Interpreting Century-Scale Changes in Southern North Sea Storm Surge Climate Derived from Coupled Model Simulations

TOM HOWARD AND JASON LOWE

Met Office Hadley Centre, Exeter, United Kingdom

KEVIN HORSBURGH

National Oceanography Centre, Liverpool, United Kingdom

ABSTRACT

This paper describes numerical experiments using a climate–storm surge simulation system for the coast of the United Kingdom, with a particular focus on the southern North Sea and the Thames estuary in southeastern England.

Time series of surges simulated in the southern North Sea by a surge model driven by atmospheric data from a regional climate model and surges simulated by the same surge model driven by atmospheric data from a global climate model are compared. A strong correspondence is demonstrated, and a linear scaling factor relating them is derived. This factor varies slowly with location. Around the Thames estuary, extreme surges are compared in the same way, and the linear scaling factor for the extremes is found to be similar to that for the full time series. The authors therefore assert that in seeking significant trends in surge at this location using this model arrangement, the regional model downscaling stage could be avoided, if observations were used to establish a suitable scaling factor for each location.

The influence of the tide–surge phase relationship is investigated, and extreme sea levels at the mouth of the River Thames from regional-model-driven simulations are compared to the extreme event of 1953. Although the simulated levels are slightly lower, they are found to be comparable given the observational uncertainty.

The assumption that time-mean sea level changes can be added linearly to surge changes is investigated at this location for large changes in time-mean sea level. The authors find that the primary effect of such an increase is on the speed of propagation of tide and surge, supporting the case for a simple linear addition of mean and extreme sea level changes.

1. Introduction

The coast plays a major role in the culture and economy of the United Kingdom, which has the longest coastline in the European Union. Concern about climate change effects on this national asset motivated the production of the government-funded United Kingdom Climate Impacts Programme report "UK Climate Projections 2009 marine scenarios" [henceforth UKCP09 (Marine); Lowe et al. 2009).

In a shallow well-mixed sea, the surface wind stress can be considered to act by accelerating the full depth of the water column; thus, to a first approximation, the shallower the water, the greater the effect of wind stress.

Thus, the southeast coast of England is prone to flooding by storm surge owing to its location near the almost-closed southern end of the shallow North Sea. The River Thames is tidal all through central London, and thus the Thames estuary is a particular hot spot of coastal flooding sensitivity. Of an estimated £150 billion (230 billion U.S. dollars) assets at risk from coastal flooding in the United Kingdom, half are located on this estuary. This is the motive that provided the greater part of our funding, through the U.K. Environment Agency's Thames Estuary 2100 project (TE2100 2010), which seeks to provide a flood risk management plan for the Thames out to the end of the twenty-first century. Thus, our particular focus here is on the results of regionalized climate modeling studies of this key location.

Some of the work presented here augments the work reported in Howard et al. (2008). Our objectives are threefold. First, we investigate the relationship between

Corresponding author address: Tom Howard, Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, United Kingdom.
E-mail: tom.howard@metoffice.gov.uk

surges modeled in the southern North Sea and particularly at the Thames estuary with and without the atmospheric downscaling step provided by a regional climate model (RCM). This issue is not discussed in either of the reports mentioned above. Second, we investigate the influence of the tide–surge phase relationship on surface elevation at this location. This augments the results presented in Howard et al. (2008). Last, we consider the assumption that mean sea level change can be added linearly to changes in surge climate for this location. This also augments the results presented in Howard et al. (2008).

2. Climate modeling system and experiments

The third climate configuration of the Met Office Unified Model (HadCM3; Gordon et al. 2000; Pope et al. 2000) is a general circulation model that has previously been shown to have considerable skill at simulating the global climate (e.g., Stott et al. 2000). The atmosphere model has 19 levels in the vertical and a horizontal resolution of $2.5° \times 3.75°$. The ocean model has 20 vertical levels and a horizontal resolution of $1.25° \times 1.25°$. Processes at scales smaller than the grid size are usually parameterized, that is, represented by simple relationships between the large scales and these smaller scales. The parameters in these relationships are often not precisely known, and the model can be run with a range of parameter values and still credibly reproduce observed climate.

To estimate uncertainty in projections of the future, we can run these plausible model versions beyond present day and examine the spread of the results (Murphy et al. 2007; Collins et al. 2006). This approach of taking a single model structure and varying the model parameters within that structure we refer to as the perturbed physics ensemble (PPE).

The resolution of HadCM3 is not sufficient to provide the regional impact studies of climate change that are reported in UKCP09 (Murphy et al. 2009), so the global climate model PPE is used to provide atmospheric boundary conditions for a one-way nested regional atmospheric climate model, the Hadley Centre Regional Model, version 3 (HadRM3; see Murphy et al. 2007), which is set up to simulate climate over Europe in more detail.

The global coupled climate model ensemble (PPE) of 17 members was set up to simulate the changing climate from 1860 to 2100. First, the global model ensemble members were spun up to a perpetual 1860-like climate. During the long spin-up stage (hundreds of years), the ocean heat distribution comes into balance with the surface fluxes, so that any signal seen in the next stage can be attributed to forcing (rather than model drift due to ongoing transient changes in the ocean heat distribution).

From 1860 to 2000, the applied greenhouse gas concentrations were based on observations, and from 2000 to 2100 they were based on the Special Report on Emissions Scenario A1B (Nakicenovic and Swart 2000). The climate model includes a sulfate aerosol model; the level of sulfate emissions is also based on the A1B scenario. The regional atmospheric model ensemble was used to downscale the global transient results for the period 1950–2100 (for further details, see Murphy et al. 2009, chapter 3). Only 11 of the 17 regional atmospheric model ensemble members were selected to drive the surge model to provide the surge projections for TE2100 (Howard et al. 2008) and UKCP09 (Marine; Lowe et al. 2009). The other six RCM simulations were found to be deficient in their simulations of storms, and this was traced to the influence of a parameter perturbation that did not scale satisfactorily between the GCM and RCM simulations. Full details of this can be found in Murphy et al. (2009, chapter 3).

The domain of the RCM is chosen with the aim of satisfying two conditions (Jones et al. 1995): 1) that the RCM circulation is constrained to stay close to the driving GCM on scales that are skillfully resolved by the GCM and 2) that at finer scales the development of features in the RCM is not damped or distorted by the lateral boundary forcing. Hourly winds and surface pressure from HadRM3 are used to drive the Proudman Oceanographic Laboratory 12-km depth-averaged storm surge model of the continental shelf (POLCS3).

Implicit in this approach is the requirement that the physical link between large-scale climate processes and small-scale effects is strong. That this requirement is fulfilled in practice, at least for "well flushed" (Weisse and von Storch 2009) regions, such as northwest Europe, has been demonstrated by Denis et al. (2002). A second implicit assumption is that the feedback from the small to large scale is important only in its long-term statistics rather than hour by hour, since only the long-term statistics can be represented by the parameterizations in the global model.

An alternative to this "dynamical downscaling" approach is statistical downscaling (see, e.g., van der Linden and Mitchell 2009; Weisse and von Storch 2009; Haylock et al. 2006), which aims to provide local detail based on statistical relationships between the large-scale climate and local effects. We investigate this alternative for our particular problem of storm surge modeling in section 5a.

Effect of spatial resolution of the atmospheric model

The surge trend results projected in TE2100 (Howard et al. 2008) and in UKCP09 (Marine) (Lowe et al. 2009) are based on the POL CS3 12-km surge model driven by

a 25-km regional atmospheric climate model, which in turn is driven by a coupled global climate model with a grid size of approximately 2.5° × 3.75°, as described above.

It is of interest to know whether useful surge projections could have been made without the use of the regional model for three reasons: first, because running a regional climate model is expensive; second, because of the appeal of simplicity and the exclusion of unnecessary detail; and third, because there exist many other global climate model simulations that are not downscaled with a regional model, but which could, in principle, be used to drive a surge model directly if the regional model downscaling step could be shown to be unnecessary for the location of interest. With our simulation using both global and regional models, we are privileged to be able to address this question to inform future work on surge modeling for the Thames estuary. Sterl et al. (2009) use global climate model data to force a surge model of the North Sea. In discussing this they say,

> . . . water level is the integral result of the forcing over the whole North Sea. We therefore anticipated that the water level is primarily determined by the large-scale pressure and wind fields, which are well-represented in global models, and that the fine structure of the wind field is only important for individual cases, but not for the long-term statistics.

