## *Original*

Dangendorf, S.; Mueller-Navarra, S.; Jensen, J.; Schenk, F.; Wahl, T.; Weisse, R.:

**North Sea storminess from a novel storm surge record since AD 1843**

In: Journal of Climate (2014) AMS

DOI: 10.1175/JCLI-D-13-00427.1

# North Sea Storminess from a Novel Storm Surge Record since AD 1843*

SÖNKE DANGENDORF,[+] SYLVIN MÜLLER-NAVARRA,[#] JÜRGEN JENSEN,[+]
FREDERIK SCHENK,[@] THOMAS WAHL,[&] AND RALF WEISSE**

[+] *Research Institute for Water and Environment, University of Siegen, Siegen, Germany*
[#] *German Maritime and Hydrographic Agency (BSH), Hamburg, Germany*
[@] *Institute for Coastal Research, Helmholtz Zentrum Geesthacht, Geesthacht, Germany, and Linné Flow Centre, Department of Mechanics, Royal Institute of Technology, Stockholm, Sweden*
[&] *College of Marine Science, University of South Florida, St. Petersburg, Florida, and Research Centre Siegen (FoKoS), University of Siegen, Siegen, Germany*
** *Institute for Coastal Research, Helmholtz Zentrum Geesthacht, Geesthacht, Germany*

(Manuscript received 19 July 2013, in final form 6 January 2014)

ABSTRACT

The detection of potential long-term changes in historical storm statistics and storm surges plays a vitally important role for protecting coastal communities. In the absence of long homogeneous wind records, the authors present a novel, independent, and homogeneous storm surge record based on water level observations in the North Sea since 1843. Storm surges are characterized by considerable interannual-to-decadal variability linked to large-scale atmospheric circulation patterns. Time periods of increased storm surge levels prevailed in the late nineteenth and twentieth centuries without any evidence for significant long-term trends. This contradicts with recent findings based on reanalysis data, which suggest increasing storminess in the region since the late nineteenth century. The authors compare the wind and pressure fields from the Twentieth-Century Reanalysis (20CRv2) with the storm surge record by applying state-of-the-art empirical wind surge formulas. The comparison reveals that the reanalysis is a valuable tool that leads to good results over the past 100 yr; previously the statistical relationship fails, leaving significantly lower values in the upper percentiles of the predicted surge time series. These low values lead to significant upward trends over the entire investigation period, which are in turn supported by neither the storm surge record nor an independent circulation index based on homogeneous pressure readings. The authors therefore suggest that these differences are related to higher uncertainties in the earlier years of the 20CRv2 over the North Sea region.

## 1. Introduction

Storm surges represent a serious hazard for coastal areas and are expected to become more severe in a warming climate (von Storch 2014) because of the effects of rising mean sea level (MSL; Slangen et al. 2012) and potential changes in regional wind fields (Woth et al. 2006). Global MSL has risen through the last century (e.g., Church and White 2011) and is expected to rise also through the twenty-first century, potentially at an accelerated rate (e.g., Slangen et al. 2012), shifting the entire distribution of extreme sea levels on a higher base level (Hunter 2010). MSL changes in the North Sea region were recently reviewed by Wahl et al. (2013). While an ongoing sea level rise is evident, changes in atmospheric circulation and storminess are presently more uncertain (Weisse and von Storch 2010, and references therein). This is partly because the detection of past changes in storminess is often hampered by inhomogeneous or biased wind measurements (e.g., Lindenberg et al. 2012). Hence, the scientific community proceeded to the evaluation of more homogeneous storminess proxies, which have been observed over longer time periods. Typical examples for such proxies are storm indices calculated from single station pressure readings (e.g., Bärring and von Storch 2004; Hanna et al. 2008), high annual percentiles of geostrophic winds derived through

* Supplemental information related to this paper is available at the Journals Online website: http://dx.doi.org/10.1175/JCLI-D-13-00427.s1.

*Corresponding author address:* Sönke Dangendorf, Research Institute for Water and Environment, University of Siegen, Paul-Bonatz-Str. 9-11, 57076 Siegen, Germany.
E-mail: soenke.dangendorf@uni-siegen.de

DOI: 10.1175/JCLI-D-13-00427.1

triangulation of pressure readings (Schmidt and von Storch 1993; Alexandersson et al. 1998, 2000), or storm surge records from tide gauge measurements (von Storch and Reichardt 1997; Zhang et al. 2000).

The North Atlantic European sector is probably the most intensively studied and discussed region for that topic. From analyzing annual geostrophic wind statistics back to 1876, Schmidt and von Storch (1993) found pronounced decadal variability in the southern North Sea but no evidence for a significant long-term trend. These results were confirmed for the larger North Atlantic–European region—for example, by Alexandersson et al. (1998, 2000) and Matulla et al. (2008)—with the common finding that storminess was high at the end of the nineteenth century and subsequently declined until about 1960, followed by a strong upward trend until the mid-1990s. Since then, up until now, a return to average conditions is evident.

In contrast to annual storm statistics, Wang et al. (2009) pointed to differences in seasonal 99th geostrophic wind percentiles linking the high values at the end of the nineteenth century to a summer maximum and the high values in the 1990s to increasing storminess during winter. Apart from these seasonal differences over the northeast Atlantic and North Sea region, Wang et al. (2009) confirmed the absence of any robust long-term trend for high annual percentiles of geostrophic wind speeds since 1874. The study showed in addition that the periods of high storm activity furthermore coincided with positive decadal trends in the North Atlantic Oscillation (NAO) index (Hurrell 1995). A similar NAO link and strong decadal trends are also visible in extreme sea levels in the southeastern North Sea (Dangendorf et al. 2013a; Mudersbach et al. 2013).

Using data from the Twentieth-Century Reanalysis (20CRv2; Compo et al. 2011), Donat et al. (2011b) detected—contrary to observational analyses—significant upward long-term trends in the occurrence of extreme storms based on daily wind speeds since 1871 over Europe and suggested that the increase could (at least partly) be a response to enhanced greenhouse gas emissions during the past decades. Brönnimann et al. (2012) showed that the variance of the 20CRv2 ensemble increases back in time, leading to a better representation of trends after 1950 than before. Krueger et al. (2013b) highlighted that the upward trends detected by Donat et al. (2011b) are inconsistent with observations and suggested this being mainly an artifact of assimilating less surface pressure data into 20CRv2 back in time, leading to larger inconsistencies before 1940. This finding is currently controversially discussed (Krueger et al. 2013a; Wang et al. 2014) partly because of the usage of different storminess measures but also inhomogeneous observations, which were still present in the Krueger et al. (2013a,b) study. By removing some potential inhomogeneities in the pressure records, Wang et al. (2014) were able to bring the storm indices from 20CRv2 and from observations closer together but fail to remove the major inconsistencies indicating that there still might be severe homogeneity issues in the early years of the (i) 20CRv2, (ii) observations, or (iii) both.

In the light of these competing results, the aim of the present study is to analyze an independent and homogeneous storm surge record from the tide gauge of Cuxhaven located in the southeastern North Sea, covering the period from 1843 to 2012. Since storm surges are generated by low atmospheric pressure and intense winds over the ocean, surges generally exhibit a comprehensive, independent, and more homogeneous archive (Zhang et al. 2000) of information about storminess. Although some tide gauge records reach back into the seventeenth and eighteenth centuries, storm surge records have been surprisingly rarely analyzed in the last decades and earlier assessments in the North Sea region just focused on the twentieth century (Ullmann and Monbaliu 2010). This is mainly attributed to the limited availability of continuous hourly measurements, which are required for a harmonic analysis in order to remove the deterministic tidal water level components from the tide gauge data (Pawlowicz et al. 2002). In this study, we use a method first introduced by Horn (1948, 1960) that allows us to reconstruct a homogeneous storm surge record based on tidal high and low water levels only instead of hourly data, which are just available after 1918. This enables us to extend the storm surge record in Cuxhaven back to 1843, enhancing information about storminess of conventional proxies in that region (Schmidt and von Storch 1993) by more than 30 yr. The method is globally applicable and will open potential for the assessment of storminess. Here, we introduce, to our knowledge, the longest contemporary storm surge record in the world and (i) analyze long-term trends in the upper percentiles of storm surges, (ii) investigate the relationship between storm surges and large-scale atmospheric patterns, and (iii) use the empirical relationship between local winds and sea level pressure (SLP) to compare observations with the 20CRv2 data.

## 2. Data and methods

Tide gauges are well suited for the assessment of storminess as they measure, beside tides and longer-term MSL changes, the direct response of the ocean to the atmosphere. Especially in the North Atlantic region, some gauges have measured sea level for hundreds of years (Amsterdam, the longest record, starts in 1682;

Woodworth et al. 2011). Here, we focus on a long sea level record observed at the tide gauge of Cuxhaven, which is located in the southeastern North Sea. Sea levels have been observed in Cuxhaven since 1843. Until 1899, the measurements were taken as readings of high and low water levels 4 times per day. Since then continuous curves were registered on tidal charts and only the peaks were handwritten in logbooks. Digital recordings are available since the mid-1990s. To get the full information about the tidal curves before that time, extensive digitization works at the German Maritime and Hydrographic Agency made hourly data available back to 1918. In the present study we combine both data types (i.e., tidal peaks and hourly observations) to get insights into the history of all available measurements. The datasets were carefully checked for outliers, datum shifts, missing values, and time drifts in earlier studies (Wahl et al. 2010, 2011). The overall quality was found to be good. For example, Dangendorf et al. (2013a) demonstrated a high coherence of sea levels between the Cuxhaven record and 12 additional records all located in the German Bight.

