in the input data lead to a time-varying bias (spurious trends) of all ensemble members as recently shown for the mentioned reanalysis (Compo et al. 2011) by Krueger et al., (2013).

Right. We will mention that the ensemble spread is conditioned by the reconstruction method itself, and we have not investigated the sensitivity to the forcing SLP dataset from Kington (1988).

p. 5159, l. 11 ff.: The main motivation is a bit diffuse from the introduction relative to what is done in the study. Maybe you could improve this point by more clearly separating the different aspects like (i) why are the periods after Icelandic eruptions import/what was their impact, (ii) how you want to reconstruct it and (iii) how and why the analog method can be used to test the hypothesis that two specific weather periods (don't) share some similar characteristics.

Thank you for this. The text will be clarified accordingly.

p. 5160, l. 1ff.: Besides the type and strength (VPI) of volcanic eruptions, the magnitude and duration of local to global impacts are very dependent on tropical vs. extratropical origin. This should be shortly mentioned somewhere. In principal I do not see why there should be any similarity in terms of atmospheric circulation triggered by the Laki and Eyjafjallajökull eruptions. Nevertheless, I agree that the analog approach combined with evaluating the sampling frequency from a specific period is a very useful approach even though the Iceland volcanoes might not be an optimal example. Perhaps you could highlight and motivate the methodological aspect more clearly in the paper independent from your Laki test case.

We actually do not know whether those volcanic eruptions triggered the atmospheric circulation. As stated earlier, we just want to assess whether the atmospheric circulation after the Laki resemble the one after the Eyjafjallajökull, from the observations (or pseudo observations). We will clarify this in the text.

p. 5160, l. 23 ff.: Better use sth. like "Such unknown fields of the past are reconstructed using the historical SLP data by Kington (1988) to predict/search for analogous fields in NCEP reanalysis." Please enhance the introduction related to the method applied. At least the main idea and concept of the reconstruction method should be introduced already here. For readers unaware of the analog method, a short description of applications like Zorita and von Storch (1999), Schenk and Zorita (2012) and your own work (Vautard and Yiou 2009, Yiou et al., 2012) might help to understand the general idea behind the analog approach and how it can be used for your and other studies.

The comment from D. Wheeler and this one indeed prompt for a more precise description of the methodology. This will be done in the revised manuscript, in a methodological appendix.

p. 5161: Although gridded SLP is used later, could you shortly mention in the data section how the stations used by Kington (1988) are geographically distributed over the domain (roughly how many stations existed over regions like the North Atlantic, Scandinavia, central Europe etc.).

We will recall what Kington did in his book.

p. 5161, l. 1: "as our target reconstruction set". The "target" can be a bit misleading in this context. Perhaps you could use "predictor" in case of the Kingston SLP and predictant (target field) for the NCEP fields. Also in 2.3 the "base" = predictor. Maybe you could use both options as many readers might be familiar with the predictor/predictant definitions.

OK, we can use a more standard terminology.

p. 5161, l. 21-22: The usage of interpolated gridded historical SLP as predictor might be not an optimal choice here. I wonder whether the reconstruction could not be done from the Kington SLP station data without the interpolation or based on first leading EOFs of the SLP data series. Usually, the interpolation does not add additional information or it adds artificial features e.g. over data free areas over the North Atlantic. When searching for analogs, potential inadequate grid information might deteriorate the prediction of the analogs while this would be not the case when only using the station data directly. As shown e.g. by Guiot et al (2010) and Schenk and Zorita (2012), the aim of analog upscaling is to find analogous fields for sparse local observations rather than interpolating the local data onto gridded fields. These aspects should be shortly mentioned (see also p. 5164, l. 18-20). Would the station data be accessible to test that in a future study?

In principle, we totally agree with this remark. P. Yiou immediately asked P.D. Jones for the original data that J. Kington used to produce the gridded dataset (J. Kington is now retired from CRU, whose head is P.D. Jones). P.D. Jones stated that there are no computer files for those data, which only exist in paper form, and that anybody is welcome to visit Norwich to digitize them. Apparently, J. Kington drew maps like a weather forecaster would draw them, and someone at CRU digitized the maps. This means that this (interesting) recommendation cannot be done in a reasonable amount of time, with reasonable efforts. This also means that there are potential human errors (or uncertainties) in those maps. At present, the only thing that can be done is to clarify this caveat and explain that there is ample room for improvement.

p. 5161, l. 23 ff.: Is this the first time the temperature data has been digitized/used? How could interested readers get the data?

The data that are plotted will be made available, with proper documentation.

p. 5162, l. 11: I do not find any supplementary movie?