We investigated this assumption in our model by forcing the surge model again, this time with 10-m wind and sea level pressure data taken directly from the global climate model. In both the global and regional cases, we used data taken from the driving models once per model hour, with no averaging performed.

We know that small differences in timing of a meteorological event relative to the astronomical tide can lead to large differences in surge when tidal interaction is included (see section 5c, particularly Fig. 10). We anticipate some differences in the timing of meteorological events between the global and regional models given that the regional model is correctly setup. So as not to be excessively constrained by its boundaries (as discussed above) and to study the effect on surge of changing from regional to global forcing separately from the effects on the tide of changes in meteorological timing, we performed both simulations with the surge model in "surge only" mode (i.e., without any tidal forcing). Thus, *in section 5a*, the word "surge" is unambiguous: it refers to the modeled water elevation in the absence of the tide.

3. Surge modeling system

POL CS3 is used operationally to provide coastal flood warning in the United Kingdom under the auspices of UK Coastal Monitoring and Forecasting. The model produces a numerical solution of the discretized nonlinear shallow water equations with friction, and a lower-resolution version is described in detail by Flather (1976), Williams and Flather (2000), and Flather and Williams (2004). Validation of the operational model is performed monthly by comparison with observed sea level data from the U.K. National Tide Gauge Network and an annual summary of performance is published (UKCMF 2009). The operational model has been shown to perform particularly well during extreme storm surges in the southern North Sea (Horsburgh et al. 2008). Here we have a different arrangement: the surge model is driven by a climate model instead of an operational weather forecasting model or reanalyzed data. So, independent validation of this different arrangement was made against observations. The results of this validation are reported in Howard et al. (2008) and Lowe et al. (2009).

The tide–surge model covers the entire northwestern European continental shelf, as shown in Fig. 1. Tidal input at the model open boundaries consists of the 26 largest constituents. Modeled surges are derived by subtracting a tidal model simulation from one forced by both gravitational tide (applied as a forcing at the surge model boundaries) and atmospheric forcing from the regional climate model. Surge lateral boundary conditions are applied by adding the so-called inverse barometer effect to the gravitational tide level at lateral boundaries.

The model of the relationship between near-surface wind speed vector (taken from the atmospheric model) and surface wind stress vector $\boldsymbol{\tau}$ (applied to the coastal shelf model) is based on the formulation suggested by Charnock (1955) and can be described by an implicit equation:

$$\frac{1}{C_d} = \left[\frac{1}{\kappa} \log \frac{g\,z}{\beta C_d U^2} \right]^2,$$ (1)

which is solved iteratively for the drag coefficient C_d, where

$$\tau = C_d \varrho_a \mathbf{U} U;$$ (2)

$\mathbf{U}$ is the wind vector at height z above the sea surface; $U = |\mathbf{U}|$ is the wind speed at height z above the sea surface; g is the acceleration due to gravity; κ is von Kármán's constant (taken to be 0.41); ϱ_a is the density of air at mean sea level; and β, chosen by tuning the operational surge model for optimal surge results, is 0.0275; z is always 10 m in this work.

FIG. 1. Domain of the surge model (shaded). The outer square shows the RCM domain, and the cross shows latitude 55°N, longitude 5°E—the location referred to in the discussion of wind speed verification.
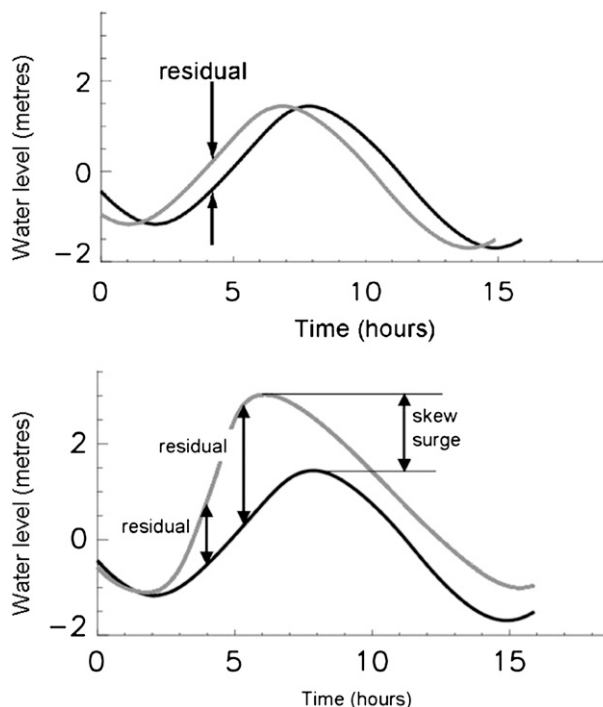

FIG. 2. Schematics showing skew surge. In each panel, the black curve shows the astronomical tide and the gray curve shows the sea level, including meteorological forcing. (top) Earlier arrival of the tide during a positive surge may be associated with substantial residuals that nevertheless have little flooding effect. (bottom) Schematic contrasting residual and skew surge (in this case for a high-impact positive surge event).

4. The use of skew surge as a diagnostic

In the simulations used to create the projections given in TE2100 (Howard et al. 2008) and UKCP09 (Marine; Lowe et al. 2009), the surge model is run with meteorological forcing from the regional climate model, producing the combined response to winds, surface pressure gradients, and tides. This captures the tide–surge interaction, where the principal effect of the surge on the tide is to alter the times of high and low water and the effect of the tide on the surge is the modulation of surge production. The model is also run in tide-only mode without any meteorological forcing. One way of measuring the meteorological component of the sea surface elevation, then, is to subtract the tide-only elevation from the combined-response elevation. This difference is the nontidal residual, which for brevity we will refer to henceforth as the residual. Wind stress is most effective at generating surge in shallow water. Also, the earlier arrival of the tide during a positive surge may be associated with large but unimportant residuals (Fig. 2, top panel). Thus, peak residuals (the maximum value of the residual during the surge event) are consistently obtained 3–5 h prior to the predicted high water (Horsburgh and Wilson 2007). For this reason, a more significant and practical measure than the peak residual is the skew surge (Fig. 2, bottom panel), which is the difference between the elevation of the predicted astronomical high tide and the nearest experienced high water (de Vries et al. 1995).

An indication of the improved statistical usefulness of skew surge over residual is given in Fig. 3, which shows a scatterplot of 693 modeled skew surges against tidal elevation at the Thames grid box for a 1-yr simulation and a corresponding plot for the 611 largest residuals from the same year (subject to the requirement of 12-h separation between any two). There is an insignificant correlation between skew and tide (P value >90%), but a significant negative correlation between residual and tide (P value <1%). This is associated with nonlinear tide–surge interaction (e.g., Horsburgh and Wilson 2007), but a simple physical interpretation is that residuals are greater close to low water.

5. Results

a. Correspondence of extremes in regionally and globally driven simulations

We consider a region of coastline around the southern North Sea and regress 19 yr of the regional-model-forced hourly elevations against the global-model-forced elevations for the same hour. We could suppose that the
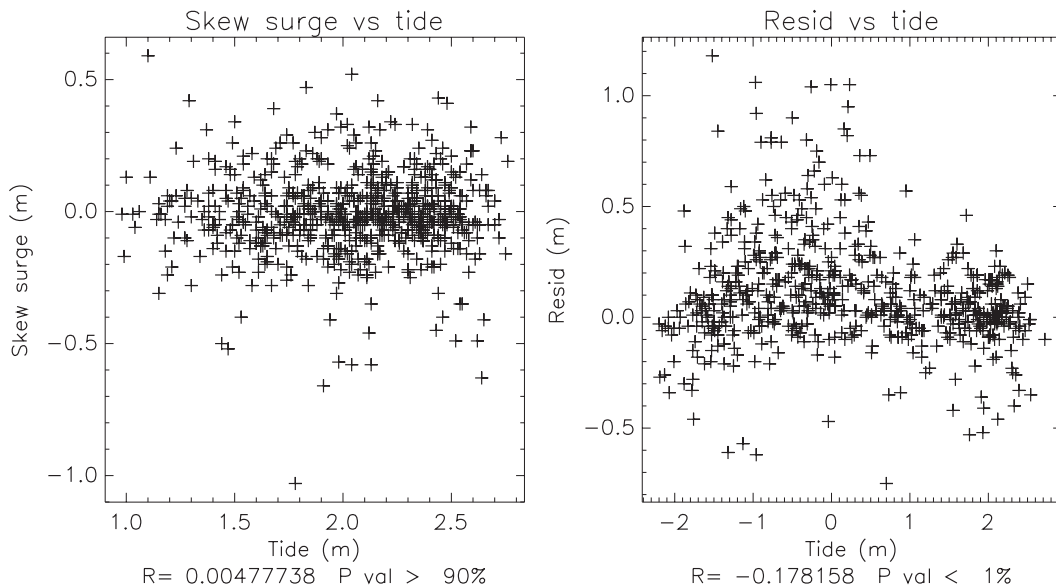
FIG. 3. Scatterplot to investigate correlation between modeled tidal elevation and two indicators of storm surge—(left) skew surge and (right) residual—for the Thames grid box of POL CS3.