Storm surges measured in the region can be decomposed into an external and a local/regional component; both caused by meteorological disturbances over the northeast Atlantic and the entire North Sea basin, respectively (Müller-Navarra and Giese 1999). Hence, it is reasonable to suggest that strong storms occurring over the northeast Atlantic or North Sea region will leave a fingerprint in the surge record of Cuxhaven. To illustrate the genesis of such a storm event, Fig. 1 shows the meteorological and oceanographic situation in the region during January 2007. In this month, Northern Europe was affected by a series of strong storms (Fink et al. 2009), leading also to a series of strong storm surge events (Fig. 1). In total the long-term 95th percentile of daily surges was exceeded eight times in only 22 days. The genesis of storm surge events often starts with a low pressure system over the North Atlantic traveling westward to the larger Baltic and Scandinavian area. On their way such pressure systems may trigger waves in the deep ocean northeast of Scotland, which then propagate into the North Sea elevating the water levels in the German Bight approximately 15 h later (Rossiter 1958). Such external surges may increase a single high water event by up to 1 m (Bruss et al. 2010). However, the most important factor for the surge generation is related to strong local winds from northwesterly directions blowing over the shallow shelf areas in the German Bight (depths $\leq$40 m; along the coastlines $\leq$10 m) occurring when the low pressure systems travel farther eastward into the Scandinavian/Baltic area. These winds cause an effect of water pile up with surges of up to more than four meters. In January 2007, the meteorological situation was characterized by a strong pressure gradient with pressure anomalies 16 hPa below the long-term mean over Scandinavia and exceeding it by 9 hPa west of the Iberian Peninsula (Fink et al. 2009). These conditions lead to the generation of a series of particular large surge events in the German Bight (Fig. 1).

To study the long-term behavior of such storm events—our first aim in the present study—short-term periodic and long-term MSL changes have to be eliminated from the observational data first (Pugh 2004). A common way of separating tides from surges is to apply a harmonic analysis to the raw data. However, this method has two general restrictions:

1) First, for a harmonic analysis at least hourly observations are needed. This restriction often hampers the evaluation of long tide gauge measurements back into the eighteenth and nineteenth centuries, since before 1900 most observations are limited to readings of tidal high and low water levels (and times).
2) Second, for tide gauges located in shallow continental shelf seas (e.g., Cuxhaven), nonlinear shallow water effects often bias the harmonic representation of tides, leaving unwanted periodic constituents in values of nontidal residuals (surges) (Pawlowicz et al. 2002).

To overcome these restrictions, Horn (1948, 1960) developed a more sophisticated technique that allows accurate predictions of the astronomical tides just on the basis of tidal peaks. The approach is based on a harmonic representation of inequalities (for detailed information of different computation steps, see Müller-Navarra 2013), usually resolving the decomposition of tides and surges more accurately in very shallow seas than the "traditional" harmonic analysis (Pansch 1989). The German Bight has distinct semidiurnal tides and under this assumption it is favorable to represent tidal high and low water heights and times as harmonic deviations of mean heights and intervals. This has also been best practice in the operational service of the German Maritime and Hydrographic Agency for decades (Müller-Navarra 2013). Additionally, as also mentioned above for the Cuxhaven record, digitized hourly observations are limited to the period from 1918 onward. Horn's method, however, allows us to extend the tidal analysis back to 1843 without being dependent on hourly data only available after 1918.

After applying the tidal analysis to the raw data of observational peaks, we generate a surge record by subtracting the astronomical tides from the original signal. One should note that there is a difference between surges generated with hourly observations and tidal peaks.

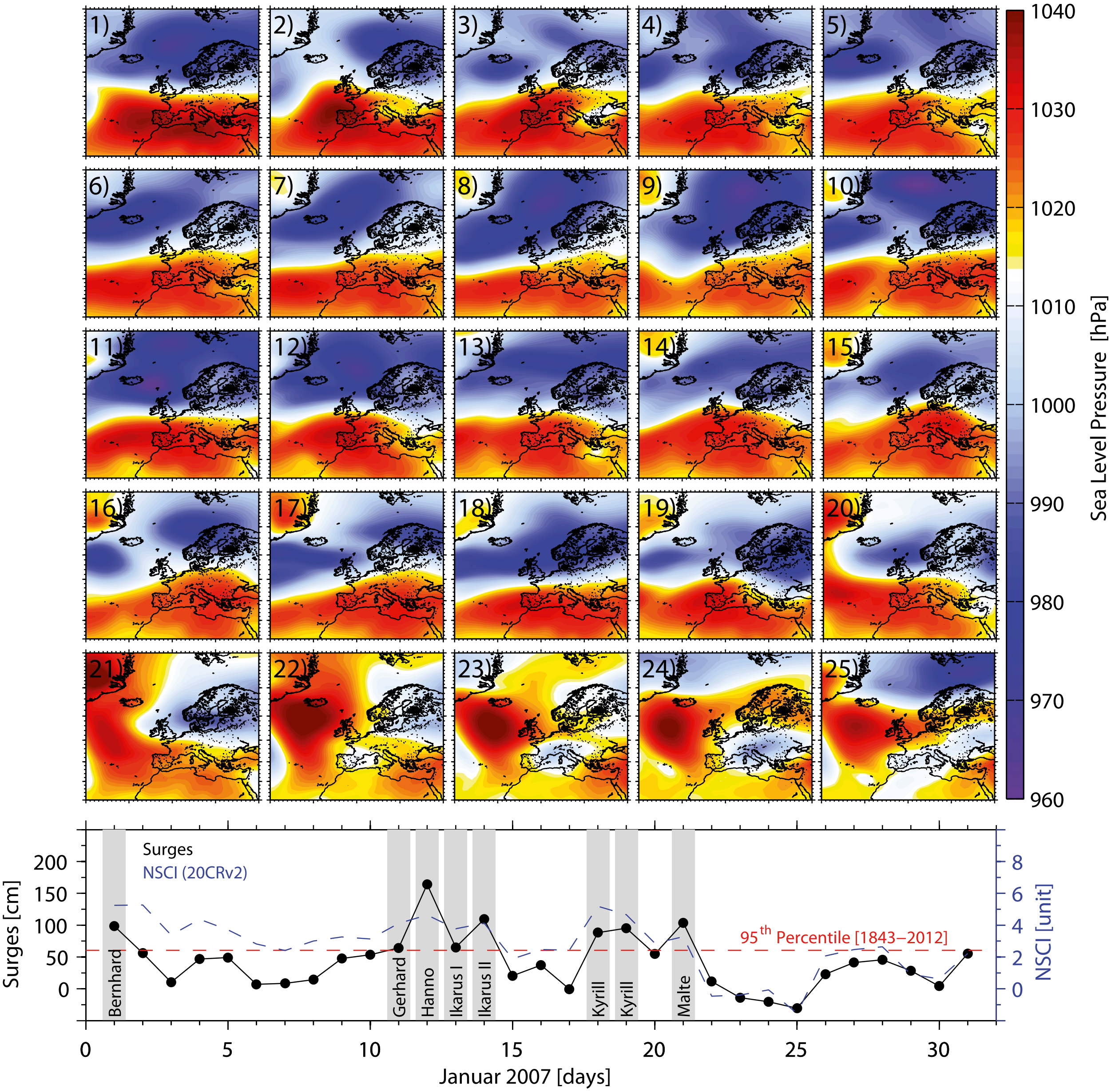


FIG. 1. Visualization of (top) the daily meteorological situation during January 2007 and (bottom) the related daily storm surges measured at the Cuxhaven tide gauge. In the vector plots (top), the colored contours represent areas of similar pressure, while the vectors show the wind speed and direction. For presentation purposes, only the first 25 days are shown. In the graph (bottom), daily surges (black) and daily NSCI index (blue; see also section 3; based on 20CRv2 data) are shown. The red line marks the long-term (1843–2012) 95th percentile of daily (skew) surges, while the gray areas represent single events exceeding this threshold.

