Answer to review #1 of manuscript: "Could old tide gauges help estimate past atmospheric variability?" for *Climate of the Past*

Paul Platzer, Pierre Tandeo, Pierre Ailliot, Bertrand Chapron

March 2024

Reviews in black, answers in blue, ***quotes from revised manuscript in bold.***

Summary: The study explores the possibility of using tide gauge records to reduce uncertainty in reconstructing past (19th century) sea-level pressure fields. The authors focus on the 20CR reanalysis, a long reanalysis product covering parts of the 19th century, and two tide gauge records in Northwestern France. The main conclusion is that the tide records can provide valuable information to constrain the 20CR reanalysis ensemble by attaching lower probabilities to the members of the ensemble that are less compatible with the implied sea level at those two tide records.

Recommendation: I think the idea of the study is interesting, and the study is, to a large extent, technically sound. I have some suggestions that the authors may want to consider in a revised version. My recommendation of 'major revisions' is more dictated by my interest in the revised version. The needed revisions are, in my opinion, between 'minor' and 'major'

We thank Reviewer #1 for the fruitful remarks and suggestions. The discussion has led to major improvements of the paper and we believe that the revised manuscript is more clear and precise.

Following also the remarks and suggestions of reviewer # 2, and in order to make the reviewed version of the paper more clear, we have made several modifications to the paper, which are listed below to make it clear to both reviewers.

i. Pre-processing of the data has changed. The effect of mean-sea-level variations on surge was removed by removing a yearly median value of the surge instead of a time-constant + time-linear trend. The pressure used to assess the relationship with the sea-level is not the anomaly with respect to a given climatology anymore, but the difference between the local sea-level pressure at Brest and the sea-level pressure averaged over the whole North-Atlantic ocean.

ii. The statistical relationship between sea-level and pressure is now estimated with a new model. First, this model is not linear, but local-linear. Is does not include the effects of pressure gradients (winds) anymore. This time we model pressure as a function of sea-level (the true objective of the paper), while in the original submission we modeled the sea-level as a function of the pressure and pressure gradients (winds). More details are given in section 3.1. of the revised manuscript.

iii. As a consequence, section 3.2. was ruled out.

iv. We kept only the tide gauge record of Brest, and we are not using the one of Saint-Nazaire anymore. In the way the algorithm was implemented and with our simple statistical relationship, using two tide gauges at a time was barely more informative than using just one. Therefore, for simplicity purposes we kept only the Brest record which is the longest one.

v. For a better interpretation of applying our method in the 19th century, we have used independent pressure data for the city of Brest from the EMULATE project and recently release archive data from Météo-France. This data starts in 1855, while the Brest tide gauge starts in 1846. This data also greatly helps the interpretation of Figures 9, 12, 13 and 14 of the revised manuscript which shows specific events.

General points

1. This is somewhat a cosmetic comment, but the the manuscript refers to 'storm surges' whereas actually, the study considers the sea-level record after filtering out the tides and the long-term sea-level rise. In my understanding,

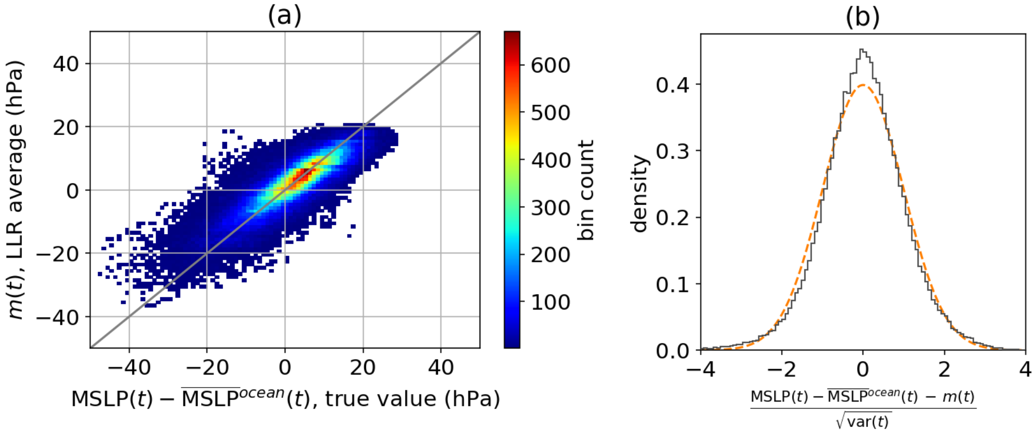

Figure 4 in revised manuscript. **Evaluation of the local-linear regression (LLR) on the period 1981-2015 using a leave-one-out procedure. (a) Histogram scatter-plot of the average estimate from the LLR versus true value of the MSLP difference with the reference ocean-averaged MSLP. (b) Density-histogram of normalized average error of the LLR estimate (black line) compared to theoretical probability density function of standard Gaussian random variable (dashed orange line).**

