Clim Dyn (2009) 32:189–203
DOI 10.1007/s00382-008-0442-2

Detection of external influence on trends of atmospheric storminess and northern oceans wave heights

Xiaolan L. Wang · Val R. Swail · Francis W. Zwiers · Xuebin Zhang · Yang Feng

Received: 24 May 2007 / Accepted: 26 June 2008 / Published online: 16 July 2008

Abstract The atmospheric storminess as inferred from geostrophic wind energy and ocean wave heights have increased in boreal winter over the past half century in the high-latitudes of the northern hemisphere (especially the northeast North Atlantic), and have decreased in more southerly northern latitudes. This study shows that these trend patterns contain a detectable response to anthropogenic and natural forcing combined. The effect of external influence is found to be strongest in the winter hemisphere, that is, in the northern hemisphere in January–March and in the southern hemisphere in July–September. However, the simulated response to anthropogenic and natural forcing combined, which was obtained directly from climate models in the case of geostrophic wind energy and indirectly via an empirical downscaling procedure in the case of ocean wave heights, is significantly weaker than the magnitude of the observed changes in these parameters.

Keywords Atmospheric storminess · Ocean wave heights · Climate extremes · Non-stationary generalized extreme value analysis · Trend analysis · Natural and anthropogenic forcing · Detection analysis · Climate change

X. L. Wang (✉) · V. R. Swail · F. W. Zwiers · X. Zhang · Y. Feng
Climate Research Division, Atmospheric Science and Technology Directorate, STB, Environment Canada, 4905 Dufferin Street, Toronto, ON M3H 5T4, Canada
e-mail: Xiaolan.Wang@ec.gc.ca

1 Introduction

After the hurricane Katrina disaster in 2005 and the tsunami tragedy that struck Asia at the end of 2004, public awareness of the potential impacts of storm surges and ocean waves on human society, the environment and ecosystems has increased. One of the issues of concern is whether or not external influences on the climate system, especially human influence, have affected the storm and ocean wave climate. Although this issue is yet to be addressed, there is evidence of significant change in extratropical cyclone activity and ocean wave heights in the boreal cold seasons of the last half century. For example, a significant increasing trend in winter (January–March) strong-cyclone activity over the high-latitude North Atlantic has been identified in the sea level pressure (SLP) fields taken from both the ERA40 reanalysis (ERA40 hereafter; Uppala et al. 2005) and the NCEP-NCAR Reanalysis (Kalnay et al. 1996; Kistler et al. 2001) for the 1958–2001 period. There is also a significant decreasing trend in strong-cyclone activity over the mid-latitude North Atlantic over this period (Wang et al. 2006). These changes are associated with a northward shift of the mean position of the North Atlantic storm track of about 180 km (Wang et al. 2006). Consistent with the increase in strong-cyclone activity, the northeast North Atlantic ocean has also been found to have roughened in winter during the 1958–1997 period, while significant decreases of ocean wave heights are identified in the subtropical North Atlantic (Wang and Swail 2006a; Wang and Swail 2001, 2002; WASA Group 1998; Bacon and Carter 1991).

There have been numerous studies on the detection, attribution, and quantification of the influence of external forcing on a range of climate variables, such as surface air temperature, atmospheric pressure, free atmosphere

temperature, tropopause height, ocean heat content, sea ice extent, and precipitation (e.g., Hegerl et al. 1997; Tett et al. 1999; Zwiers and Zhang 2003; Stott 2003; IDAG 2005; Zhang et al. 2006; Gillett et al. 2003, 2005; Jones et al. 2003; Santer et al. 2003; Barnett et al. 2005; Hegerl et al. 2007; Zhang et al. 2007). However, previous studies have not addressed the question of whether external factors have influenced atmospheric storminess and ocean wave heights in the extra-tropics. This study attempts to do exactly that by comparing observationally based estimates of changes in atmospheric storminess and ocean wave heights with estimates derived from multi-model simulations of climate change driven with historical external forcing from both anthropogenic sources (e.g., greenhouse gases and aerosol forcing) and natural sources such as solar and volcanic forcing.

This article is structured as follows. The data sets used and their preparation are described in Sect. 2. The detection analysis method is described in Sect. 3. The results are presented and discussed in Sect. 4, followed by some concluding remarks in Sect. 5.