When using hourly values, astronomical tidal water levels are subtracted from contemporaneous observed water levels. When using only two tidal high and low water levels per day, astronomical peaks are subtracted from the nearest observed peaks. The two may be up to a few hours apart from each other due to the wind influence, and therefore the resulting surges are known as "skew surges." Hence, we first evaluated whether the skew surges (averaged to daily means) differ in the long-term from original surges (or nontidal residuals) derived from hourly values for the overlapping period from 1918 to 2008 (Fig. 2). For that purpose, we have computed different annual percentiles (0.1%, 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98%, 99%, and 99.9%; Fig. 2a)—the percentiles are computed for each year separately—and the corresponding

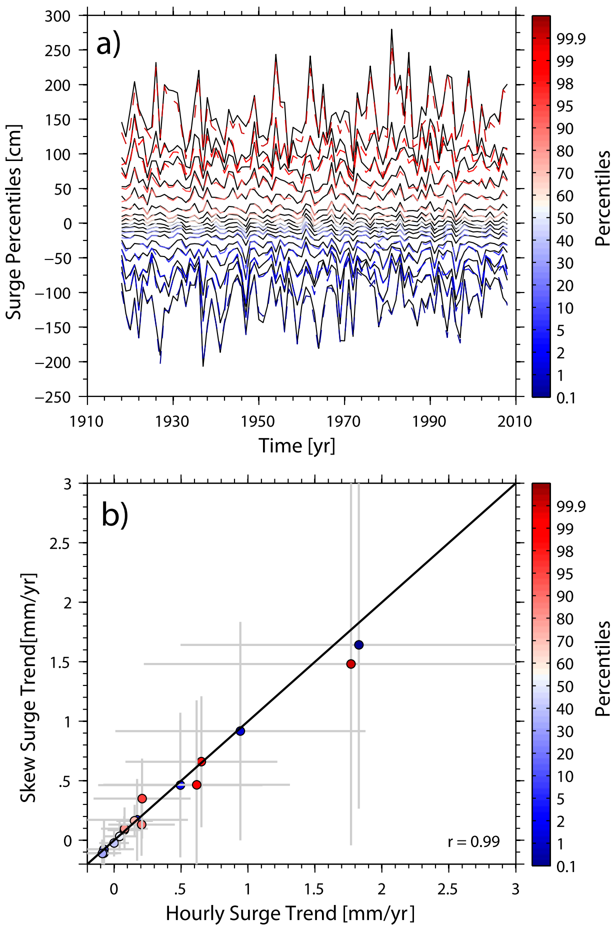


FIG. 2. Comparison of the statistics of daily surges based on hourly observations (black) and the skew surge record (colored and dotted) over the common period from 1918 to 2008 at the tide gauge of Cuxhaven. (a) Annual percentiles and (b) linear trends of annual percentiles as a correlation plot. The gray lines mark the SEs of each trend.

linear trends (Fig. 2b). The figure clearly demonstrates that the percentiles derived from both data sources show virtually the same characteristics in terms of both variability (Fig. 2a) and linear trends (Fig. 2b). The highs and lows in the resulting time series are of similar characteristic: that is, they show the same temporal development and also match in magnitude. This is further confirmed by the correlations in Table 1, which are all larger than 0.94 for the four upper percentiles (which are hereafter investigated in detail).

We investigate storminess by computing annual and seasonal [October–March for the cold season (winter); April–September for the warm season (summer)] 95th, 98th, 99th, and 99.9th percentiles of daily surges. Since no gaps are present in the record, there are no restrictions for the analysis of linear trends. We quantify long-term changes by applying the ordinary least squares regression (OLS). The significance of linear trends is assessed using standard errors (SEs) considering serial correlation of the time series by reducing the number of degrees of freedom as suggested by Santer et al. (2000). It may happen in time series of extreme events that the trends are largely biased by outliers. In such cases, robust regression methods such as the Theil–Sens slope (Gilbert 1987) are more appropriate. We have compared the results from a range of methods and could not find any differences in the trend estimates. This is mainly attributed to the fact that the time series considered here are long and that there are no obvious outliers in the record. Hence, we decided to proceed in the analysis with the common OLS method.

TABLE 1. Pearson correlation coefficients between daily skew surges and daily nontidal residuals (i.e., surges based on hourly measurements) over the period 1918–2008. Significant correlations (*t* test) are marked in boldface.

| | Daily skew surges | | | |
|---|---|---|---|---|
| | 95th | 98th | 99th | 99.9th |
| Daily nontidal residuals | **0.96** | **0.96** | **0.94** | **0.98** |

With respect to our second aim (i.e., comparing storm surges with the variability of large-scale atmospheric circulation patterns), we make use of three additional datasets:

1) The NAO index provided by Jones et al. (1997): The NAO index is a proxy describing large-scale atmospheric circulation over the North Atlantic region. It is calculated by the differences of pressure anomalies taken from stations in southern Iceland and Gibraltar, Spain. The updated index was downloaded from the webpage of the University of East Anglia, United Kingdom (http://www.cru.uea.ac.uk/cru/data/nao/).
2) 20CRv2 wind and pressure fields (Compo et al. 2011): 20CRv2 is the newest generation of global reanalysis products covering a long period from 1871 to 2010. By assimilating daily SLP observations into a state-of-the-art climate model with monthly mean sea surface temperatures and sea ice as boundary conditions, 20CRv2 provides an ensemble of 56 equally likely best estimates of the atmospheric state at a given time step with a temporal resolution of 6 h and on a global grid with a resolution of 2°. For the present investigations, we have downloaded daily data from the webpage of the National Oceanographic and Atmospheric Administration (NOAA), Boulder, Colorado (http://www.esrl.noaa.gov/psd/data/gridded/data.20thC_ReanV2.htm; http://portal.nersc.gov/project/20C_Reanalysis/). Both each individual ensemble member and the ensemble mean are analyzed.

3) Pressure readings from homogenized station records [European and North Atlantic Daily to Multidecadal Climate Variability Project (EMULATE); Ansell et al. 2006]: These data will be used to compare the storm surge record with homogeneous SLP observations covering the past approximately 160 yr on a large scale (see section 3 for more details).

The third objective of our study is to compare the long-term behavior of surges with that of reanalysis wind fields. In a shallow shelf sea such as the North Sea, the variability is clearly dominated by the wind stress: that is, the downward transfer of the momentum from the air into the water. Surges in the southeastern North Sea are caused by atmospheric disturbances over the ocean and can be accurately predicted in the region by the use of simplified statistical–empirical wind surge formulas (Müller-Navarra and Giese 1999). The model used here describes surges $S(t)$ by a number of functions $g_j$ with coefficients $a_j$ and residuals $e(t)$,

$$S(t) = \sum_{j=0}^{n} a_j g_j(t) + e(t), \tag{1}$$

whereas here six functions of $g_j$ based on quadratic and cubic wind stress and SLP fluctuations are linearly fitted with the least squares method to the surges. The functions are given by

$$g_0 = 1, \tag{2}$$

$$g_1 = f^2 \cos(\beta), \tag{3}$$

$$g_2 = f^2 \sin(\beta), \tag{4}$$

$$g_3 = f^3 \cos(\beta), \tag{5}$$

$$g_4 = f^3 \sin(\beta), \quad \text{and} \tag{6}$$

$$g_5 = p - 1015\,\text{hPa}, \tag{7}$$

where $g_0$ is a constant term, $g_1$ and $g_2$ are the quadratic wind stress, $g_3$ and $g_4$ are the cubic wind stress, and $g_5$ is the static response of the water column to SLP changes. The variables $f$ and $\beta$ represent the wind speed and direction, respectively.

We use the empirical relationship to analyze (i) whether the increasing trends in the 20CRv2 data, detected by Donat et al. (2011b), are reflected in the statistical connection between winds, SLP, and surges and (ii) whether the predicted surges differ (on decadal and longer time scales) from the observations. To do so, we apply the wind surge formulas to daily wind and SLP data from the 20CRv2 $u$ and $v$ winds and mean SLP (MSLP) from the nearest grid point at 54°N, 8°E. Since surges measured in Cuxhaven are the cumulative response to changes in the wind field over the North Sea, we have also tested whether using additional grid point time series in the regression model (by using a stepwise regression) may improve the results. No grid point time series was able to increase the model performance significantly. Hence, we decided to use only one grid point for the analysis. Regression coefficients are estimated for the period from 1950 to 2010, a period for which the 20CRv2 was proven to be of good quality (Compo et al. 2011; Krueger et al. 2013b).

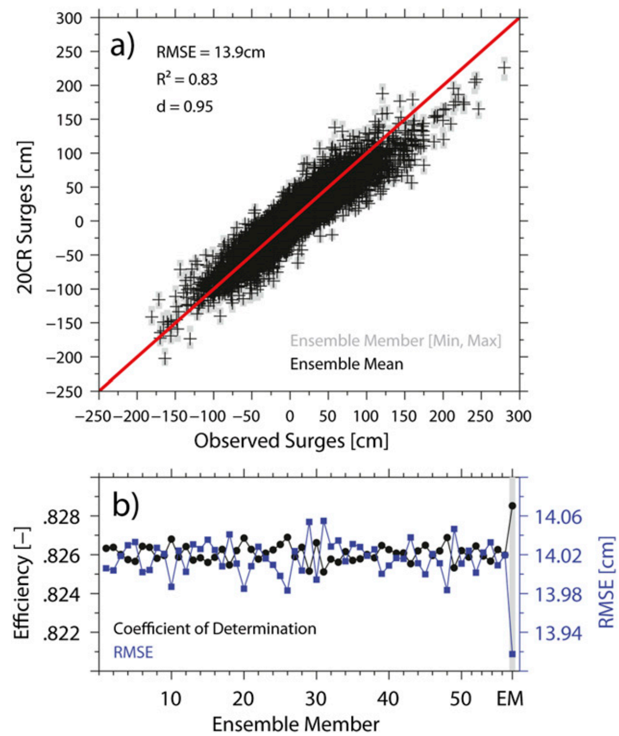


Fig. 3. (a) Correlation plot for observed and modeled surges at the tide gauge of Cuxhaven over the period from 1950 to 2010. The black crosses represent the result by using the ensemble mean as input data, while the gray dots give the minimum and maximum range as a result of evaluating each ensemble member itself. (b) Coefficient of determination (i.e., squared correlation coefficient) and RMSE for each ensemble member and the ensemble mean (gray shaded).