regionally driven elevations (think of the y axis) represent "truth" and the globally driven elevations (think of the x axis) represent deviations from this truth. Alternatively, the data could be viewed as realizations of $y = \alpha x$ in which both x and y are subject to independent, symmetrically distributed random errors. In this case, each approach (either globally or regionally driven) would be an equally valid way of simulating surges (albeit the surges are on different scales in the global and regional simulations), with both approaches having a level of noise (again, on different scales). Leng et al. (2007) discuss approaches to regression in the case where both variables are subject to error. Following their recommendations, we find geometric-mean regression to be appropriate for our data. Since we know that in conditions of flat calm there would be no elevation under either forcing, it is inappropriate to allow a nonzero intercept for the regression line. When we insist that the intercept be zero, the gradient of the geometric-mean regression line becomes

$$\alpha = \text{sign}[\text{cov}(x, y)]\sqrt{\frac{\text{var}(y)}{\text{var}(x)}}. \qquad (3)$$

This quantity is mapped in Fig. 4. Gradient, α, varies slowly with location between 1.3 and 1.5 within the region chosen. Thus, the modeled surges forced by the global model are smaller than those forced by the regional model. To investigate the causes of this, we compare the distribution of surface wind speeds in the two models with each other and also with data presented by Sterl et al. (2009) for a location

(55°N, 5°E) over the North Sea, in Fig. 5. For ease of reference, the location is indicated on Fig. 1. The top panel of Fig. 5 shows that the Hadley Centre model extreme wind speeds are smaller than most of the Ensemble Simulations of Extreme Weather Events under Nonlinear Climate Change (ESSENCE) and GCM extreme wind speeds presented by Sterl et al., but that the Hadley Centre RCM extreme wind speed distribution compares very well with the 40-yr European Centre for Medium-Range Weather Forecasts Re-Analysis (ERA-40) and National Centers for Environmental Prediction–National Center
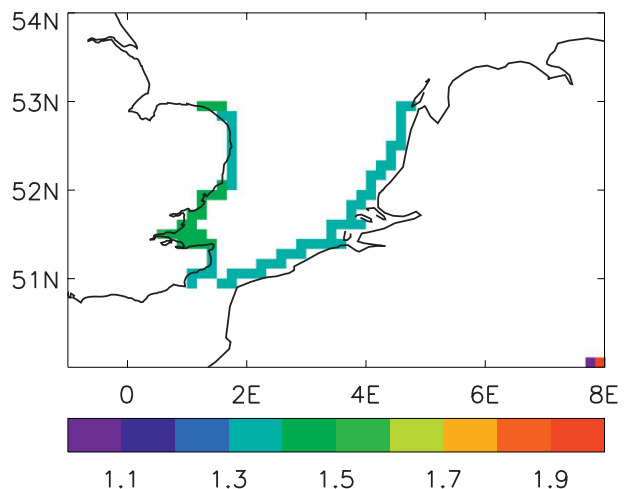

FIG. 4. Gradient of the fitted regression line of RCM-driven elevation against GCM-driven elevation for the model coastline grid cells of the southern North Sea.
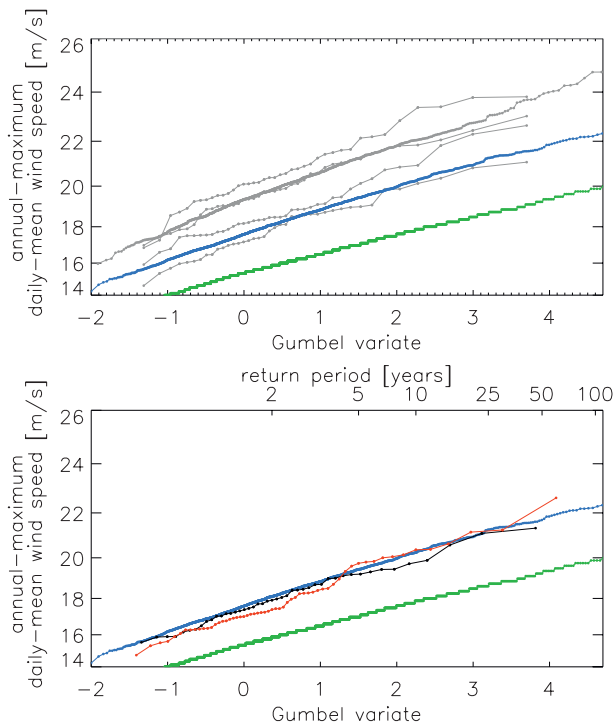
FIG. 5. Following Sterl et al. (2009), we present Gumbel plots of annual maximum daily-mean wind speed at 55°N, 5°E. (top) Gray lines: ESSENCE and GCMs data from Fig. 1 in Sterl et al. Green line: 17×150 yr of Hadley Centre global climate model perturbed physics ensemble treated as one simulation. Blue line: 11×150 yr of Hadley Centre RCM perturbed physics ensemble treated as one simulation. (bottom) Green and blue lines as in (top). Red line: NCEP reanalysis as presented by Sterl et al. Black line: ERA-40 as presented by Sterl et al. Only the 11 well-validating regional models are included in this assessment.

for Atmospheric Research (NCEP–NCAR) reanalyses. As shown in the bottom panel of Fig. 5, the Hadley Centre RCM distribution agrees with the reanalyses about as well as the reanalyses agree with each other.

However, the Hadley Centre global model extreme wind speeds, at least over this sea point, appear to be biased low; we suggest that this bias, assuming it applies more widely, will account for the unrealistically small surges in our global-model-driven simulations. The low bias in HadCM3 wind speeds at 55°N, 5°E, well offshore in the North Sea, makes an interesting contrast, not only with the results of Sterl et al. (2009), but also with some experiments investigating the "added value" (see, e.g., Weisse and von Storch 2009; Christensen et al. 2007) of regional climate models. Sotillo et al. (2005), Winterfeld (2008), and Winterfeld et al. (2010) all support the idea that regional climate modeling gives added value to marine near-surface winds near the coast, particularly where there is complex orography, but not over the open sea. In contrast we find a significant improvement at

55°N, 5°E over the open sea for our regional climate model, and this is consistent with our improved regional climate model surge simulations.

We are particularly interested in extreme positive surge events and the diagnosis of trends in these from a climate-model-driven simulation. To study this we focus again on the modeled elevation at the surge model grid point corresponding to the mouth of the Thames estuary. One approach is to regress the 19 regional-model-driven annual maxima at the mouth of the Thames against the corresponding 19 global-model-driven annual maxima at the mouth of the Thames (Fig. 6, top left panel). This is a simple approach but the uncertainties are large. We find the gradient of the fitted regression to be around 1.6 for the annual maxima and also for lower ranks—some of which are shown in the other panels of Fig. 6. This is consistent with the values found above for regression of the whole time series at this point (see Fig. 4).

In every case the correlation is highly significant (P value less than 0.5%).

b. Identification of trends in regionally and globally driven simulations

The major theme of the TE2100 (Howard et al. 2008) and UKCP09 (Marine; Lowe et al. (2009) surge work is the identification of statistically significant trends in the extremes of the surge distribution. Typically, significance is assessed by comparing the size of the estimated trend in the 50-yr return level with the uncertainty in this estimation. Return levels and trends are determined from the five largest events each year following the procedure advocated by Coles (2001, chapter 6). Not surprisingly, no such significant trend is found in our short (19 yr) simulation using either the global or the regional model. However, we can make random resamples of the 19 yr of surge height projections from the globally and regionally driven simulations (we make the same resampling of each set to preserve the global–regional correspondence, since it is this correspondence that we wish to study). Occasionally we will happen to resample them such that the trend determined from the resampled regionally driven simulation will be positive and statistically significant. If the resampled globally driven simulation also exhibits a positive trend with a similar level of significance, then we can infer that we need not have performed the regional model simulation to diagnose the existence of a trend in this resampled case.