2 Data sets and preparation procedure

This study analyzes two climate elements, a measure of atmospheric storminess and ocean wave heights. The atmospheric and ocean wave height data sets that we use are described in the following two subsections.

2.1 Atmospheric data and preparation

Many indices have been used to represent atmospheric storminess, including (but not limited to) cyclone count statistics (e.g., Pettersen 1956; Whitaker and Horn 1984), eddy variance/covariance statistics (e.g., Blackmon 1976; Hoskins and Hodges 2002), and geostrophic wind speeds (e.g., Matulla et al. 2008; Alexandersson et al. 1998, 2000). In this study, we also use an index that represents the geostrophic wind energy, namely the squared seasonal mean SLP gradient, to provide a measure of atmospheric storminess. We analyze time series of anomalies of the squared SLP gradient, G_t (here t denotes years) expressed relative to the 1961–1990 mean, to determine whether or not external influence on change in atmospheric storminess is detectable. The technical details of how the atmospheric storminess index G_t is derived are described in Appendix A.

For comparison with previous detection studies on SLP (Gillett et al. 2003, 2005), we also analyze time series of seasonal mean SLP anomalies expressed relative to the 1961–90 mean, P_t, which represent variability in the mean SLP field.

We use observationally based proxies derived from ERA40 for the 1958–2001 period (Uppala et al. 2005) and also use HadSLP2 (Allan and Ansell 2006). While the latter contains gridded SLP observations for the period 1900–2004 (available online at http://hadobs.metoffice.com/gmslp/hadslp2/index.html), we focus most of our effort on the most recent 50 years (1955–2004), which have significantly better spatial coverage than the earlier part of the record.

In order to detect the influence of historical external forcing in the observations, climate model based estimates of the response to external forcing and of natural internal climate variability are required. Thus, we also use simulations of SLP from the nine coupled ocean-atmosphere models listed in Tables 1 and 2. These simulations were obtained from the multi-model data archive at PCMDI (https://esg.llnl.gov:8443/index.jsp). An ensemble of simulations of the twentieth-century with historical external forcing is available for each of these models, with the individual ensemble sizes ranging from 3 to 7. In total, we use 41 such simulations (Table 1). Note that the twentieth-century simulations generally finished in 1999 or 2000. As in Gillett et al. (2005), we have extended these simulations to 2004 using output for 2000–2004 from integrations using the so-called SRES A1B, A2 or B1 emission scenario (Nakicenovic and Swart 2000), so that the model simulations can be compared to observations up to 2004 (see Table 1). The SRES emission scenarios are used extensively in the IPCC (2001, 2007) reports. There should not be significant differences among different scenarios for the period 2000–2004 (see, for example, Fig. 10.4 in Meehl et al. 2007). Control integrations from each of these models, which are listed in Table 2, are also used in this study.

The ERA40 SLP data are available on a 2.5°-by-2.5° lat-long grid, while the HadSLP2 data are only available on a 5°-by-5° lat-long grid and the multi-model SLP data were archived on different grids for different models. All of these data were interpolated onto the same 5°-by-5° lat-long grid, with the first gridpoint centered at (2.5°E, 87.5°S), and the last, at (357.5°E, 87.5°N).

Our detection analysis will focus on G_t and P_t change in each of the four seasons, separately: winter (JFM), spring (AMJ), summer (JAS), and fall (OND). We choose this definition of the seasons for convenience and because the resulting JFM data series is one year longer than the DJF data series. It is shown later in Sect. 4 that this definition does not notably affect the results of our detection analysis.