Figure 2 shows the results of the cross validation between observed and predicted surges. The model is able to reproduce the observations during the validation period from 1950 to 2010, as demonstrated by a high correlation of 0.91 and a small root-mean-square error (RMSE) of 13.9 cm for the ensemble mean (Fig. 3a). The RMSE is close to those of hydrodynamic models applied in the region (e.g., Weisse and Plüß 2006), which also confirms the skill of the model. While the performance is particularly high for moderate values, Fig. 3a also suggests an underestimation of the most extreme events. This underestimation may be caused by different factors. First, the ensemble mean may potentially average a few single extreme events out. Second, because of the low temporal and spatial resolution of the model forcing (the reanalysis has a temporal and spatial resolution of 6 h and 2°, respectively) some extreme events which are present in the observations may be lost. Finally, model inaccuracies may be responsible for the underestimation of the most extreme events. To test whether the first reason mentioned above is responsible for the differences, we also tested the predictive skill for each ensemble member separately. While for some individual extreme events the model performance can be slightly improved (Fig. 3a), the general performance cannot be increased significantly. Overall, the ensemble mean shows the highest skill (Fig. 3b). It should be noted that the underestimation of some extreme events is not a unique shortcoming of the statistical model but also present in numerical models (Weisse and Plüß 2006; Arns and Jensen 2010). This is why we suggest that the second reason discussed above (i.e., the temporal and spatial resolution of the model forcing) is most likely responsible for the deviations in the highest percentiles. Because of their stochastic occurrence, it is unlikely that these differences affect the long-term behavior, which is further analyzed in the next sections.

As introduced our main aim is to determine whether the statistically significant relationship between surges and 20CRv2 wind and pressure fields remains stationary in time. Hence, we apply the regression coefficients from Eq. (1) to the entire reanalysis period from 1871 to 2010. To test the stationarity of the relationship, we compute different efficiency criteria—namely, the coefficient of determination ($R^2$) and the RMSE—between observed and predicted surges for each year back to 1871.

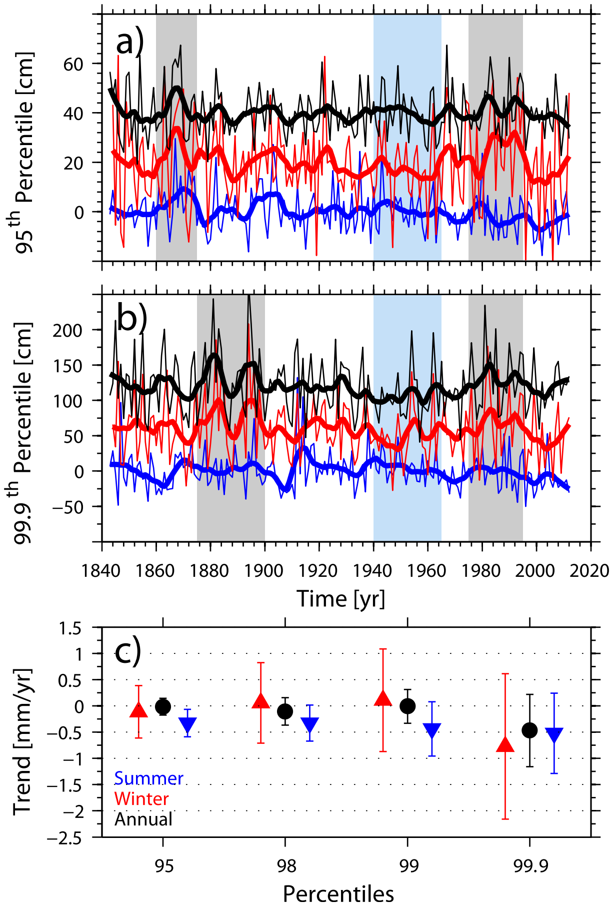


FIG. 4. Normalized (i.e., the long-term average has been removed) annual (black) and seasonal [red is October–March (ONDJFM); blue is April–September (AMJJAS)] time series of the (a) 95th and (b) 99.9th storm surge percentiles. Low-pass-filtered time series [10-yr locally weighted scatterplot smoothing (LOWESS) filter] are shown as thick lines. The gray and blue shaded areas represent periods of increased and decreased storminess, respectively. For presentation purposes, the time series are shown with an arbitrary offset [0, 20, and 40 cm in (a) and 0, 60, and 120 cm in (b)]. (c) Linear trends $\pm 2\sigma$ SE of four upper percentile time series for the period 1843–2012.

## 3. Results

### a. Storm surge trends and variability

We present the storm surge record as a time series of normalized seasonal and annual 95th and 99.9th percentiles together with low-pass-filtered versions of the time series in Fig. 4a,b, respectively. Figure 4c shows linear trends for four upper seasonal and annual percentile (95th, 98th, 99th, and 99.9th) time series. The time series are characterized by a considerable interannual-to-multidecadal variability, while there is no evidence for any significant long-term trend: neither for the seasonal nor for the annual percentile time series. Generally, surge levels are considerably higher and more variable during winter (both on intra-annual and interannual time scales) compared to the summer season. Periods of particular high surges are found at the end of the nineteenth and twentieth centuries (note that the periods at the end of the nineteenth century are slightly different in timing between the 95th and 99th percentiles but at the moment we are not able to explain these differences). In both cases the high rates are dominated by the winter season. Between both peaks, the highest percentiles of surges especially are marked by a gradual decline until the mid-1960s, as noted earlier for central European storminess (Matulla et al. 2008). After the mid-1990s maximum, surges returned to more moderate values.

TABLE 2. Pearson correlation coefficients between winter (ONDJFM) surge percentiles and winter (ONDJFM) SLP indices. Significant correlations (*t* test) are marked in boldface.

| | Surge percentiles | | | |
|---|---|---|---|---|
| Indices | 95th | 98th | 99th | 99.9th |
| NAO | **0.45** | **0.41** | **0.40** | **0.37** |
| NSCI | **0.66** | **0.57** | **0.51** | **0.44** |

The largest contribution to the observed variability in the storm surge record can be found on time scales up to a few decades. From a variety of studies, it is well known that especially during the winter season a considerable fraction of sea level variability can be explained by the NAO (e.g., Yan et al. 2004; Dangendorf et al. 2012). It is also obvious that this relationship does not only exist for mean but also for extreme sea levels (Woodworth et al. 2007; Dangendorf et al. 2013a). We therefore examined the relationship for the winter season by comparing the updated station-based NAO index from Jones et al. (1997) to the four upper percentiles of storm surges (Table 2). In all cases, the comparison exhibits a weak but significant correlation ($r$ = 0.37–0.45) between the time series, being slightly lower for the highest percentiles. This relationship is not stationary over time; it shows considerable fluctuations over the entire period (Figs. 5b,d). In agreement to earlier studies between the NAO and MSL over the Northern European shelf (Jevrejeva et al. 2005) the correlations are high during the mid-nineteenth century, decreasing to insignificant values until the 1960s and then returning back to particular high values at the end of the twentieth century up to the present. This suggests that (i) other factors besides the NAO play an important role for the variability of surges as found earlier also for storminess (Matulla et al. 2008), (ii) the statistical relationship stagnates in times of low large-scale atmospheric variability (i.e., bathymetric effects on the surge generation become more influential), and/or (iii) the influence of the NAO on surges depends on the position of the NAO centers of action (Kolker and Hameed 2007).

To further examine the mechanisms behind this variability we computed the cross correlations between daily surges in Cuxhaven and daily pressure fields from the 20CRv2 (Compo et al. 2011) over the larger geographic area from 60°W to 40°E and from 20° to 80°N. To keep the results unbiased by the increasing uncertainties of reanalysis data in the early decades (Krueger et al. 2013a,b), we evaluated the data over the period from 1950 to 2010. The correlation analysis suggests a dipole-like pattern between surges and SLP with significant negative correlations over Scandinavia and positive correlations over Iberian Peninsula (Fig. 5a). This pattern is also known from MSL time series (Dangendorf et al. 2013b; Dangendorf et al. 2014) in that region and represents the mean weather situation triggering strong storm surges (Heyen et al. 1996). Composite plots (not shown) suggest an increased westerly flow if surges deviate positively from the mean, while the opposite is true for particular negative surges. The dipole-like pattern generally shows similarities to the NAO, but it has a more regional character with a more robust link to the local climate of the German Bight that is also able to reproduce surges in response to serial clustering of extratropical cyclones, such as in January 2007 (Fig. 1; Pinto et al. 2013).