After examining many such random resamples, we can produce a contingency table showing global versus regional diagnoses of positive, negative, or no trend. This table for the 6-yr return level is shown, against a background plot of the data for 10 000 random resamples, in Fig. 7. For each data point, each trend estimate has been
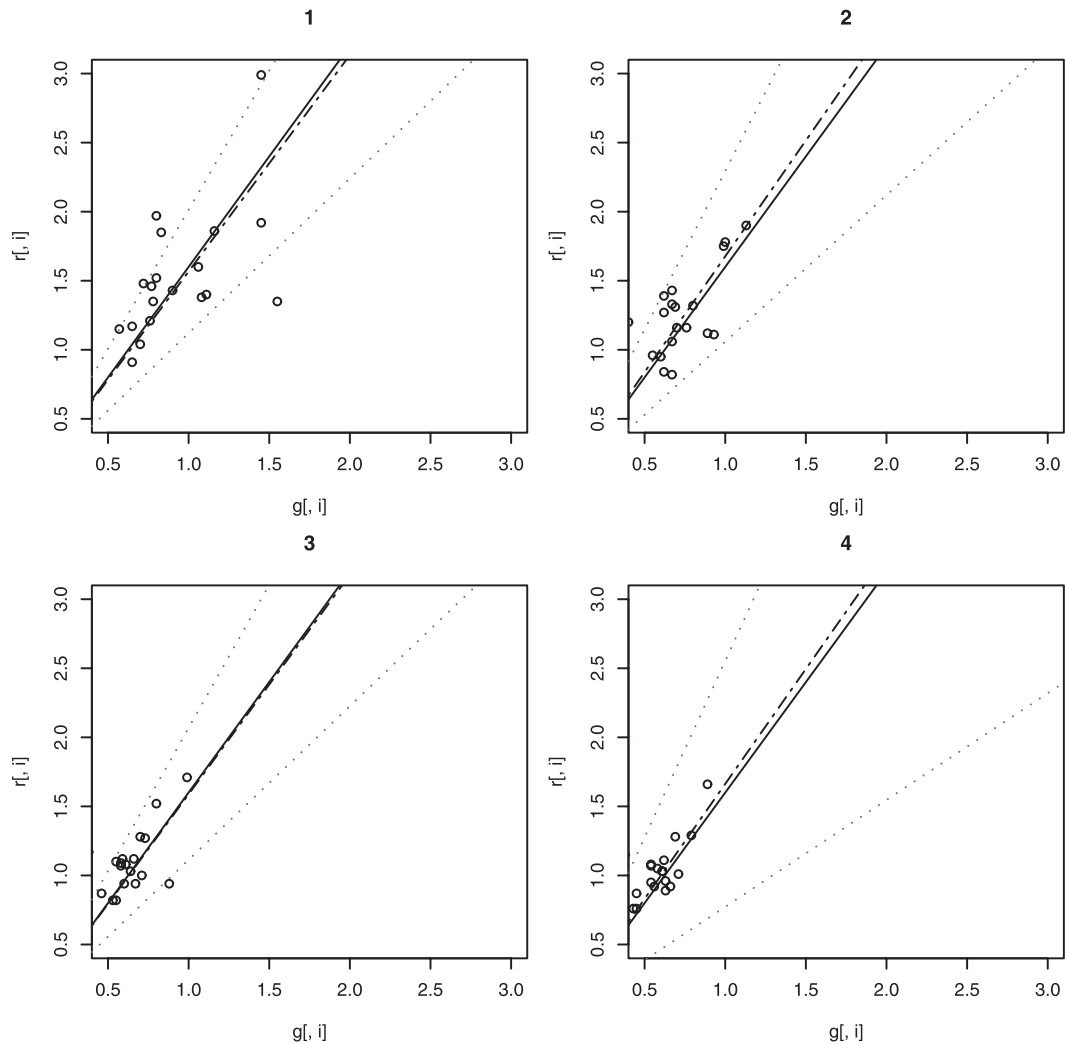
FIG. 6. (top left) Scatterplot and regression lines for the 19 annual maxima. The x axis shows annual maxima from the globally driven model; y axis shows annual maxima from the regionally driven model. The dashed line shows the maximum likelihood estimation of the gradient of the geometric mean regression. Dotted lines show the 90% confidence intervals. The solid line shows a gradient of 1.6. (top right and bottom) Similar plots for the annual second, third, and fourth largest modeled surges. Notice that in every case, there is good agreement between the solid and dashed lines.

normalized by its own uncertainty estimate (standard error) and thus a value of 2 represents a P value of approximately 5% under an assumption of normality. So, using a 5% test level, we would diagnose, for example, a positive trend from the globally driven simulation (x axis) but no trend from the regionally driven simulation (y axis) on 357 occasions (right-hand column, middle row) out of 10 000 comparisons (3.6%). Diagnosis of a positive trend from the globally driven simulation but a negative trend from the regionally driven simulation does not occur in this 10 000-strong sample. The converse occurs only once. We studied the 6-yr return level because the TE2100 (Howard et al. 2008) study focused on

the 50-yr return level; 6 yr is to our 19-yr sample what 50 yr is to a 149-yr sample: about one-third. Owing to the normalization, Fig. 7 is strongly symmetrical about the lines $y = x$ and $y = -x$. Certainly x and y are very strongly correlated (P value much less than 0.1%).

As with the simpler case of the annual maxima above, we seek a transfer relationship, but this time between the magnitude of a trend in the globally driven simulation and the magnitude of a trend in the regionally driven simulation, again using geometric-mean regression. Again we find the gradient of the regression to be 1.6, just as it was for the annual maxima, but this time with smaller uncertainties associated with the use of
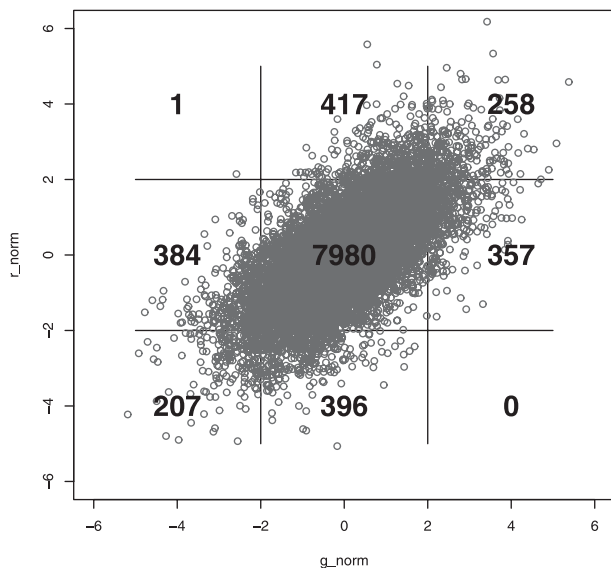
FIG. 7. Scatterplot of normalized trend in 6-yr return level for 10 000 random resamplings of 19 yr of globally driven (x axis) and regionally driven (y axis) modeled surges at the mouth of the Thames. The contingency table is superimposed. For full details see the main text.

more data: five ranks annually instead of just the annual maxima. For our data this result is robust across trends in different return levels (Fig. 8).

In summary, then, we find that for our mouth of the Thames location, our global climate model is suitable for driving our surge model *for the purpose of identifying century-scale changes* in extreme surge climate, and our evidence suggests that the signal-to-noise ratio in such trends will be similar whether the global or regional climate data are used to drive the surge model.

Furthermore, a factor of 1.6 is suitable for amplifying the globally driven extreme surges to reach the scale of the regionally driven extreme surges for our model and location, although other locations, other models, or other values of the Charnock parameter might require a different factor. This second finding makes an interesting contrast to the results of Sterl et al. (2009). Their globally driven surge model water levels at Hoek van Holland, Netherlands, are directly comparable with observations (Sterl et al. 2009, their Fig. 4). This contrast might be associated with their use of a larger value of the Charnock parameter, which has the effect of increasing the stress on the sea surface for a given wind in a one-way nested model. Sterl et al. use a Charnock parameter of 0.032; we use 0.0275. Or the contrast might be associated with the underprediction of wind speed (see section 5a, particularly Fig. 5) by our GCM, or both.