the residuals are not the 'storm surge' component. The storm surge is, by definition, caused by the passage of a storm. The mechanisms connecting the storm atmospheric forcing and the sea surface elevation are more complex in cases of storms than those for the 'normal' sea level. For instance, depending on the coastline topography, the wind may cause a direct elevation of sea-level in the direction of the wind. This is clear for estuaries, and it is the main cause of storm surges in many North Sea locations. The study statistically analyses the full range of sea-level variations, not only the extremes, and it is unclear why the text highlights 'storms'. This is only the cause for the last section, and this is related to some additional problems (see next point).

1. We have modified all occurrences of "storm surge" with simply "surge" to account for this comment.

2. The studies used a linear statistical regression method to link atmospheric predictors and sea-level elevation (after tidal and long-term trend filtering). The skill of the linear model is shown in Figure 4. It is visually clear that the linear model underestimates sea-level extremes, both high and low. This is very usual for linear models. But, while using tide gauges to constrain the reconstructed SLP would be permissible for 'normal' situations, it is clearly problematic for extremes. Thus, the last section needs a bit more attention, in my opinion. It could very well happen that the SLP from the ensemble members of the 20CR reanalysis passed through the linear model underestimates the storm surge (in this case, it is indeed a storm surge), and therefore, more ensemble members appear not compatible with the sea-level observations than they are. I suggest testing this approach with a recent storm when the 20CR has assimilated sufficient SLP observations to be considered accurate. In that recent storm, if I am correct in my concern, the study's approach would also find that the 20CR ensemble is biased towards implied lower storm surges. This would test the study's main claim, namely that 'storm surges' can constrain the 20CR reanalysis.

Even if the test results fall against the study's claim, the study itself can still be useful, albeit in a modified form. Long records would still be useful to constrain the 20CR, but this constraint in case of storms would be too restrictive, attaching too low a probability to too many ensemble members of the 20CR reanalysis. Alternatively, a non-linear model should be set up to more accurately estimate extreme-level from atmospheric forcing.

2. This is a very important remark and we thank reviewer #1 for pointing this out.

Indeed, using a linear model induces bias in the extremes. In order to partially remedy this problem, we have chosen to slightly change the relationship used in the study. Now, we use a local-linear regression (LLR) to predict the local MSLP in Brest from the observed surges. This was motivated by your remark, and by the fact that we observed a deviation from the linear relationship which depends on both the value of the 12h-averaged surge (see Fig. A) and on the value of the 12h-difference of the surge (see Fig. B).

As obvious from Fig. 4 of the revised manuscript (reproduced here), our local-linear model still induces biases in the extremes, although it is less biased for moderate values while the linear model was. However, we believe that this bias is less significant for our application purposes than the one induced by the ambiguity of the pressure reconstruction from surges due to the effects of wind. Looking at Fig. 9 of the revised manuscript (reproduced here), it is obvious that the surge-based LLR may actually either overestimate or underestimate the amplitude of pressure variations, depending on the amplitude and direction of the wind stress at the time of the storm. Therefore, we choose to still use a model that is biased in the extremes.

We expect the hypothesis of reviewer # 1 to be validated in case of typical (i.e. average) wind conditions: the