For each gridpoint we analyze, we obtain observed G_t and P_t time series from the ERA40 and HadSLP2 datasets, and also obtain corresponding time series from each of the climate model simulations. Linear trends are estimated from each time series for each season, over the periods 1958–2001 and 1955–2004, using the least squares method

Table 1 The nine coupled ocean–atmosphere models used in this study and the number of twentieth century runs conducted with each of these models (the period of simulation given in parentheses)

Model	Number of runs	Forcing	Scenario simulations used to extend the data to 2004		
			A2	A1B	B1
ECHO-G	Run 1–3 (1900–2000)	GSN	Run 1–3 for 2001–2004		
ECHO-G	Run 4–5 (1900–2000)	GSN		Run 1–2 for 2001–2004	
GFDL-CM2.0	Run 1 (1900–2000)	GSNO	Run 1 for 2001–2004		
GFDL-CM2.0	Run2 (1900–2000)	GSNO		Rrun 1 for 2001–2004	
GFDL-CM2.0	Run 3 (1900–2000)	GSNO			Run 1 for 2001–2004
GFDL-CM2.1	Run 1 (1900–2000)	GSNO	Run 1 for 2001–2004		
GFDL-CM2.1	Run 2 (1900–2000)	GSNO		Run 1 for 2001–2004	
GFDL-CM2.1	Run 3 (1900–2000)	GSNO			Run 1 for 2001–2004
GISS-EH	Run 1–3 (1900–1999)	GSNO		Run 1–3 for 2000–2004	
GISS-ER	Run 1 (1900–2003)	GSNO	Run 1 for 2004		
GISS-ER	Run 2–6 (1900–2003)	GSNO		Run 1–5 for 2004	
GISS-ER	Run 7 (1900–2003)	GSNO			Run 1 for 2004
MIROC3.2(medres)	Run 1–3 (1900–2000)	GSNO	Run 1–3 for 2001–2004		
MRI CGCM2.3.2	Run 1–5 (1900–2000)	GSN	Run 1–5 for 2001–2004		
NCAR CCSM3	Run 1–5 (1900–1999)	GSNO	Run 1–5 for 2000–2004		
NCAR CCSM3	Run 6, 7, 9 (1900–1999)	GSNO		Run 1–3 for 2000–2004	
NCAR PCM	Run 1–4 (1900–1999)	GSNO	Run 1–4 for 2000–2004		

The letters G, S, N and O denote Greenhouse gases, Sulphate aerosols, Natural and stratospheric Ozone forcing, respectively

Table 2 The control integrations (runs) used in this study

Model	Number of years available (used)
ECHO-G	340 (1–300 in 6 segments of 50-year or 1–308 in 7 segments of 44-year)
GFDL-CM2.0	500 (1–500 in 10 segments of 50-year or 1–484 in 11 segments of 44-year)
GFDL-CM2.1	500 (1–500 in 10 segments of 50-year or 1–484 in 11 segments of 44-year)
GISS-EH	400 (1–400 in 8 segments of 50-year or 1–396 in 9 segments of 44-year)
GISS-ER	500 (1–500 in 10 segments of 50-year or 1–484 in 11 segments of 44-year)
MIROC3.2(medres)	500 (1–500 in 10 segments of 50-year or 1–484 in 11 segments of 44-year)
MRI-CGCM2.3.2	350 (1–350 in 7 segments of 50-year or 1–308 in 7 segments of 44-year)
NCAR CCSM3	230 (1–200 in 4 segments of 50-year or 1–220 in 5 segments of 44-year)
NCAR CCSM3	500 (1–500 in 10 segments of 50-year or 1–484 in 11 segments of 44-year)
NCAR PCM	589 (1–550 in 11 segments of 50-year or 1–572 in 13 segments of 44-year)
Total	86 non-overlapping 50-year segments or 96 non-overlapping 44-year segments

(this method is used throughout this paper). These two overlapping periods are chosen because the ERA40 covers only the period 1958–2001 while HadSLP2 and the climate model simulations cover the longer period 1955–2004. Our detection analyses, which will compare the observed and simulated trend patterns, will be based on 1958–2001 when using ERA40, and 1955–2004 when using HadSLP2.

The linear trend patterns in G_t over the common period 1958–2001 are shown in Fig. 1. The corresponding 1955–2004 trends (not shown) are similar, though generally somewhat smaller. The multi-model simulated trend patterns in G_t (and P_t, not shown) resemble their observed counterparts in the boreal cold seasons, although the magnitude of the observed trend in the North Atlantic storm track region is substantially under-estimated (see Fig. 1c, e or d, f). The observed and simulated trend patterns are also similar in the southern hemisphere in the austral cold seasons (not shown).