PY forgot to put a link to an animated gif file. It is available at (62Mb file):

https://cloud.lsce.ipsl.fr/public.php?service=files&t=cd64552c3de72b5156f7021dbc15f929

p. 5163, l. 13: What means "rapidly" in terms of number of days here? SLP has usually a quite high serial correlation on daily scale with a relatively flat slope compared to other variables. So wouldn't it be rather slow than rapid in the context of geophysical variables? In Yiou et al. 2012 you mentioned the sometimes occurring "score flaws" for the reconstruction in terms of pattern correlation and argue that this is a result of your applied continuity constraint (so w=5 in this study). This is consistent with the finding by Schenk and Zorita (2012) who show already a quite visible drop in the temporal daily correlation with w=5. Both studies conclude that the trade-off is caused by the fact that analogs optimized for several days contain less likely the best analog. Could you indicate how big the loss in temporal and also in serial correlation is for your reconstruction if you use the best analog (so comparing the w=1 NCEP field reconstruction with w=5)? I.e. if you aim at reconstructing rare specific events like conditions after the volcanic eruptions, the best analog could be more important than the improvement of the flow continuity potentially destroyed by the time-invariant analog reconstruction. These aspects should be also part of the discussion.

This will be clarified and discussed in the methodology and discussion. In particular, we will explain the loss of correlation between w=1 to w=5.

p. 5164, l. 9: "better scores in summer". As you correctly state in the next sentence, the reason is seasonality. Relative to the mean seasonal variance, the "better score" is more likely achieved for winter if you divide the RMS by the seasonal mean variance - the latter being much higher in winter than in summer. Please clarify the differences in seasonal skills also with respect to Figure 2.

OK, this will be explained in the text.

p. 5164, l. 18-20: Could the small region of optimization close to the land area be caused by the density/location of the Kington stations? This would underline my concerns regarding the use of interpolated SLP as predictor mentioned before. This should be discussed in the paper.

The Kington "original" time series are on European land (including Iceland). The Kington gridded dataset actually contains a lot of "empty" cells where there is no data. The calibration was done on a subregion of the Kington maps, i.e. over 5x5 cells with no missing data. The actual coverage of the gridded data can be given as supplementary material.

p. 5165, l. 27-28: I do not fully understand what you do here. Please explain.

Indeed, this is not clear. We computed monthly means of temperature anomalies for each analogue, and plotted the median of the 20 monthly means, between May 1783 and April 1784.

p. 5166, l. 8: temperature anomaly relative to which period? Please specify; see also figures.

Composites of temperature anomalies are determined from anomalies with respect to 1971-2000.

p. 5167, l. 1ff: As the historical temperature data was not published before, could you shortly indicate the cross-correlations of the single temperature time series with each other and with the average of all historical time series? This could help to quantify the comparison in Fig. 6 a bit more. You could also mention that the analog method can help to estimate the quality of historical time series or even fill in gaps (Guiot et al., 2010).

We will add a table with the cross correlations.

p 5168, l. 16ff.: "only anomalies" - I understand that you only want to compare anomalies for your case study. The disadvantage might be however that e.g. the potential thermal impact of Laki is to some extent lost by subtracting the mean of that short period. In principal, I don't see a problem here to reconstruct absolute SLP and absolute wind fields by using absolute SLP predictor data. The absolute SLP Kington data can be used to find closest analogous fields of absolute SLP from NCEP (apart from a potential systematic bias between the datasets). Due to the strong physical link to SLP, also absolute wind fields can be reconstructed although the skill will be lower in summer due to increased local to regional thermal influence (see Schenk and Zorita, 2012). I agree that in contrast the climatological mean for temperature cannot be inferred from SLP alone. However, you could use your monthly mean station data to predict analogous temperature fields if the absolute T2m observations are very similar to the absolute T2m values of NCEP. These different options should be mentioned here as they may motivate further applications of the analog method for the reconstruction of historical fields.

Indeed, everything that is directly related to pressure (wind, geopotential height) can have an absolute reconstruction, while this might not be the case for temperature or precipitation. This will be mentioned in the text, although we prefer to avoid mixing reconstructions of absolute and anomaly fields in this paper, which could be confusing.

p. 5169, 11ff.: Again, the generally differing impact of volcanoes should be shortly mentioned. So far, there is little evidence that extra-tropical volcanoes have a significant impact on the atmospheric circulation in contrast to major explosive eruptions in the tropics. On the other hand, climate models have also problems to realistically simulate the dynamical influence of tropical eruptions. This i.e. the case for high northern latitudes (e.g. Driscoll et al., 2012). So it might be also less clear for extra-tropical volcanoes which underlines the importance of your analog reconstruction as an alternative to reconstruct flow fields from observed analogs (reanalysis).