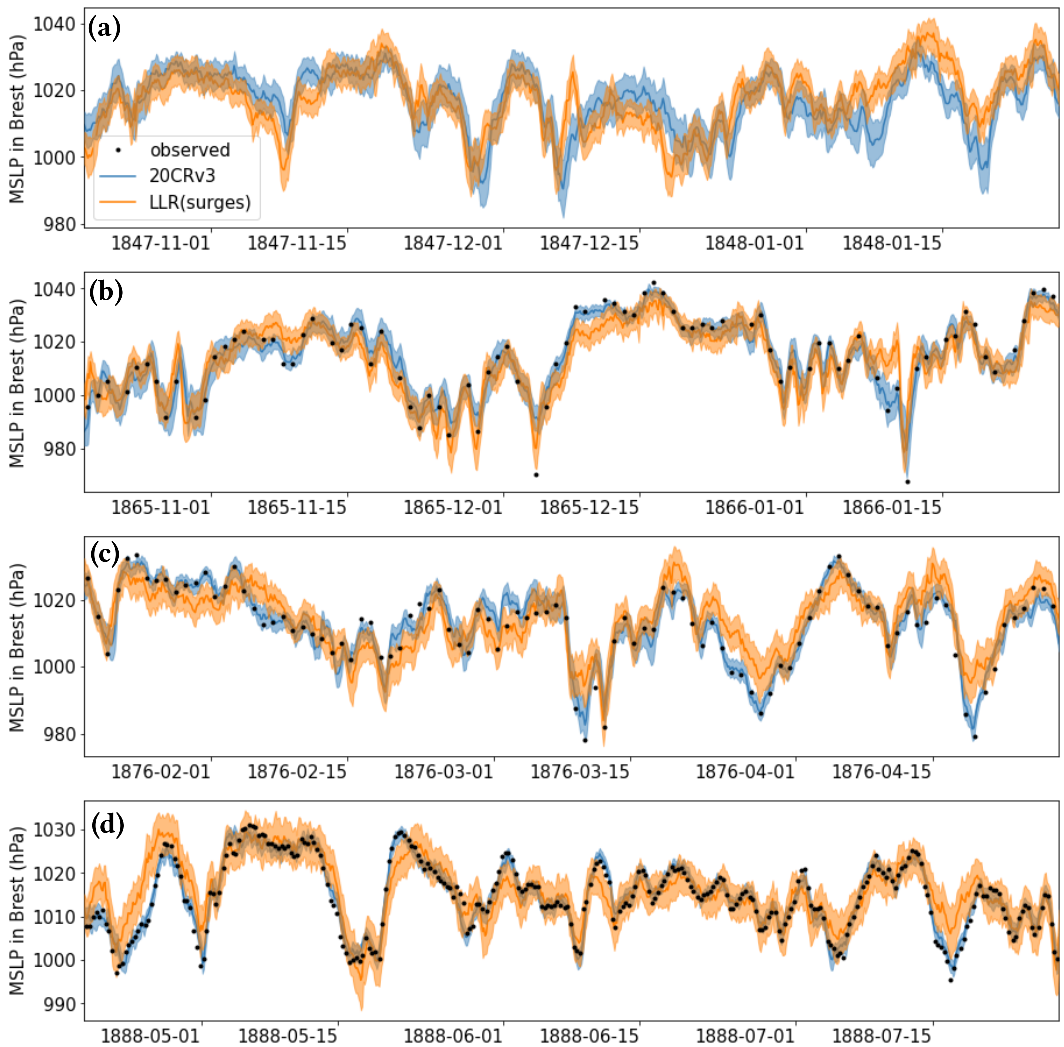

Figure 9 in revised manuscript. **Comparison of MSLP estimation in Brest from 20CRv3 (blue), LLR based on surges (orange), and independent observations (black dots) that were not used to build the orange and blue curves, for three periods surrounding the events studied in this section. Full lines correspond to average values while shaded areas correspond to ± one standard-deviation around the average.**

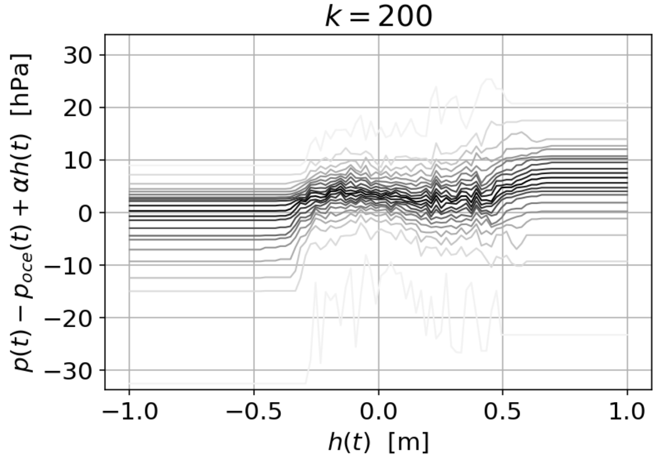

Figure A. Not included in revised manuscript. Quantiles [0. , 0.05, 0.1 , 0.15, 0.2 , 0.25, 0.3 , 0.35, 0.4 , 0.45, 0.5 , 0.55, 0.6 , 0.65, 0.7 , 0.75, 0.8 , 0.85, 0.9 , 0.95, 1.] of error of estimation from a simple linear regression of MSLP(t) based on 12h-averaged surge $h(t)$, as a function of $h(t)$, estimated by selecting the 200 closest values of $h(t)$ in the period 1981-2015. This figure is not part of the revised manuscript and is here only for discussion purposes.