Importantly, the storminess trends inferred from the geostrophic wind energy index G_t are consistent with the findings of previous studies using other types of storminess indices such as indices that are based on applying objective cyclone detection and tracking algorithms to 6-hourly SLP fields (Wang et al. 2006; Gulev et al. 2001). This

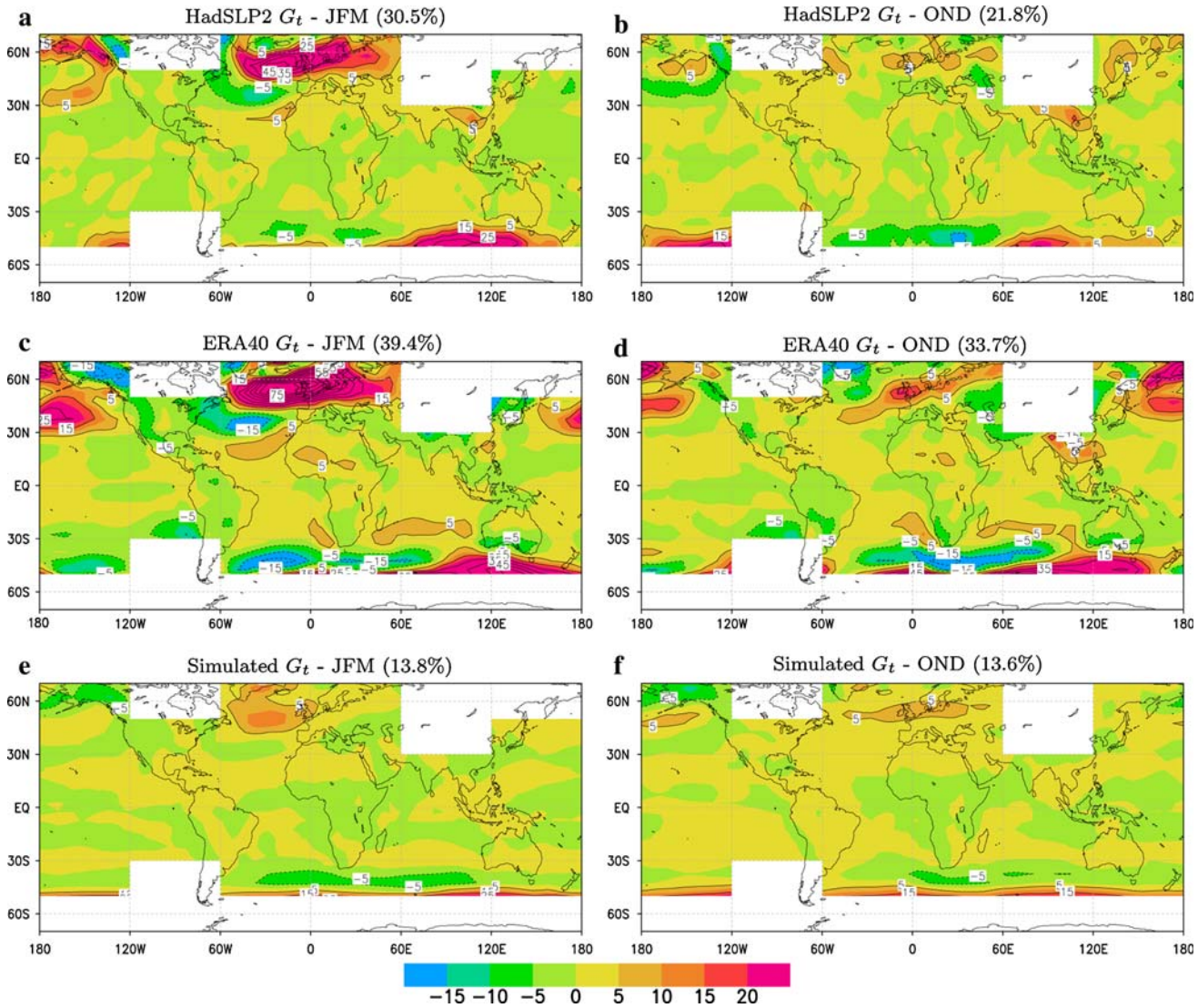

Fig. 1 The linear trend patterns of the geostrophic wind energy index G_t (unit: $(\text{hPa})^2/5°$) as derived from ERA40, HadSLP2, and the multi-model/ensemble mean simulations of G_t for the period 1958–2001. *Solid contours* and *yellow-red shadings* indicate upward/positive trends, and dashed ones and *green-blue shadings* downward/negative trends (unit: $(\text{hPa})^2/5°$ per century). The contourline interval is 10 units (note that *contour lines* are drawn every other shading level). The percentage of *grid-boxes* that were found to have a significant trend is given in the *parentheses* above

