

We thank the referees for their constructive criticism and helpful remarks and we will address their points in a revised version as indicated below. The referees' questions are replicated in italics, followed by our response in regular font.

Referee 1

1) I suggest the author further explore the mechanistic relationship between the wind and mean sea level in the Dutch Wadden Sea at interannual time scale

We address this point in the revised version by elaborating on what we already mentioned in Section 3.1 with regard to the shorter time scales of a tidal period. The *mechanistic* relationship between wind and surges or depressions of sea level comes into play at the time scales of *hours* and models are well able to capture this (e.g., Zijl et al. 2013 and Duran-Matute et al. 2016). In an annual-mean sense, a relation is still found (as shown by this paper and earlier studies that we mention), but this simply reflects the aggregate of the mechanistic connections that occur on a much shorter time scale of hours.

2) To help readers have a better understanding of the setting, it would be beneficial if the authors can plot the climatological annual mean wind and/or sea level from observation or reanalysis on top of the current map.

About the wind one may say that the long-term dominant wind direction is SW/W (see Fig. 4a), but to plot "the" annual mean wind or annual mean sea level is not really possible precisely because the values vary so much from year to year, as figures 4b and Fig 2 show. We have however added in Fig 1 the location of the reanalysis data that we will analyze in response to a suggestion by Referee 3.

3) I suggest the authors describe the data sets used in the present study and methods in more details, so that readers can reproduce results in the manuscript. For example, the temporal resolution and any preprocessing of the sea level data should be provided in Section 2.1. Also, more details should be provided on how the confidence interval and the (effective) sample size is calculated.

We have provided more details on the underlying data and methods. In Section 2.1: The sea level data was obtained from the PSMSL website as monthly mean values. For the confidence interval, we adopted a method from the reference mentioned in our paper (Montgomery and Runger, 2003); we now mention the formulas. In Section 2.2: the webpage is mentioned from which the wind data was obtained.

4) Figure 5 shows the annual mean wind energy at weather station Vlieland, based on which the authors conclude that the annual mean wind energy for all directions combined stays fairly constant throughout the 20 years. I find this statement unconvincing, at least with current analysis. I suggest the authors replot their Figure 5 with a better (smaller) vertical scale. The current one is misleading. The authors further conclude that "it is not so much the total wind energy that varies between years, but rather the share each of the directions gets from this total" (Page 8, Line 5). This statement is also superficial and unsubstantiated.

We have substantiated our statements in more detail in Section 2.3, based on Figs 4 and 5. Specifically, we calculated the relative standard deviation (standard deviation divided by the mean, and expressed as a percentage). This gives, for each of the individual directions: 28% (N), 34% (NE), 49% (E), 26% (SE), 36% (S), 25% (SW), 26% (W), 26% (NW). Doing the same for the total energy (all directions summed), we find 13%, which is much smaller than the number for any of the individual directions. These numbers bear out our statement that "it is not so much the total wind energy that varies between years, but rather the share each of the directions gets from this total" and confirm what is already visually reflected in Figs 4b and 5.

5) When conducting multiple regression in Section 3.4, the authors treat the west-east and south-north winds and the annual mean atmospheric pressure as three independent variables. This can be easily examined by calculating their correlation at the interannual (or yearly)

timescale.

We checked this. Based on the 20-yr record, the correlation between WE and SN wind-energy levels is +0.46; between pressure and WE winds, -0.03; and between pressure and SN winds, -0.39. In the first and third cases, there is a significant, although not very high, correlation. We emphasize that the inclusion of SN winds and pressure anyway contributes very little to the reconstruction of mean sea level, since it is already largely explained by the WE contribution alone.

6) *With the 20-year data, year 1996 is very different from other years and looks like an outlier in a statistical sense. I wonder how the results (statistics) change if data from 1996 is removed. This practice is useful for checking the robustness of the present findings, especially considering the limited sample size of the current data.*

The year 1996 was an exceptional year in this record because of the anomalously strong winds from the E and weak SW/W winds (as already discussed in Section 2.3). However, it is *not* an outlier in the relation between WE winds and annual mean sea level, as is demonstrated by Fig 7, where 1996 is represented by the point in the lower left corner but does not stray away from the general trend indicated by the grey line. Consequently, that line barely changes if the year 1996 is excluded.

We have corrected the small typos.

Referee 2

We thank the referee for his/her positive appraisal that this study contributes to the understanding of regional climate change.

1) *A mechanism is needed of winds on inter-annual variability sea level variability in this particular small region. The analysis so far is a good start, whereas the authors need to explore the physical processes and mechanism via a regional climate model.*

The model used in this paper (Section 4), was already used in earlier studies to examine the effects of wind on residual circulation and sea level (Duran-Matute et al 2016, see also Duran-Matute et al 2014, both referenced in our paper). The model was forced by wind stress at the surface obtained from reanalysis data (we now mention this in Section 4). As we emphasized already in response to Referee 1, Q1, the *direct causal mechanistic* relation between wind stress and sea level plays on the time scale of *hours* (or days at most), since the sea surface responds rapidly to this forcing. In Duran-Matute et al (2016), Fig 3a, a clear relation is therefore found in a tidal-average sense. In an annual-mean sense, a relation is still found (as shown by this paper and earlier studies that we mention), but this simply reflects the aggregate of the mechanistic connections that occur on a much shorter time scale of hours.