For taking this regionally more relevant large-scale feature of atmospheric variability into account, we define an additional index that is referred to as northern Scandinavia–central Iberia index (NSCI). The index is computed in a similar manner as the station-based NAO index (Jones et al. 1997) by using homogenized daily SLP records of Stockholm and Madrid since 1850 (EMULATE; Ansell et al. 2006). Both stations are located in the closest vicinity to the centers of the correlation pattern of surges at Cuxhaven with the pressure fields (Fig. 5a).

As shown in Table 2, the correlations between the winter half-year NSCI and high storm surge percentiles exceed those of the NAO. More importantly, the link of surges to the NSCI is temporally more stationary than to the NAO. This is indicated by the fact that the running 30-yr correlations with the NSCI remain significant and relatively stable over the entire investigation period (Fig. 5d). As a locally important circulation index like the NSCI can be defined for any location, the main advantage of such an index relates to the temporal long-term robustness of the link between a local variable and the dominating large-scale atmospheric variability. The robust link (in terms stationary running correlations) of high surges at Cuxhaven to the NSCI over time back to 1850 therefore suggests the homogeneity of the surge record, since both parameters (surge levels and SLP) are measured completely independently. Note that the high correlation between both can be taken as an independent measure of homogeneity in terms of low-frequency variability and long-term trends (e.g., Arns and Jensen 2010; Gouriou et al. 2013), while a partial disagreement could be explained either by inhomogeneities or changes in wind circulation (e.g., direction). Such periods of disagreement are, however, not detectable in the presented time series.

### *b. Differences between observations and reanalysis data*

Since we have shown that the storm surge record has a stationary link to the NSCI back to 1850 and is also

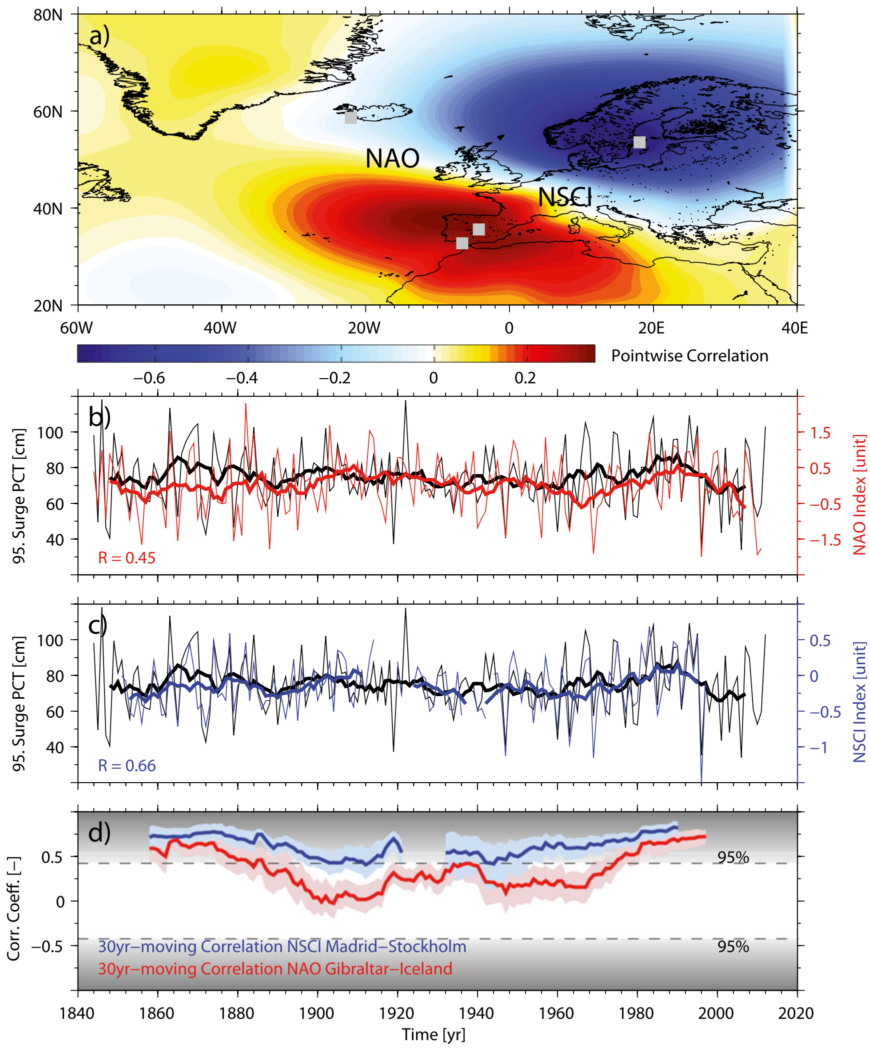


FIG. 5. (a) Pointwise correlations between daily surges in Cuxhaven and gridded daily SLP fields from the 20CRv2 (1950–2010). (b) 95th percentile time series of daily surges in Cuxhaven from 1843 to 2012 in comparison to the seasonal NAO index. A 10-yr moving average is also shown by the thick line. (c) As in (b), but in comparison to the seasonal NSCI index [SLP anomalies Madrid minus SLP anomalies Stockholm, shown as gray squares in (a) together with the stations used for the NAO: i.e., Reykjavik and Gibraltar]. (d) 30-yr moving correlations between 95th percentile time series of daily surges in Cuxhaven and the seasonal NAO/NSCI index. The SEs of the correlation coefficients (blue and red shaded areas) have been computed using a bootstrap method with 500 simulations (Efron and Tibshirani 1993). All time series from (b) to (d) are computed for the storm surge season from October to March.

representative for North Sea storminess, we now turn our attention to the long-term comparison with reanalysis wind fields. For that purpose, we use the statistical relationship described in section 2 to test whether the significant trends in reanalysis wind fields, which were recently detected by Donat et al. (2011b), are consistent with the storm surge record. The evaluation of the reconstructed surge record (i.e., based on 20CRv2 wind and pressure time series from the ensemble mean) confirms a stationary and high predictive skill back until approximately 1910, which is indicated by the efficiency criteria shown in Fig. 6. Before that time, the efficiency criteria point to a decreasing predictive skill: that is, the $R^2$ decreases and the RMSE increases significantly. While during the calibration period from 1951 to 2010 over 83% of the observed variability can be explained by wind and pressure effects, before 1910 the performance gradually decreases to values between 70% and 80% with an absolute minimum in 1871 of 60%. Also for the RMSE the absolute minimum occurs in the early decades (RMSE = 20.64 cm

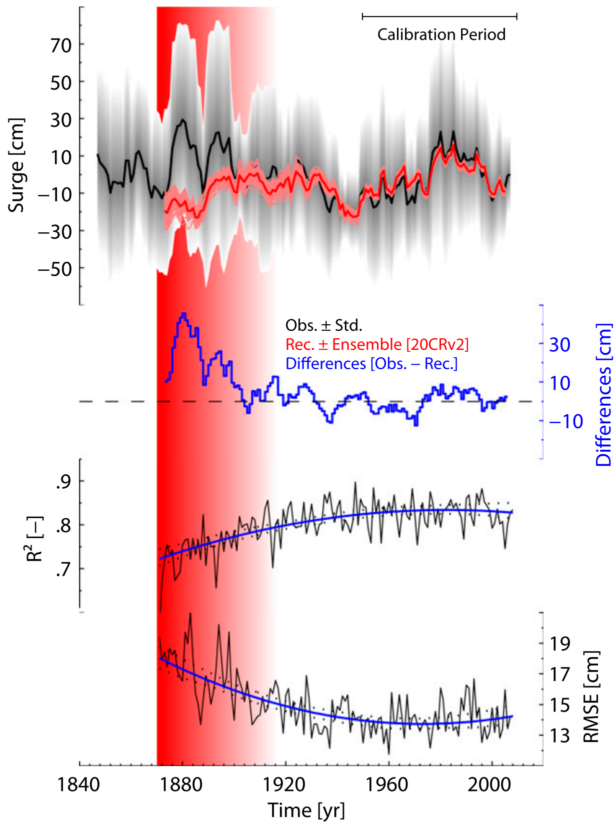


FIG. 6. Reconstruction of surges based on 20CRv2 winds and SLP. The 10-yr moving averages of the 99.9th percentiles of observed surges ± standard deviations (black line with gray shaded area; the standard deviation has been computed as a measure of variability over each 10-yr window) and their reconstructions based on 20CRv2 (light red: individual ensemble members; dark red: ensemble mean; both normalized to a common period from 1950 to 2011: i.e., the mean has been removed) are presented. Differences between both are shown in blue. Annual efficiency criteria between observed and reconstructed daily surges are presented in black. The red shaded area marks the period for which significant differences between observations and 20CRv2 are detected. The different shades demonstrate the gradual increase of inconsistencies before the 1910s.

for the year 1883), exceeding the mean of the calibration period by approximately 150%. As shown before with the stationary correlation of observed surge levels with the NSCI since 1850 (Fig. 5b), the deviations of observed surge levels with those predicted through 20CRv2 are unlikely to be caused by the observational record (the NSCI and the storm surge record are measures independently).