LLR estimation of the most extreme pressure variations would be biased towards the mean value of MSLP, and our algorithm would select the members with the smallest MSLP deviation from the mean. We have not performed the test directly, but the discussion of section 5. of the revised manuscript sheds light on the limitations of our method line 368:

Our claim that the wind variations are responsible for the persistent biases between the LLR pressure estimation and the reanalysis is supported by Fig. 13.f, where we also show the direction and amplitude of the 10m-wind intensity as given by the average over all reanalysis members and interpolated at the city of Brest. In March 1876, two low-pressure systems passed to the North of Brest's tide gauge, one around March 10th and a second around March 12th, as indicated by the reanalysis members and the independent pressure observations (Fig. 13.e). However, the first low-pressure system did not induce a surge as strong as the second one. One key difference between the two events is the wind amplitude, which reached 15m/s during the first event and then decreased to 5-10m/s during the second event, with almost steady wind direction. Although wind intensity and direction estimated from the reanalysis must be taken with care, the value of 15m/s is rarely exceed (only 7 in 1000 times in the period 1981-2015, not shown), indicating exceptional wind intensity during the event, and justifying the inaccuracy of the LLR which is based on already observed events and therefore biased towards typical wind conditions. Our interpretation relies on the fact that the effect of wind on extreme surges acts at small time scales (daily or sub-daily), which is backed by recent work (Pineau-Guillou et al., 2023).

We emphasis these limitations but still believe that the study can be useful if improved, or that the methodology can be used for other variables than sea-level data.

Minor comments *(→ We added letters to allow cross-reference between reviews.)*

A. 'Anomalies are considered with a reference climatology computed as an average over all members and over ±30 calendar days, ± 3 day hours '.

This sentence is unclear.

This was modified in the revised manuscript. Moreover, since we are using a second type of preprocessing of MSLP, as suggested by reviewer # 2, we have added a whole subsection entitled "Preprocessing of mean-sea-level pressure" which is copied here (line 90)

In this work, we are using only the mean sea-level pressure (MSLP) variable from 20CRv3. We make two different preprocessings of this variable.

A first preprocessing is used for the statistical relationship between the local pressure and the surge. As the latter is driven by a physical phenomenon called the "inverse barometer effect" which will be introduced in the next section, we consider the difference between the MSLP interpolated at the city of Brest (4.49504°W, 48.3829°N), and the MSLP averaged over all members of 20CRv3 and over the North-Atlantic ocean (using the reanalysis' land mask and averaging from 98°W to 12°E and from 0°N to 69°N), similarly to Ponte (1994). This spatial-averaged pressure is noted $\overline{\text{MSLP}}^{ocean}(t)$ **and depends only on time.**

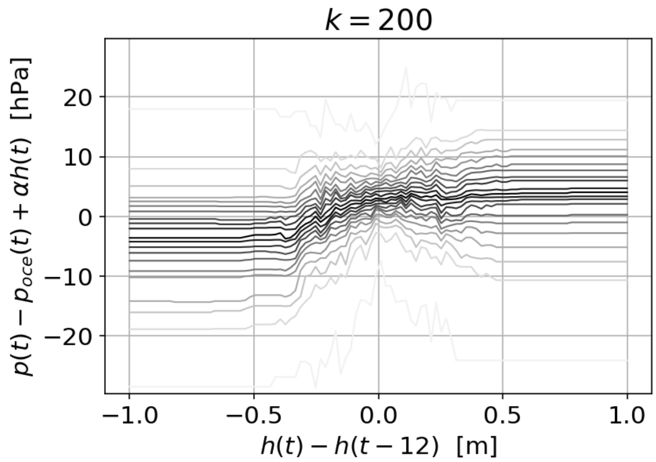

Figure B. Not included in revised manuscript. Same as previous figure but as a function of the difference in surge $\Delta\overline{h}^{3h}(t)$ as defined in the revised manuscript. This figure is not part of the revised manuscript and is here only for discussion purposes.

A second preprocessing of MSLP is used to compute the probability of transition from one member of the reanalysis to another in the Hidden Markov Model (HMM) presented in section 3.2. For this purpose, we consider seasonal anomalies of MSLP with respect to a climatology computed from the period 1847-1890, because the HMM is run only for those years. The reference MSLP climatology for calendar day d and hour h is given by the average over days between $d - 30$ and $d + 30$, hours between $h-3$ and $h+3$, and all years 1847-1890. This reference MSLP is noted $\overline{\mathrm{MSLP}}^{clim}$ and depends on latitude and longitude.

B. Table 1, Table 3 units are missing

These tables were removed as we do not use the linear regression anymore (see description above).

References

Lucia Pineau-Guillou, Jean-Marc Delouis, and Bertrand Chapron. Characteristics of storm surge events along the north-east atlantic coasts. *Journal of Geophysical Research: Oceans*, 128(4):e2022JC019493, 2023.

Rui M Ponte. Understanding the relation between wind-and pressure-driven sea level variability. *Journal of Geophysical Research: Oceans*, 99(C4):8033–8039, 1994.