The validation reported by Howard et al. (2008) suggests that water-level statistics at the Thames estuary

produced by our simulation system in regional-model-driven mode are only slightly diminished compared to those derived from tide–gauge observations or from an ERA-40-based (Uppala et al. 2005) surge model simulation. This suggests that for the purpose of identifying trends at this location, the regional modeling step could be omitted, and trends diagnosed from a globally driven surge model could be scaled using observations.

One caveat in particular should be noted in the context of the regression coefficient: we ran the surge model in surge-only mode, for reasons discussed in section 2a. It is recognized (Horsburgh and Wilson 2007; Sterl et al. 2009) that the tide has a damping effect on surge, and thus this figure may be modified when the effect of the tide is included, since we might anticipate that the damping effect would be proportionally larger for the larger regional-model-driven surges.

c. Comparison of extreme simulated surges and the observational record

An important goal for climate-model-driven surge simulations is to produce extreme storm surges that are comparable in magnitude (and therefore effect) to those observed. The worst storm surge in recent history to affect the coasts of the United Kingdom and the Netherlands occurred during the night of 31 January 1953, when coastal flooding caused more than 2000 fatalities (Baxter 2005). The skew surge at Southend-on-Sea, at the mouth of the Thames estuary, during this event has been estimated from archived paper records to be 2 m above a predicted tide of 2.63 m above ordnance datum (G. Siggers 2007, personal communication; see acknowledgments) about 48 h after the full moon. (Ordnance datum is the U.K. national standard datum of vertical height, based on the mean sea level at Newlyn, Cornwall.) In contrast, the largest skew surge in our entire PPE ensemble simulation (which represents a total of 1639 model years) was 1.78 m, and this occurred during a tidal cycle where the modeled high water was 1.69 m. The synoptic situation in these two events was rather similar. In both cases the low moved eastward to the north of the United Kingdom and then southward approximately along the eastern shores of the North Sea. Charts showing mean sea level atmospheric pressure near the time of maximum elevation at the mouth of the Thames for both the 1953 event and the largest skew event in our PPE ensemble simulation are shown in Fig. 9. The similarity in these two snapshots is apparent, although the gradients are larger in the 1953 analysis, and the low center is farther west.

Since the passage of weather systems is independent of tide, it is reasonable to ask whether, with different timings of storm movement relative to the state of the tide and the fortnightly spring-neap cycle, the model
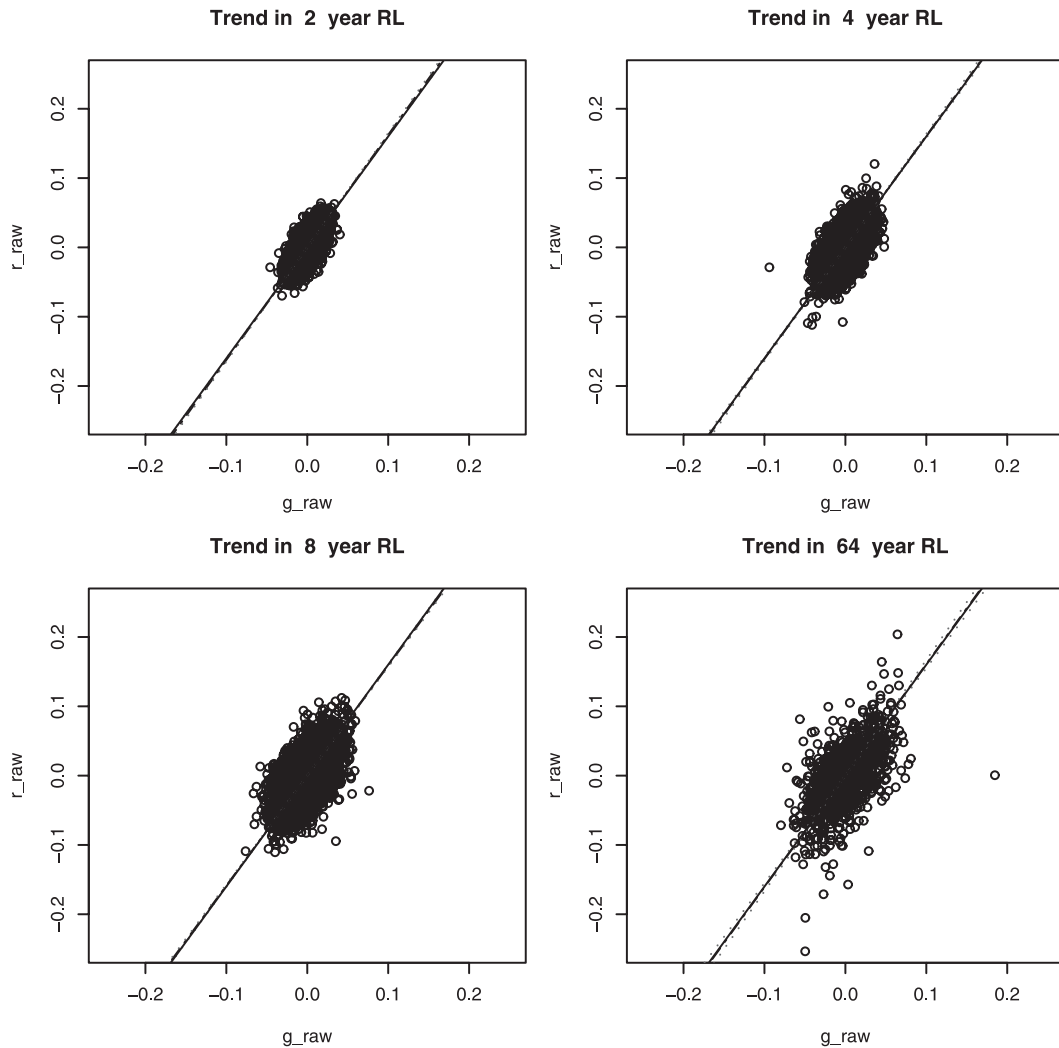
FIG. 8. (top left) Scatterplot as in Fig. 7, but *not* normalized, for the 2-yr return level. Regression lines are shown using the same key as in Fig. 6, but the agreements are so strong and the confidence intervals so small that most of the lines overlap. (top right and bottom) Similar plots for trend in the 4-, 8-, and 64-year return levels.

runs could simulate an extreme sea level in closer agreement to the observed 1953 event. It is possible to achieve this synthetically by deliberately phase shifting the tide with respect to the meteorological forcing. Horsburgh et al. (2008) compared the variability of this type of manipulation with that obtained from a 24-member short-range ensemble of surge model runs driven by a perturbed initial condition ensemble of local area models (Flowerdew et al. 2010). They found that altering the phase relationship between tide and weather for their particular event caused the skew surges at Southend-on-Sea to differ by only 5 cm, but we shall see below that much greater variation is possible with our more extreme meteorological forcing.