A similar picture is retrieved by comparing the 10-yr moving averages of the 99.9th percentile time series of observed and statistically reconstructed surges (Fig. 6). Over the past 100 yr, the reconstruction fits well to the observations. The model predicts the known decline in storminess in the mid-twentieth century, the rapid increase until the mid-1990s, and the downturn afterward. Nevertheless, in the early 1910s, the prediction starts to decrease in a manner not visible in the observations (for both the ensemble mean as well as the ensemble spread). This decrease finally results in significant positive long-term trends over the entire reanalysis period from 1871 to 2010 if 20CRv2 is taken as predictor (note that a similar behavior was also observed for the 95th percentile; Fig. S1 of the supplementary material).

Related to this, Brönnimann et al. (2012) demonstrated that the ensemble mean appears to be biased toward lower wind speeds during earlier decades. They recommended the use of single ensemble members rather than the ensemble mean when investigating long-term changes. To examine whether the results from Fig. 6 are influenced by such biases, we additionally evaluated long-term changes in each ensemble member separately. First, we calculated the differences between the percentile time series from each 20CRv2 ensemble member prediction and the observed time series. Then, in a second step, we computed linear trends for each of the residual time series. The results show that for each percentile all ensemble members point to significant positive long-term changes, which are further significantly different from the observations (Fig. 7a). Additionally, we found that the residual trends are generally increasing with the order of the percentiles: that is, highest deviations are found within the highest percentiles.

To determine the exact timing from which the 20CRv2 generated surges start to deviate significantly from the observations, we further computed linear trends for the residual time series over 30-yr moving windows. The results are shown in Figs. 7b,c for the 99th and 99.9th percentiles, respectively. While the trend estimates scatter around zero back to approximately 1910, before that time statistically significant differences are found for the ensemble mean as well as each individual ensemble member. While we can confirm the bias of the ensemble mean reported by Brönnimann et al. (2012) (Fig. 7b), our results also illustrate that using individual members cannot improve the results significantly (when assessing the long-term behavior of storm surges in the German Bight). The reanalysis is significantly biased toward a lower occurrence of extreme values in the period prior to 1910 in both the ensemble mean as well as all members (in this region).

The decreasing coherence between reanalysis forcing and observed surges is generally in line with increasing uncertainties in the reanalysis because of fewer assimilated observational data in the earlier periods (Compo et al. 2011; Krueger et al. 2013b). The results therefore partly confirm the inconsistencies between storm

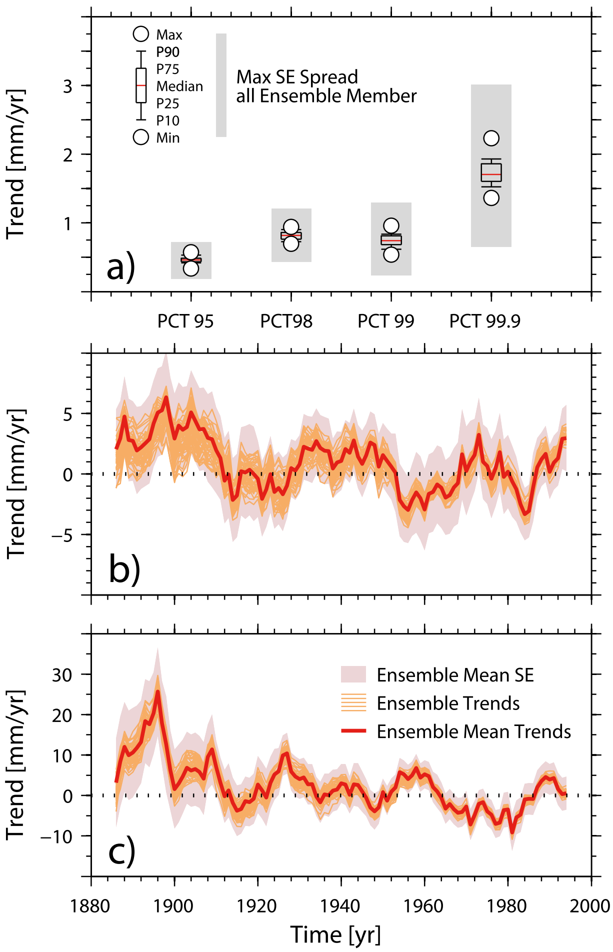


FIG. 7. (a) Box plot of residual percentile trends (20CRv2 generated surges minus observation) from all ensemble members over the period from 1871 to 2011. The gray shaded areas represent the maximal range of SEs (95% confidence level): that is, residual trends are significant for all ensemble members and percentiles. (b) 30-yr moving trends for the 99th percentile residuals (20CRv2 generated surges minus observation). All 56 ensemble members are shown in orange. The ensemble mean together with the related SE's is given by the red line and the red shaded area, respectively. (c) As in (b), but for the 99.9th percentile residuals.

observations and 20CRv2 found by Krueger et al. (2013a,b) over the North Atlantic region prior the 1940s. In contrast to this earlier study, the inconsistencies with surge levels at Cuxhaven are less pronounced and become clearly visible only for the period before the 1910s. It should be noted that the surges itself differ to the conventional storminess proxies insofar that they measure both changes in wind speed and direction. However, there are three reasons why this does not affect our main conclusion that the surge record is representative for storminess in the region:

1) The rank correlation (Spearman) between wind speeds and absolute surges has a value of 0.53, which is significantly different statistically from zero (see Fig. S2 of the supplementary material). Since during periods of low wind speeds the influence of bathymetric effects on the surge generation increases relative to the winds, the correlation between the extreme events, which are analyzed here, increases noticeably. For example, the correlation between the annual 95th percentile time series of both factors increases to a value of 0.64 (not shown).
2) Dangendorf et al. (2013c) investigated changes in the frequency of different wind directions between January and March from 1871 to 2008 and could not find any evidence for significant changes during the period of interest.
3) The fact that Krueger et al. (2013b) previously detected similar differences in conventional proxies (although for a slightly different region and with larger differences than detected here) and Donat et al. (2011b) pointed to significant trends in the wind speed further supports the reliability of the surge record as a measure of storminess.

Additionally, one could argue that the use of a single grid point time series rather than the entire wind field over the North Sea as input for the empirical wind surge model could bias the results. When looking at Fig. 3 in Donat et al. (2011b), it can clearly be seen that the linear trends are evident over the entire North Sea area. In fact, strongest trends in wind speeds have been found for the grid points not used as input data in the present study. Therefore, using other additional grid points as input parameters for the wind surge formulas would lead to even larger inconsistencies between 20CRv2 and surge observations.

On this basis, we conclude that the significant trend detected by Donat et al. (2011b) was less a result of the large decadal trends in storminess in the last decades but rather reflects lower occurrence of extreme values in the early decades of the reanalysis. The latter are supported by neither pressure-based storm indices (Krueger et al. 2013a,b) nor the surge observations at Cuxhaven since 1843.

## 4. Discussion and conclusions

We have established a new storm surge record for the tide gauge of Cuxhaven, extending conventional records back to 1843. This could be achieved by using a generally known but in the last decades less considered method (Horn 1948, 1960) that enables us to decompose measurements of tidal high and low water levels into tides, MSL, and (skew) surges. The method can open a worldwide available but in terms of surges unappreciated data source for investigating surges and hence storminess. For Europe, tide gauges located in Liverpool, Amsterdam, and Brest provide readings of tidal high and/or low water levels back into the seventeenth century. Additionally, Talke and Jay (2013) recently pointed to the availability of records in the Pacific and western Atlantic region going back to the mid-nineteenth century with more than 50 records available from 1900 onward. The analysis of these valuable and more or less robust oceanographic measurements could be of high importance for the analysis of storms going into times for which only a few conventional proxies for storminess exist.

Despite the fact that surges measure changes in wind speed and direction, the analysis of high annual percentiles of surges at Cuxhaven confirms earlier observational studies on storminess over the European–Atlantic region with conventional proxies (Schmidt and von Storch 1993; Alexandersson et al. 1998, 2000; Bärring and von Storch 2004; Matulla et al. 2008; Hanna et al. 2008; Wang et al. 2009) in terms of both variability and trends. Consistent with the different pressure-based storm indices of the last up to 150 yr, periods of increased storminess with higher occurrence of extreme storm surges prevailed at the end of the nineteenth and twentieth century, with very low levels in the 1970s. While we could identify a considerable interannual-to-multidecadal variability, which is significantly correlated to large-scale atmospheric variability over the North Atlantic and European region, no robust long-term trend could be detected in surges since 1843 at the tide gauge of Cuxhaven. The absence of any robust long-term trends in annual storminess in observations of the last up to 170 yr over the Euro–Atlantic region seems to support global modeling results of externally forced coupled atmosphere–ocean general circulation models (AOGCM). They indicate no long-term changes in storminess (Fischer-Bruns et al. 2005) or cyclone characteristics (Xia et al. 2012) for the Northern Hemisphere through the last millennium, although most AOGCM's point to an increase in storminess under enhanced greenhouse gas emission in a future climate (e.g., Gastineau and Soden 2009; Pinto et al. 2007; Donat et al. 2011a).