2) *Is this regional change of winds and sea level related to large-scale atmosphere and ocean changes? What is the atmosphere circulation pattern associated with the wind change over the Dutch Wadden Sea? Some effects from the westerly jet shift or ENSO?*

We briefly mentioned already that there is no obvious connection in our 20-yr record between total wind energy and NAO or ENSO (section 2.3), but we now discuss in the Introduction in more detail what other studies have found, notably Frederikse et al. (2016), who identified a small contribution from NAO at some tide gauges.

Referee 3

We thank the referee for his/her appraisal of our paper as being very well written and the appreciation of the novel feature of using measured atmospheric data in this sea-level variability study.

1) *The issue of the standard error size was not first mentioned by Zervas (2009), but already discussed in Douglas (1991) and in particular acknowledged for the North Sea including Dutch stations (and the effect of an atmospheric correction) in the already referenced literature.*

Zervas (2009) conveniently summarized in a graph the standard error as a function of the length of the time series. There appears to be no similar figure or table in the paper by Douglas (1991), although he does of course mention standard errors in specific individual cases. However, we agree that it is important to add the reference to Douglas and have done so in the first paragraph of the Introduction.

2) *One of your foci is on the important issue of uncertainty of trend estimates. However, I am missing information on how you calculated the uncertainties. Did you calculate them based on uncorrelated noise? If yes, this needs to be corrected in a revised version of the ms. As shown by Dangendorf et al. (2015, <https://www.nature.com/articles/ncomms8849>) and many other studies in the recent years, sea level records show a high degree of autocorrelation, which can only be accurately accounted for by adjusting the degrees of freedom. While traditionally this has been undertaken using a simple AR1 process, Dangendorf et al. (2015) have shown that in case of the North Sea this results in a serious underestimation of the true uncertainty. Hence, I suggest that the authors should at least use an ordinary AR1 process (since they use annual data, see Bos et al. <https://academic.oup.com/gji/article-lookup/doi/10.1093/gji/ggt481>) in describing the uncertainty and discuss the underestimation related to that.*

As indicated in Section 2.1, we calculate the 95% confidence interval on slope as explained in the statistics textbook by Montgomery and Runger (2003). We now provide the formula in the revised version of our paper, for the convenience of the readers (Section 2.1, second paragraph). Linear regression, as we used it, is a very commonly applied method in the recent literature (see, eg, Frederikse et al. 2016). Moreover, our 100-yr annual mean trends Table 1 are fully in agreement with other recently published results by Wahl et al. 2013 (as we mention in Section 2.1). We also note that the merits of Detrended Fluctuation Analysis appear to be a matter of debate (Bryce and Sprague, Sci. Rep. 2, 315; DOI:10.1038/srep00315 (2012). However, we added a comment (Section 2.1, second paragraph) noting that alternative methods have recently been proposed.

Finally, we note that this paper has not as its primary goal to determine trends but rather to examine the regional variability of the sensitivity of annual mean sea level to wind climate.

3) *One of the novel aspects in the ms is the use of wind and pressure observations. However, I am missing a comparison to equivalents from reanalysis data to demonstrate the real effect. This could be, for instance, made by comparing the amount of explained variability for the different products. I would suggest that the observations should probably show a bit better agreement than reanalysis products.*

We thank the reviewer for this valuable suggestion; in the revised version of this paper, we included a comparison with reanalysis data (Section 5.5). The analysis indicates that the results are qualitatively similar but that the measured atmospheric data from the weather station provides a slightly more accurate reconstruction of annual mean sea level.

4) *Along with point three I would like to suggest making use of one the available tide+surge model outputs for the North Sea, which should give in general a proper description of the effects of atmospheric forcing. You could for instance compare your estimates using the same reanalysis wind/pressure forcing as in the model and compare the results.*

In Section 4, we already included model results and in the revised version we now explicitly mention that the model was in fact forced at the surface by meteorological reanalysis data. So,

the comparison between modeled and measured inter-annual sea level variability already involves the use of reanalysis data.

5) *Freshwater Discharge: The authors suggest that once of the missing components could be freshwater discharge. While I agree that this could be one particular factor, this is incomplete. The first factor, which comes into my mind, is the steric effect (including its ocean bottom pressure component), which has been estimated, for instance, by Frederikse et al. (2016).*

We rephrased the sentences on the influence of freshwater discharge to emphasize that we are considering *inter-annual* and *small-scale regional* variabilities in this paper (Section 5.4, last paragraph). It would seem to us that global steric effects can barely be significant in either of these variations, because they act on longer temporal and larger spatial scales. This contrasts with local and seasonal variations of freshwater discharge.