We adjusted the phase relationship between the meteorological forcing and the tide at hourly increments around several different selected tides. Residuals were derived in the usual way (by subtracting a "tide only" run) to capture the tide–surge interaction. As an example of the results of this sort of adjustment, the importance of nonlinear interactions is underscored in Fig. 10, where we can see that, for example, the same meteorological forcing that produces a residual of nearly 4 m (but a low impact) if it arrives at an early stage of the rising tide produces only a 2-m residual (but a high impact) if it arrives at high tide. This figure also makes a striking illustration of the improved impact relevance of skew surge over residual: for the high-impact event, the skew surge is large and the maximum residual is
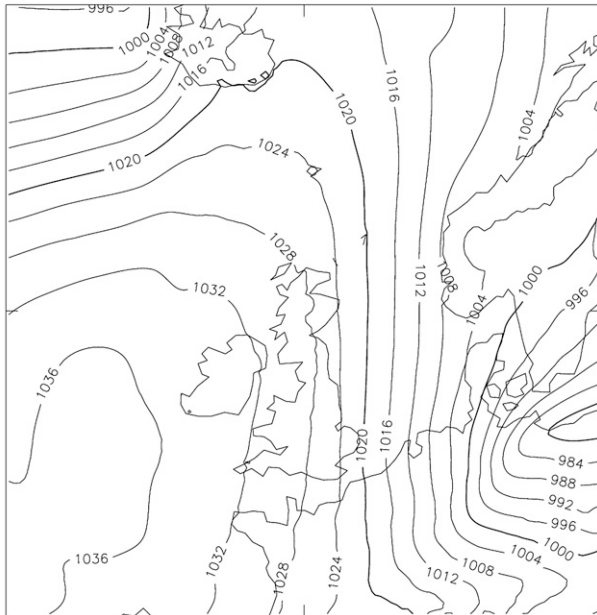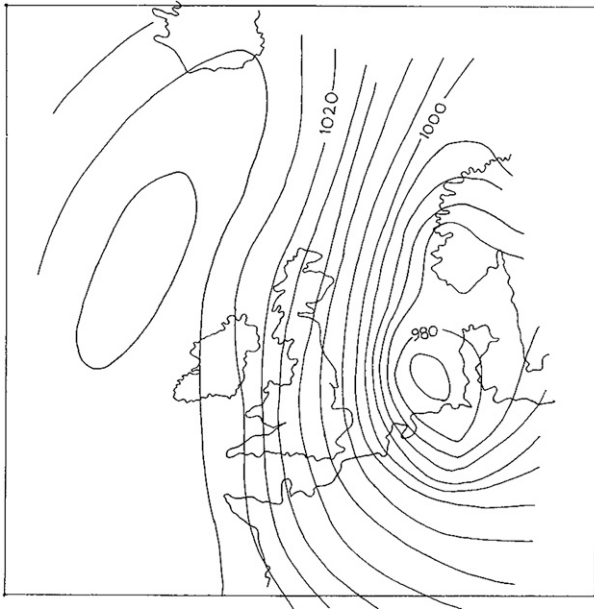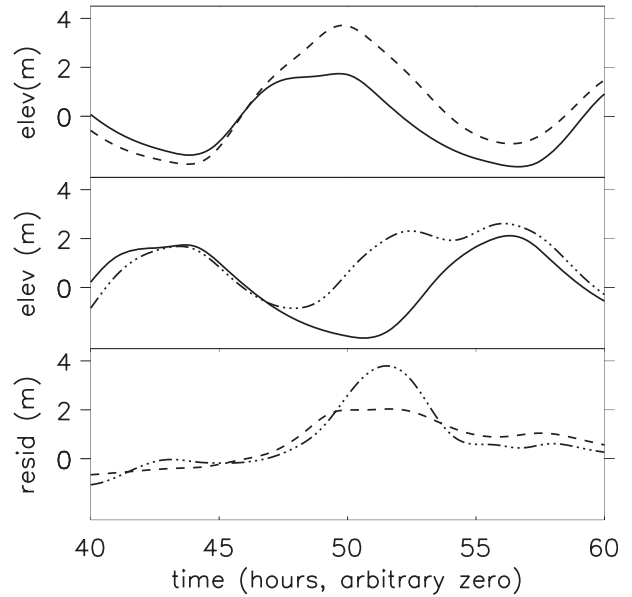
FIG. 10. Example of phase adjustment: the effect on water level at Southend-on-Sea of the same (extreme) meteorological forcing applied at different phases of the same tide. (top) A high-impact event. Solid line: tide only simulation; dashed line: water level, including meteorological forcing. An extreme sea level of nearly 4 m arises, reflected by a skew surge of nearly 2m. (middle) A low-impact event: the same meteorological forcing applied at a different phase of the same tide. Solid line: tide only simulation; dotted–dashed line: water level, including meteorological forcing. (bottom) Residuals of each event. Dashed line: high impact; dotted–dashed line: low impact.

FIG. 9. (top) Isobars of mean sea level atmospheric pressure for 0000 UTC 1 Feb 1953 (i.e., near the time of maximum elevation at the mouth of the Thames) based on the daily weather report (Met Office 1953). Isobar interval 4 hPa. (bottom) Isobars of mean sea level atmospheric pressure near the time of the most extreme simulated skew surge at the mouth of the Thames in the PPE. Isobar interval 4 hPa.

small; for the low-impact event, the skew surge is small and the maximum residual is large.

Extreme water levels, astronomical high tides, and the emergent skew surge for the same meteorological forcing applied at four different phases of tide are summarized in Table 1. After phase adjustment around its associated

modeled tide of 1.69 m, the weather system giving the largest skew surge overall in our ensemble now produced a skew surge of 2.0 m (case A in Table 1). When the same storm was subsequently phase shifted around a modeled spring tide of 2.35 m, the maximum skew surge obtained was 1.82 m, corresponding to a total water level of 4.17 m (case A1C in Table 1).

One of our selected model tides had a very similar astronomical high (2.65 m) to that in the analysis by G. Siggers (2007, personal communication) based on observations of the 1953 event. The model tide-only tidal range was also very similar (see Fig. 12). After phase adjustment around this tide, the weather system giving the largest skew surge overall in our ensemble now produced a skew surge of 1.7 m (case "A2" in Table 1). A comparison of this phase-adjusted modeled event and the analysis by Siggers is shown in Figs. 11 and 12.

In the analysis of the 1953 event presented by Horsburgh et al. (2008), the maximum water level is 4.57 m. The corresponding model maximum is 4.37 m. Siggers, in his analysis, puts the peak level at 4.63 m and estimates the observational uncertainty in the peak levels to be around 0.2 m. Thus, we can say that our simulated extreme water

TABLE 1. Results of sensitivity tests in which the surge–tide phase relationship was adjusted. The same meteorological forcing is used in each of the model simulations. It is the largest meteorological forcing found in all of the Met Office Hadley Centre ensemble runs, as diagnosed by the size of the skew surge. The rows labeled "Obs" show analyses of paper records of the 1953 event from Horsburgh et al. (2008; H) and G. Siggers (2007, personal communication; S). The model cases are as follows: R: raw (1.69-m tide). This is the simulation exactly as it arose in the ensemble. A: Adjusted (1.69 m tide). Here we have adjusted the phase relationship to maximize the skew surge on a tide of 1.69 m (i.e., the same tide as case R). A1C: adjusted (2.35-m tide) with correction. Here we have adjusted the phase relationship to maximize the skew surge on a tide of 2.35 m (i.e., the model underrepresentation of the tide during the 1953 event). The figure in parenthesis (4.61) is the elevation given by applying the tidal correction that is used operationally (see main text). A2: Adjusted (2.65-m tide). Here we have adjusted the phase relationship to maximize the skew surge on a tide of 2.65 m (i.e., a model tide that approximates the tide during the 1953 event, without the tidal correction). All figures are meters; elevations are above ordnance datum Newlyn. Here "tide" refers to the astronomical component of sea level.

Source	Case	High tide	Skew surge	Max elevation
Model	R	1.69	1.78	3.47
	A	1.69	2.00	3.69
	A1C	2.35	1.82	4.17 (4.61)
	A2	2.65	1.71	4.37
Obs	H	2.47	2.10	4.57
	S	2.63	2.0	4.63

level is on the low side of the 1953 event analyses but comparable given the observational uncertainty. It could be argued that by adjusting the phase, we have effectively increased the empirical return period (because we have adjusted the joint probability of surge and tide) so that it is now substantially greater than the total number of simulated years in the entire ensemble, which is 1639 yr, whereas the 1953 event is usually regarded as a less-than-500-yr event. This is consistent with our more comprehensive validation, based on a comparison of model and observation statistics for ports around the U.K. coastline, shown in TE2100 (Howard et al. 2008). It is also consistent with a more general paradigm of increasing realism as we move from a "globally driven" simulation [unrealistically small surges (section 5a)] to a "regionally driven" simulation (more realistic but still small) and finally to an analysis of observations (most realistic).