By applying a simple statistical model to reanalysis (20CRv2) winds and SLP over the North Sea, we have further reconstructed storm surges in the German Bight over the entire reanalysis period since 1871. Based on the ensemble mean as well as the ensemble spread, we demonstrated that the reanalysis data have a high predictive skill back to the 1910s, while previously the model skill decreases considerably leading to lower occurrence of extreme storm surges in the first four to five decades of the reanalysis. This decrease in storminess is visible in neither surge observations at Cuxhaven nor different pressure-based storm indices (Krueger et al. 2013a,b) over the European–Atlantic region. Hence, the significant positive trends detected in 20CRv2 storminess by Donat et al. (2011b) appears to be less a result of the large decadal trends in storminess in the last decades but rather reflects the lower occurrences of extreme winds in the early decades of the reanalysis. In contrast to the results from Krueger et al. (2013b), which are partly from a different region, our study points to increasing inconsistencies between reanalysis and observation data before the 1910s. The inconsistencies and their dating are supported by each ensemble member as well as the ensemble mean. As the link of surge levels at Cuxhaven with the fully independent NSCI remains stationary back to 1850, the discrepancies with 20CRv2 are unlikely to be explained with inconsistencies in the surge record. Thus, we conclude that the 20CRv2 represents a useful database for the North Sea region from the beginning of the twentieth century, but one has to be careful by computing linear trends, particularly when periods before 1910 are included in the analysis.

We further recommend the presented methods as an independent quality check of reanalysis and tide gauge data in other regions of the world. Especially in regions where meteorological observations are sparse, the cross validation with homogeneous tide gauge data might provide information on the consistency of reanalysis data on longer time scales. In turn, in regions where the tide gauge network is sparse, homogenous reanalysis data may provide information on the homogeneity of sea level measurements. In case of inconsistencies in the long-term variations between reanalysis and sea level observations, an independent pressure index like the NSCI for the German Bight can be established from observations to study the robustness of the local observations or reanalysis data relative to this index. Whether a similar link between large-scale forcing (pressure and wind) and local surges can be also established (e.g., in tropical regions) needs to be evaluated. The homogeneity of observations and reanalysis data are indispensable for oceanographers and meteorologists to study multidecadal variations or trends in storminess or exchange processes between the atmosphere and the ocean, which is in turn an important step in understanding the predictability of the system.

*Acknowledgments.* We highly acknowledge four anonymous reviewers for their valuable comments. We further thank the Twentieth Century Reanalysis team for providing the dataset without any charge. Support for the Twentieth Century Reanalysis Project dataset is provided by the U.S. Department of Energy, Office of Science Innovative and Novel Computational Impact on Theory and Experiment program, and Office of Biological and Environmental Research, and by the National Oceanic and Atmospheric Administration.

REFERENCES

Alexandersson, H., T. Schmith, K. Iden, and H. Tuomenvirta, 1998: Long-term variations of the storm climate over NW Europe. *Global Atmos. Ocean Syst.,* **6,** 97–120.

——, H. Tuomenvirta, T. Schmith, and K. Iden, 2000: Trends of storms in NW Europe derived from an updated pressure data set. *Climate Res.,* **14,** 71–73, doi:10.3354/cr014071.

Ansell, T. J., and Coauthors, 2006: Daily mean sea level pressure reconstructions for the European–North Atlantic region for the period 1850–2003. *J. Climate,* **19,** 2717–2742, doi:10.1175/JCLI3775.1.

Arns, A., and J. Jensen, 2010: Developing sustainable coastal protection- and mangement strategies for Schleswig-Holstein's Halligen considering climate change (ZukunftHallig). CoastDoc 2010, Mitteilungen des Forschungsinstituts Wasser und Umwelt, Universität Siegen, Heft 2, 24–31.

Bärring, L., and H. von Storch, 2004: Scandinavian storminess since about 1800. *Geophys. Res. Lett.,* **31,** L20202, doi:10.1029/2004GL020441.

Brönnimann, S., O. Martius, H. Von Waldow, C. Welker, J. Luterbacher, G. P. Compo, P. D. Sardeshmukh, and T. Usbeck, 2012: Extreme winds at northern mid-latitudes since 1871. *Meteor. Z.,* **21,** 13–27, doi:10.1127/0941-2948/2012/0337.

Bruss, G., G. Gönnert, and R. Mayerle, 2010: Extreme scenarios at the German North Sea coast: A numerical modeling study. *Proc. 32nd Conf. on Coastal Engineering,* Shanghai, China, American Society of Civil Engineers, 26. [Available online at 10.9753/icce.v32.currents.26.]

Church, J. A., and N. J. White, 2011: Sea-level rise from the late 19th to the early 21st century. *Surv. Geophys.,* **32,** 585–602, doi:10.1007/s10712-011-9119-1.

Compo, G. B., and Coauthors, 2011: The Twentieth Century Reanalysis project. *Quart. J. Roy. Meteor. Soc.,* **137,** 1–28, doi:10.1002/qj.776.

Dangendorf, S., T. Wahl, H. Hein, J. Jensen, S. Mai, and C. Mudersbach, 2012: Mean sea level variability and influence of the North Atlantic Oscillation on long-term trends in the German Bight. *Water,* **4,** 170–195, doi:10.3390/w4010170.

——, C. Mudersbach, J. Jensen, G. Anette, and H. Heinrich, 2013a: Seasonal to decadal forcing of high water level percentiles in the German Bight throughout the past century. *Ocean Dyn.,* **63,** 533–548, doi:10.1007/s10236-013-0614-4.

——, ——, T. Wahl, and J. Jensen, 2013b: Characteristics of intra-, inter-annual and decadal variability and the role of meteorological forcing: The long record of Cuxhaven. *Ocean Dyn.,* **63** (2–3), 209–224, doi:10.1007/s10236-013-0598-0.

——, T. Wahl, C. Mudersbach, and J. Jensen, 2013c: The seasonal mean sea level cycle in the southeastern North Sea. *J. Coastal Res.,* **65,** 1915–1920.

——, ——, E. Nilson, B. Klein, and J. Jensen, 2014: A new atmospheric proxy for sea level variability in the southeastern North Sea: Observations and future ensemble predictions. *Climate Dyn.,* doi:10.1007/s00382-013-1932-4, in press.

Donat, M. G., G. C. Leckebusch, S. Wild, and U. Ulbrich, 2011a: Future changes in European winter storm losses and extreme wind speeds inferred from GCM and RCM multi-model simulations. *Nat. Hazards Earth Syst. Sci.,* **11,** 1351–1370, doi:10.5194/nhess-11-1351-2011.

——, D. Renggli, S. Wild, L. V. Alexander, G. C. Leckebusch, and U. Ulbrich, 2011b: Reanalysis suggests long-term upward trends in European storminess since 1871. *Geophys. Res. Lett.,* **38,** L14703, doi:10.1029/2011GL047995.

Efron, B., and J. R. Tibshirani, 1993: *An Introduction to the Bootstrap. Monogr. on Statistics and Applied Probability,* No. 57, Chapman & Hall, 456 pp.

Fink, A. H., T. Brücher, V. Ermert, A. Krüger, and J. G. Pinto, 2009: The European storm Kyrill in January 2007: Synoptic evolution, meteorological impacts and some considerations with respect to climate change. *Nat. Hazards Earth Syst. Sci.,* **9,** 405–423, doi:10.5194/nhess-9-405-2009.

Fischer-Bruns, I., H. von Storch, J. F. González-Rouco, and E. Zorita, 2005: Modelling the variability of mid-latitude storm activity on decadal to century time scales. *Climate Dyn.,* **25,** 461–476, doi:10.1007/s00382-005-0036-1.

Gastineau, G., and B. J. Soden, 2009: Model projected changes of extreme wind events in response to global warming. *Geophys. Res. Lett.,* **36,** L10810, doi:10.1029/2009GL037500.

Gilbert, R. O., 1987: Sens's nonparametric estimate of slope. *Statistical Methods for Environmental Pollution Monitoring.* John-Wiley and Sons, 217–219.

Gouriou, G., M. B. Miguez, and G. Wöppelmann, 2013: Reconstruction of a two-century long sea level record for the Pertuis d'Antioche (France). *Cont. Shelf Res.,* **61–62,** 31–40, doi:10.1016/j.csr.2013.04.028.

Hanna, E., J. Cappelen, R. Allan, T. Jónsson, F. le Blancq, T. Lillington, and K. Hickey, 2008: New insights into north European and North Atlantic surface pressure variability, storminess, and related climatic change since 1830. *J. Climate,* **21,** 6739–6766, doi:10.1175/2008JCLI2296.1.

Heyen, H., E. Zorita, and H. von Storch, 1996: Statistical downscaling of monthly mean North Atlantic air-pressure to sea level anomalies in the Baltic Sea. *Tellus,* **48A,** 312–323.

Horn, W., 1948: Über die Darstellung der Gezeiten als Funktion der Zeit. *Dtsch. Hydrogr. Z.,* **1,** 124–140, doi:10.1007/BF02226142.

——, 1960: Some recent approaches to tidal problems. *Int. Hydrogr. Rev.,* **37,** 65–84.

Hunter, J., 2010: Estimating sea-level extremes under conditions of uncertain sea-level rise. *Climatic Change,* **99,** 331–350, doi:10.1007/s10584-009-9671-6.

Hurrell, J. W., 1995: Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science,* **269,** 676–679, doi:10.1126/science.269.5224.676.