Results/Discussion: You mentioned in the introduction the health problems caused by aerosols after the Laki event and the cold winter in Europe. As it is difficult to say whether the winter would not have been cold anyway, would your daily reconstruction allow to say sth. on the existence of long-lasting high pressure episodes with very low wind speeds (i.e. winter)? Such conditions would be "favorable" for respiratory problems as aerosols could be trapped in the atmospheric inversion layer over wide areas and long time periods ("smog"). Maybe the result and discussion section could be extended a bit by comparing these kind of detailed weather information from the reconstruction and what the historians know from descriptions of that time.

Actually, the respiratory problems occurred in the Summer/Fall of 1783, not the Winter. Our paper does not treat this issue (although it was a motivation of the project that funded the study).

Figure 1: What are the circles standing for in the plot?

The circles are for the outliers. Classically, the upper (lower) whiskers of the boxplot are the minimum (maximum) between 1.5 times the interquartile range and the maximum (minimum) value. The circles indicate that there are skewed distribution tails.

Figure 2: " ...SLP anomalies and 20 best analogues from RMS, on monthly averages."

Do you mean monthly average of the daily correlation or the correlation for the monthly means calculated from the daily reconstruction? Same question afterwards for " for each month, for 20 analogues and daily reconstructions." Please clarify.

The daily correlation is computed for each analogue. Then we report the average of daily correlations for each (or all) analogue and each month. This will be clarified.

The Analog 1, 10 and 20 are meant here as 1 being the best and 10 and 20 being the n-1 next neighbor to the best?

Yes: Analog1 is the best (in term of distance). This will be clarified.

Figure 4: Please add the unit for the scale (anomalies in K) and the period, for which the mean has been subtracted (are the anomalies relative to the monthly means of 1781-1785?).

The legend will be corrected. The text will explain what temperature anomalies mean (see your comment above).

Figure 5: Again, anomalies relative to what. If the red line shows also anomalies, just use K for all information in the plot. The font size might be too small in some figures.

The font size will be increased to improve readability.

Figure 6: Please use K instead of C as this is standard for temperature deviations if you don't refer to absolute temperatures in C. Do you mean with "median analogue reconstruction for France" the spatial median of the field over France? Please specify.

Figure 6 is a reconstruction for France. The legend will be corrected with degrees in K.

Technical corrections:

The technical corrections will be implemented.

p. 5159, l. 14: large

p. 5166, l. 4: October

p. 5166, l. 7: missing "into account" at the end of the sentence

p. 5168, l. 8ff.: "We have applied a method of analogues of circulation..." sounds strange.

This will be rephrased to something like: "we have proposed an ensemble of atmospheric field reconstructions, that is compatible with an estimate of sea-level pressure".

p. 5175, figure caption: "i" missing in distribution

References:

Driscoll, S., A. Bozzo, L. J. Gray, A. Robock, and G. Stenchikov (2012), Coupled Model Intercomparison Project 5 (CMIP5) simulations of climate following volcanic eruptions, J. Geophys. Res., 110(D17), D17105, doi:10.1029/2012JD017607

Guiot, J., Corona, C., and ESCARSEL members (2010): Growing Season Temperatures in Europe and Climate Forcings Over the Past 1400 Years, PLoS ONE 5(4): e9972, doi:10.1371/journal.pone.0009972

Kravitz, Ben, and Alan Robock, 2011: The climate effects of high latitude volcanic eruptions: The role of the time of year. J. Geophys. Res., 116, D01105, doi:10.1029/2010JD014448.

Krueger, Oliver, Frederik Schenk, Frauke Feser, Ralf Weisse, 2013: Inconsistencies between Long-Term Trends in Storminess Derived from the 20CR Reanalysis and Observations. J. Climate, 26, 868–874, doi:10.1175/JCLI-D-12-00309.1

Oman, Luke, Alan Robock, Georgiy Stenchikov, Gavin A. Schmidt, and Reto Ruedy, 2005: Climatic response to high latitude volcanic eruptions. J. Geophys. Res, 110 (D13), D13103, doi:10.1029/2004JD005487

Schmidt, A., T. Thordarson, L. D. Oman, A. Robock, and S. Self (2012), Climatic impact of the long-lasting 1783 Laki eruption: Inapplicability of mass-independent sulfur isotopic composition measurements, J. Geophys. Res., 117, D23116, doi:10.1029/2012JD018414.