d. The effect of large mean sea level rise on modeled surge propagation

All long-term (149 yr) surge simulations performed for the TE2100 (Howard et al. 2008) and UKCP09 (Marine; Lowe et al. 2009) projects used present-day bathymetry in the surge model. The projected mean sea level changes were then added linearly to any trend in
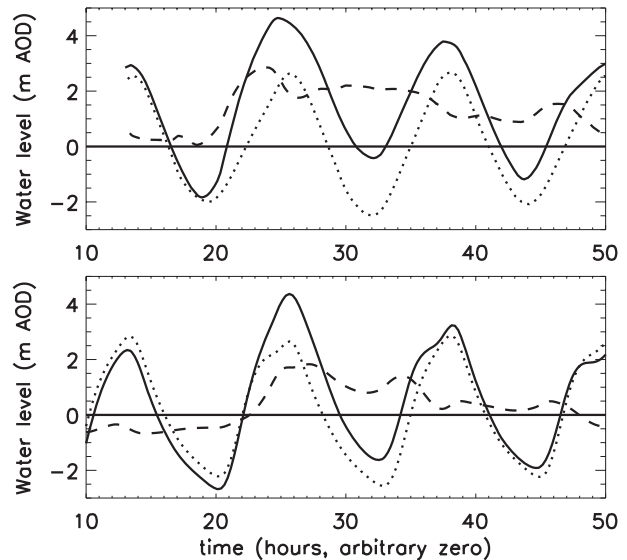

FIG. 11. Comparison of an analysis based on observations of the 1953 event with modeled water level time series for an extreme meteorological event imposed on a very similar modeled astronomical tide. (top) Analysis based on 1953 observations. (bottom) Model. Solid line: total water level; dotted line: astronomical tide; and dashed line: residual. The zero of time is 0000 31 Jan 1953 for the analysis, arbitrary for the model.

extreme surges that was identified. This is valid only if the effect of mean sea level changes on surge is linear, to a good approximation. Lowe et al. (2001) found that to a first-order approximation, time-mean sea level rise and changes in surge can be added linearly around the United Kingdom for mean sea level increase up to about 0.5 m. Sterl et al. (2009) obtained a similar result for mean sea level increase up to 2 m. Here we further test this assumption for our location of special interest at the Thames estuary by comparing surges with the same forcing but different bathymetry representing mean sea level increase up to 5 m. While it is considered extremely unlikely that such an increase will occur within the next century (Pfeffer et al. 2008, e.g.), it is consistent with a complete collapse of the West Antarctic Ice Sheet (which might occur on a longer time scale) and impact studies such as that described by Dawson et al. (2005) have used comparable values.

In the first case, we use our standard present-day bathymetry. In the second case, we add 5 m to the bathymetry, everywhere in the domain (i.e., we make the model sea 5 m deeper). In reality under a 5-m increase of mean sea level, the area covered by water would also change; we suggest that this would have a negligible effect on surge–tide interaction, since the area increase would apply to water of depth 5 m or less. We assume that the bulk of the tide and surge energy is carried in deeper water. In each
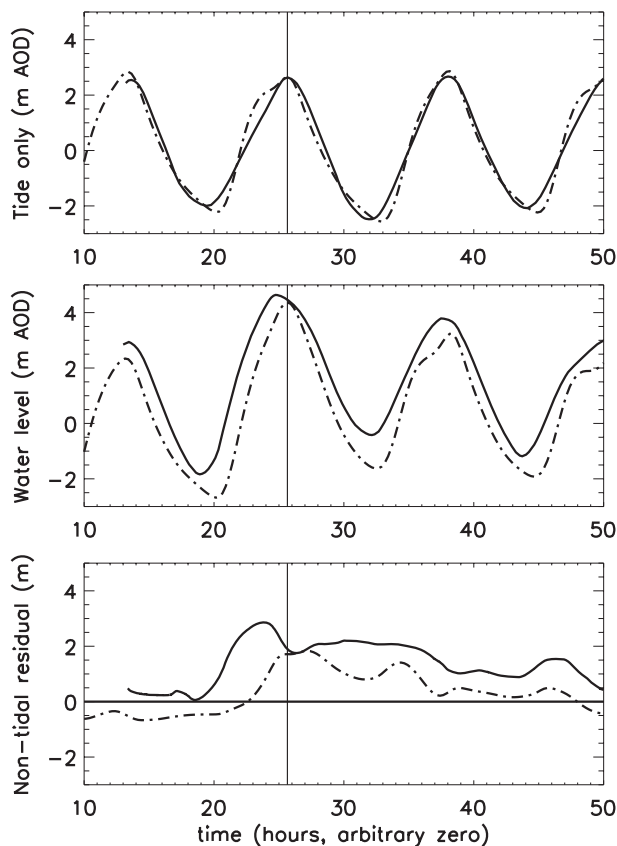
FIG. 12. Alternative presentation of the same data as in Fig. 11 to facilitate comparison. (top) Tide only. (middle) Water level. (bottom) Residual. Solid line: analysis of 1953 event; dotted–dashed line: model event.

case we also produced a tide-only simulation (i.e., without meteorological forcing).

The results in terms of the sea surface elevation time series at the Thames grid box are shown in Fig. 13. It can be seen that the primary effect is on the timing: in both meteorology-forced (shown in Fig. 13) and tide-only mode (not shown, but the pattern of change is very similar), the signal arrives about one hour earlier with the increased bathymetry. This change is physically reasonable—we would expect the tide to travel faster on deeper water— and its magnitude is consistent with the following approximate calculation. Recall that the horizontal scale of the tide is large relative to the depth of the oceans and consequently the tide is well approximated as a shallow water wave with wave speed (phase speed) $\sqrt{gh}$, where g is gravitational acceleration and h is the water depth. Assume that the tide begins its journey from Wick (approx 58°N, 3°W) at about the same time in either case. This is not an unreasonable approximation since we do not change the timing of the tidal boundary forcing and the water is, typically, deep enough between the northern and
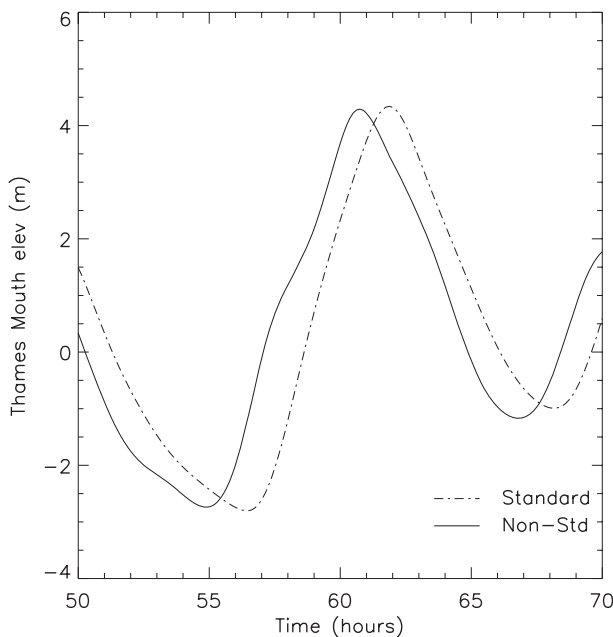

FIG. 13. The modeled primary effect of a 5-m increase in mean sea level is a change in the timing of arrival of high water at the mouth of the Thames.

eastern boundaries of the surge model domain and Wick that we may neglect the effect of the bathymetry increase for that part of the journey. Suppose the tide currently travels from Wick to Southend-on-Sea (about 860 km) in 13.5 h, giving a shallow-water phase speed of about 17.7 m s^{-1} and an implied shallow-water depth of about 32 m. Adding 5 m to give a new depth of 37 m gives a new shallow-water phase speed of about 19 m s^{-1}, allowing the tide to arrive about 57 min earlier at Southend-on-Sea for the same start time at Wick.

The effect on the total sea surface elevation is less than 10 cm, with a slightly larger effect on the skew surge; however, in either case, this may be due to the faster tide producing a change in the tide–surge phase relationship (see section 5c) and so should not be interpreted as a systematic change brought about by the change in the water depth. Smaller changes in bathymetry (not shown) produce correspondingly smaller changes in timing and elevation.

6. Conclusions

Our main conclusion is this: to diagnose the presence or absence of statistically significant trends in surge at the mouth of the Thames estuary using a coastal shelf model driven by the Met Office Hadley Centre global climate model HadCM3, we find that it is not necessary to downscale using a regional model. However, to

simulate realistic levels of surge using HadCM3, we find that the regional downscaling step is necessary. If trends in the globally driven simulation are found, then a comparison of globally and regionally driven simulations provides a scaling factor to interpret the trends to real-world scale. This scaling factor varies only slowly with location around the southern North Sea. An alternative would be to use observations to derive the scaling factor. This again would sidestep the need for a regional model.