Jevrejeva, S., J. C. Moore, P. L. Woodworth, and A. Grinsted, 2005: Influence of large scale atmospheric circulation on European sea level results based on the wavelet transform method. *Tellus,* **57A,** 183–193.

Jones, P. D., T. Jonsson, and D. Wheeler, 1997: Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland. *Int. J. Climatol.,* **17,** 1433–1450, doi:10.1002/(SICI)1097-0088(19971115)17:13<1433::AID-JOC203>3.0.CO;2-P.

Kolker, A. S., and S. Hameed, 2007: Meteorologically driven trends in sea level rise. *Geophys. Res. Lett.,* **34,** L23616, doi:10.1029/2007GL031814.

Krueger, O., F. Feser, L. Bärring, E. Kaas, T. Schmith, H. Tuomenvirta, and H. von Storch, 2013a: Comment on "Trends and low frequency variability of extra-tropical cyclone activity in the ensemble of Twentieth Century Reanalysis" by Xiaolan L. Wang, Y. Feng, G. P. Compo, V. R. Swail, F. W. Zwiers, R. J. Allan, and

P. D. Sardeshmukh, Climate Dynamics, 2012. *Climate Dyn.,* **42,** 1127–1128, doi:10.1007/s00382-013-1814-9.

——, F. Schenk, F. Feser, and R. Weisse, 2013b: 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.

Lindenberg, J., H. Mengelkamp, and G. Rosenhagen, 2012: Representativity of near surface wind measurements from coastal stations at the German Bight. *Meteor. Z.,* **21,** 99–106, doi:10.1127/0941-2948/2012/0131.

Matulla, C., W. Schöner, H. Alexandersson, H. von Storch, and X. L. Wang, 2008: European storminess: Late nineteenth century to present. *Climate Dyn.,* **31,** 125–130, doi:10.1007/s00382-007-0333-y.

Mudersbach, C., T. Wahl, I. D. Haigh, and J. Jensen, 2013: Trends in high sea levels along the German North Sea coastline compared to regional mean sea level changes. *Cont. Shelf Res.,* **65,** 111–120, doi:10.1016/j.csr.2013.06.016.

Müller-Navarra, S. H., 2013: On tidal predictions by means of harmonic representation of inequalities. Federal Maritime and Hydrographic Agency Rep. 50, 64 pp. [Available online at http://www.bsh.de/de/Produkte/Buecher/Berichte_/Bericht50/Bericht50.pdf.]

——, and H. Giese, 1999: Improvements of an empirical model to forecast wind surge in the German Bight. *Ocean Dyn.,* **51,** 385–405, doi:10.1007/BF02764162.

Pansch, E., 1988: *Harmonische Analyse von Gezeiten- und Gezeitenstrombeobachtungen im Deutschen Hydrographischen Institut.* Wissenschaftlich-Technische Berichte des DHI, 1988-1, 32 pp.

Pawlowicz, R., B. Beardsley, and S. Lentz, 2002: Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE. *Comput. Geosci.,* **28,** 929–937, doi:10.1016/S0098-3004(02)00013-4.

Pinto, J. G., U. Ulbrich, G. C. Leckebusch, T. Spangehl, M. Reyers, and S. Zacharias, 2007: Changes in storm track and cyclone activity in three SRES ensemble experiments with the ECHAM5/MPIOM1 GCM. *Climate Dyn.,* **29,** 195–210, doi:10.1007/s00382-007-0230-4.

——, N. Bellenbaum, M. K. Karremann, and P. M. Della-Marta, 2013: Serial clustering of extratropical cyclones over the North Atlantic and Europe under recent and future climate conditions. *J. Geophys. Res.*, **118,** 12 476–12 485, doi:10.1002/2013JD020564.

Pugh, D., 2004: *Changing Sea Levels: Effects of Tide, Weather and Climate.* Cambridge University Press, 265 pp.

Rossiter, J. R., 1958: Storm surges in the North Sea, 11 to 30 December 1954. *Philos. Trans. Roy. Soc. London,* **251A,** 139–160, doi:10.1098/rsta.1958.0012.

Santer, B. D., T. M. L. Wigley, J. S. Boyle, D. J. Gaffen, J. J. Hnilo, D. Nychka, D. E. Parker, and K. E. Taylor, 2000: Statistical significance of trends and trend differences in layer-average atmospheric temperature time series. *J. Geophys. Res.,* **105,** 7337–7356, doi:10.1029/1999JD901105.

Schmidt, H., and H. von Storch, 1993: German Bight storm analysed. *Nature,* **365,** 791, doi:10.1038/365791a0.

Slangen, A. B. A., C. A. Katsman, R. S. W. van de Wal, L. L. A. Vermeersen, and R. E. M. Riva, 2012: Towards regional projections of twenty-first century sea-level change based on IPCC SRES scenarios. *Climate Dyn.,* **38,** 1191–1209, doi:10.1007/s00382-011-1057-6.

Talke, S. A., and D. A. Jay, 2013: Nineteenth century North American and Pacific tidal data: Lost or just forgotten? *J. Coastal Res.,* **29,** 118–127, doi:10.2112/JCOASTRES-D-12-00181.1.

Ullmann, A., and J. Monbaliu, 2010: Changes in atmospheric circulation over the North Atlantic and sea surge variations along the Belgian coast during the twentieth century. *Int. J. Climatol.,* **30,** 558–568, doi:10.1002/joc.1904.

von Storch, H., 2014: Storm surges: Phenomena, forecasting and scenarios of change. *Procedia IUTAM,* **10,** 356–362, doi:10.1016/j.piutam.2014.01.030.

——, and H. Reichardt, 1997: A scenario of storm surge statistics for the German Bight at the expected time of doubled atmospheric carbon dioxide concentration. *J. Climate,* **10,** 2653–2662, doi:10.1175/1520-0442(1997)010<2653:ASOSSS>2.0.CO;2.

——, J. Jensen, and T. Frank, 2010: On analysing sea level rise in the German Bight since 1844. *Nat. Hazards Earth Syst. Sci.,* **10,** 171–179, doi:10.5194/nhess-10-171-2010.

——, ——, ——, and I. D. Haigh, 2011: Improved estimates of mean sea level changes in the German Bight over the last 166 years. *Ocean Dyn.,* **61,** 701–715, doi:10.1007/s10236-011-0383-x.

——, and Coauthors, 2013: Observed mean sea level changes around the North Sea coastline from 1800 to present. *Earth Sci. Rev.,* **124,** 51–67, doi:10.1016/j.earscirev.2013.05.003.

Wang, X. L., F. W. Zwiers, V. R. Swail, and Y. Feng, 2009: Trends and variability of storminess in the northeast Atlantic region, 1874–2007. *Climate Dyn.,* **33,** 1179–1195, doi:10.1007/s00382-008-0504-5.

——, Y. Feng, G. P. Compo, F. W. Zwiers, R. J. Allan, V. R. Swail, and P. D. Sardeshmukh, 2014: Is the storminess in the Twentieth Century Reanalysis really inconsistent with observations? A reply to the comment by Krueger et al. (2013b). *Climate Dyn.,* **42,** 1113–1125, doi:10.1007/s00382-013-1828-3.

Weisse, R., and A. Plüβ, 2006: Storm-related sea level variation along the North Sea coast as simulated by a high-resolution model 1958–2002. *Ocean Dyn.,* **56,** 16–25, doi:10.1007/s10236-005-0037-y.

——, and H. von Storch, 2010: *Marine Climate and Climate Change: Storms, Wind Waves and Storm Surges.* Springer Praxis, 200 pp.

Woodworth, P. L., R. A. Flather, J. A. Williams, S. L. Wakelin, and S. Jevrejeva, 2007: The dependence of UK extreme sea levels and storm surges on the North Atlantic Oscillation. *Cont. Shelf Res.,* **27,** 935–946, doi:10.1016/j.csr.2006.12.007.

——, M. Menéndez, and W. R. Gehrels, 2011: Evidence for century-timescale acceleration in mean sea levels and for recent changes in extreme sea levels. *Surv. Geophys.,* **32,** 603–618, doi:10.1007/s10712-011-9112-8.

Woth, K., R. Weisse, and H. von Storch, 2006: Climate change and North Sea storm surge extremes: An ensemble study of storm surge extremes expected in a changed climate projected by four different regional climate models. *Ocean Dyn.,* **56,** 3–15, doi:10.1007/s10236-005-0024-3.

Xia, L., H. von Storch, and F. Feser, 2012: Quasi-stationarity of centennial Northern Hemisphere midlatitude winter storm tracks. *Climate Dyn.,* **41,** 901–916, doi:10.1007/s00382-012-1543-5.

Yan, Z., M. N. Tsimplis, and D. Woolf, 2004: Analysis of the relationship between the North Atlantic Oscillation and sea level changes in northwest Europe. *Int. J. Climatol.,* **24,** 743–758, doi:10.1002/joc.1035.

Zhang, K., B. C. Douglas, and S. P. Leatherman, 2000: Twentieth-century storm activity along the U.S. East Coast. *J. Climate,* **13,** 1748–1761, doi:10.1175/1520-0442(2000)013<1748:TCSAAT>2.0.CO;2.