We also find that simulated large increases in mean sea level (up to 5 m) have very little effect on extreme surges relative to the astronomical tide at the Thames estuary, the primary effect being on the speed of propagation of tide and surge.

Finally, we support the use of skew surge as a more significant and practical measure of the effect of an extreme event.

REFERENCES

Baxter, P., 2005: The east coast Big Flood, 31 January–1 February 1953: A summary of the human disaster. *Philos. Trans. Roy. Soc. London,* **A363,** 1293–1312, doi:10.1098/rsta.2005.1569.

Charnock, H., 1955: Wind stress on a water surface. *Quart. J. Roy. Meteor. Soc.,* **81,** 639–640.

Christensen, J., T. Carter, M. Rummukainen, and G. Amanatidis, 2007: Evaluating the performance and utility of regional climate models: The PRUDENCE project. *Climatic Change,* **81,** 1–6, doi:10.1007/s10584-006-9211-6.

Coles, S., 2001: *An Introduction to Statistical Modeling of Extreme Values.* Springer, 208 pp.

Collins, M., B. Booth, G. Harris, J. Murphy, D. Sexton, and M. Webb, 2006: Towards quantifying uncertainty in transient climate change. *Climate Dyn.,* **27,** 127–147.

Dawson, R. J., J. W. Hall, P. D. Bates, and R. J. Nicholls, 2005: Quantified analysis of the probability of flooding in the Thames estuary under imaginable worst-case sea level rise scenarios. *Water Resour. Dev.,* **21,** 577–591.

Denis, B., R. Laprise, C. Caya, and J. Côté, 2002: Downscaling ability of one-way nested regional climate models: The Big-Brother experiment. *Climate Dyn.,* **18,** 627–646.

de Vries, H., and Coauthors, 1995: A comparison of 2D storm surge models applied to three shallow European seas. *Environ. Software,* **10,** 23–42.

Flather, R., 1976: A tidal model of the north-west European continental shelf. *Mem. Soc. Roy. Sci. Liege,* **10,** 141–164.

——, and J. Williams, 2004: Future development of operational storm surge and sea level prediction. Proudman Oceanographic Laboratory Internal Doc. 165, 73 pp.

Flowerdew, J., K. Horsburgh, C. Wilson, and K. Mylne, 2010: Development and evaluation of an ensemble forecasting system for coastal storm surges. *Quart. J. Roy. Meteor. Soc.,* **136,** 1444–1456.

Gordon, C., C. Cooper, C. Senior, H. Banks, J. Gregory, T. Johns, J. Mitchell, and R. Wood, 2000: The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dyn.,* **16,** 147–168.

Haylock, M., G. Cawley, C. Harpham, R. Wilby, and C. Goodess, 2006: Downscaling heavy precipitation over the UK: A comparison of dynamical and statistical methods and their future scenarios. *Int. J. Climatol.,* **26,** 1397–1415.

Horsburgh, K. J., and C. Wilson, 2007: Tide-surge interaction and its role in the distribution of surge residuals in the North Sea. *J. Geophys. Res.,* **112,** C08003, doi:10.1029/2006JC004033.

——, J. A. Williams, J. Flowerdew, and K. Mylne, 2008: Aspects of operational forecast skill during an extreme storm surge event. *J. Flood Risk Manage.,* **1,** 213–221.

Howard, T., J. A. Lowe, A. Pardaens, J. Ridley, and K. Horsburgh, 2008: Met Office Hadley Centre projections of 21st century extreme sea levels for TE2100. Met Office Hadley Centre Rep., Phase 2 EP17 Study, 54 pp.

Jones, R., J. Murphy, and M. Noguer, 1995: Simulation of climate change over Europe using a nested regional-climate model. 1. Assessment of control climate, including sensitivity to location of lateral boundaries. *Quart. J. Roy. Meteor. Soc.,* **121,** 1413–1449.

Leng, L., T. Zhang, L. Kleinman, and W. Zhu, 2007: Ordinary least square regression, orthogonal regression, geometric mean regression and their applications in aerosol science. *J. Physics,* **78,** 012084, doi:10.1088/1742-6596/78/1/012084.

Lowe, J. A., J. Gregory, and R. Flather, 2001: Changes in the occurrence of storm surges around the United Kingdom under a future climate scenario using a dynamic storm surge model driven by the Hadley Centre climate models. *Climate Dyn.,* **18,** 179–188.

——, and Coauthors, 2009: Marine and coastal projections. Met Office Hadley Centre UK Climate Projections Rep., 99 pp. [Available online at http://ukclimateprojections.defra.gov.uk/images/stories/marine_pdfs/UKP09_Marine_report.pdf.]

Met Office, 1953: Daily weather report. Met Office Rep. issued 1 February, 4 pp.

Murphy, J. M., B. B. B. Booth, M. Collins, G. R. Harris, D. M. H. Sexton, and M. J. Webb, 2007: A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles. *Philos. Trans. Roy. Soc. London,* **A365,** 1993–2028, doi:10.1098/rsta.2007.2077.

——, and Coauthors, 2009: Climate change projections. Met Office Hadley Centre UK Climate Projections Science Rep., 194 pp. [Available online at http://ukclimateprojections.defra.gov.uk/images/stories/projections_pdfs/UKCP09_Projections_V2.pdf.]

Nakicenovic, N., and R. Swart, Eds., 2000: *Special Report on Emissions Scenarios.* Cambridge University Press, 599 pp.

Pfeffer, W. T., J. T. Harper, and S. O'Neel, 2008: Kinematic constraints on glacier contributions to 21st-century sea-level rise. *Science,* **321,** 1340–1343.

Pope, V., M. Gallani, P. Rowntree, and R. Stratton, 2000: The impact of new physical parametrizations in the Hadley Centre climate model: HadAM3. *Climate Dyn.,* **16,** 123–146.

Sotillo, M., A. Ratsimandresy, J. Carretero, A. Bentamy, F. Valero, and F. González-Rouco, 2005: A high-resolution 44-year atmospheric hindcast for the Mediterranean Basin: Contribution to the regional improvement of global reanalysis. *Climate Dyn.,* **25,** 219–236.

Sterl, A., H. van den Brink, H. de Vries, R. Haarsma, and E. van Meijgaard, 2009: An ensemble study of extreme North Sea storm surges in a changing climate. *Ocean Sci.,* **5,** 369–378.

Stott, P., S. Tett, G. Jones, M. Allen, J. Mitchell, and G. Jenkins, 2000: External control of 20th century temperature by natural and anthropogenic forcings. *Science,* **290,** 2133–2137.

TE2100, cited 2010: Thames Estuary 2100. [Available online at http://www.environment-agency.gov.uk/homeandleisure/floods/104695.aspx.]

UKCMF, cited 2009: UK Coastal Monitoring and Forecasting operational report to the Environment Agency for the period 1st June 2008 to 31st May 2009. 48 pp. [Available online at http://www.metoffice.gov.uk/publicsector/environmental/UKCMF_annual_report_2008-9.pdf.]

Uppala, S. M., and Coauthors, 2005: The ERA-40 Re-Analysis. *Quart. J. Roy. Meteor. Soc.,* **131,** 2961–3012, doi:10.1256/qj.04.176.

van der Linden, P., and J. F. B. Mitchell, Eds., 2009: Climate change and its impacts at seasonal, decadal and centennial timescales: Summary of research and results from the ENSEMBLES project. Met Office Hadley Centre ENSEMBLES Tech. Rep., 160 pp.

Weisse, R., and H. von Storch, 2009: *Marine Climate and Climate Change*. Springer, 219 pp.

Williams, J., and R. Flather, 2000: Interfacing the operational storm surge model to a new mesoscale atmospheric model. Proudman Oceanographic Laboratory Internal Document 127, 20 pp.

Winterfeld, J., 2008: Comparison of measured and simulated wind speed data in the North Atlantic. Ph.D. thesis, University of Hamburg, 116 pp. [Available online at http://dvsun3.gkss.de/BERICHTE/GKSS_Berichte_2008/GKSS_2008_2.pdf.]

——, B. Geyer, and R. Weisse, 2010: Using QuickSCAT in the added value assessment of dynamically downscaled wind speed. *Int. J. Climatol.,* in press.