consistency suggests that the index G_t represents the trend component of the atmospheric storminess adequately. It should therefore be suitable for a detection analysis on change in atmospheric storminess.

2.2 Ocean wave height data and preparation

It is also of great interest to determine whether or not external influence on ocean wave heights is detectable, which is another goal of this study. Here, we focus on significant wave height (SWH). Thus, we also need both SWH observations and simulations of SWH response to historical external forcing.

There exist two global wave reanalyses: ERA40 (Uppala et al. 2005; Caires et al. 2004a) and AES40 (Cox and Swail 2001). Having inter-compared and validated these reanalyses against independent NOAA/NDBC buoy and TOPEX/Poseidon altimeter observations, Caires et al. (2004b) concluded that most of the large-scale features of observed wave height variability are equally present in these wave datasets. Thus, in this study, we use the original ERA40 SWH data as a proxy for actual SWH observations. These data were aggregated onto the same 5°-by-5° lat-long grid as is used to represent G_t and P_t (see Sect. 2.1).

Since SWH data are not directly available from the output of global climate models, estimates of the SWH

response to historical external forcing must be obtained empirically. We therefore used a statistical model to represent the observed relationship between atmospheric circulation and SWH, and applied this statistical model to climate model output in order to estimate variation and change in SWH that is induced by natural and anthropogenic forcing combined. Since the SLP-SWH relationships are weaker in the boreal warm seasons (Wang and Swail 2006b), we will focus only on the boreal winter and fall in this study. We will analyze both seasonal means and extremes of SWH in these seasons. The statistical models used for estimating seasonal mean and extreme SWH variates are described in following two subsections.

2.2.1 Seasonal mean SWH

Following Wang and Swail (2006a, b), we use the regression equation

$$H_{\mathrm{avg},\,t} = a + bG_t + cP_t + \varepsilon_t \quad (1)$$

to represent the observed relationship between seasonal mean SWH ($H_{\mathrm{avg},\,t}$) and the atmospheric circulation variates (G_t and P_t). Here, ε_t denotes a white noise process. The parameters a, b, and c are estimated using the seasonal quantities ($H_{\mathrm{avg},\,t}$, G_t, and P_t) that are derived from the ERA40 wave and SLP data for the period 1958–2001 (excluding a period of erroneous wave data from January 1992 to May 1993; Caires et al. 2004b). These estimates are denoted as $\hat{a}$, $\hat{b}$, and $\hat{c}$, respectively ($\hat{x}$ denotes an estimate of x throughout this paper). Both the predictor and predictand series were detrended before estimating these parameters in order to focus the regression on the relationship between non-systematic changes. Note that the predictor trend components were not removed when the fitted regression was subsequently used to obtain the predictand values.

This statistical model exploits the significant SLP-SWH relation that exists in most of the world oceans north of 30°N (Wang and Swail 2006b). We perform our detection analysis only in the areas where the SLP-SWH relationship is significant, which include most of the oceans in the 30°N–70°N band, as shown in Fig. 2. In these selected areas, the relationship $\hat{H}_{\mathrm{avg},\,t} = \hat{a} + \hat{b}G_t + \hat{c}P_t$ reasonably well reproduces the observed trend patterns in $H_{\mathrm{avg},t}$, especially in winter, although it significantly under-represents the magnitude of change (see Fig. 2a, c or b, d). Thus, we think that it is reasonable to use this SLP-SWH relationship to estimate changes in $H_{\mathrm{avg},t}$ for use in a detection analysis. Detection studies rely on the assumption that climate models, or in this case the combination of climate models and statistical downscaling simulate the correct pattern of response to external forcing, but they do not require that these patterns have the correct amplitude (e.g., Hegerl et al. 2007).

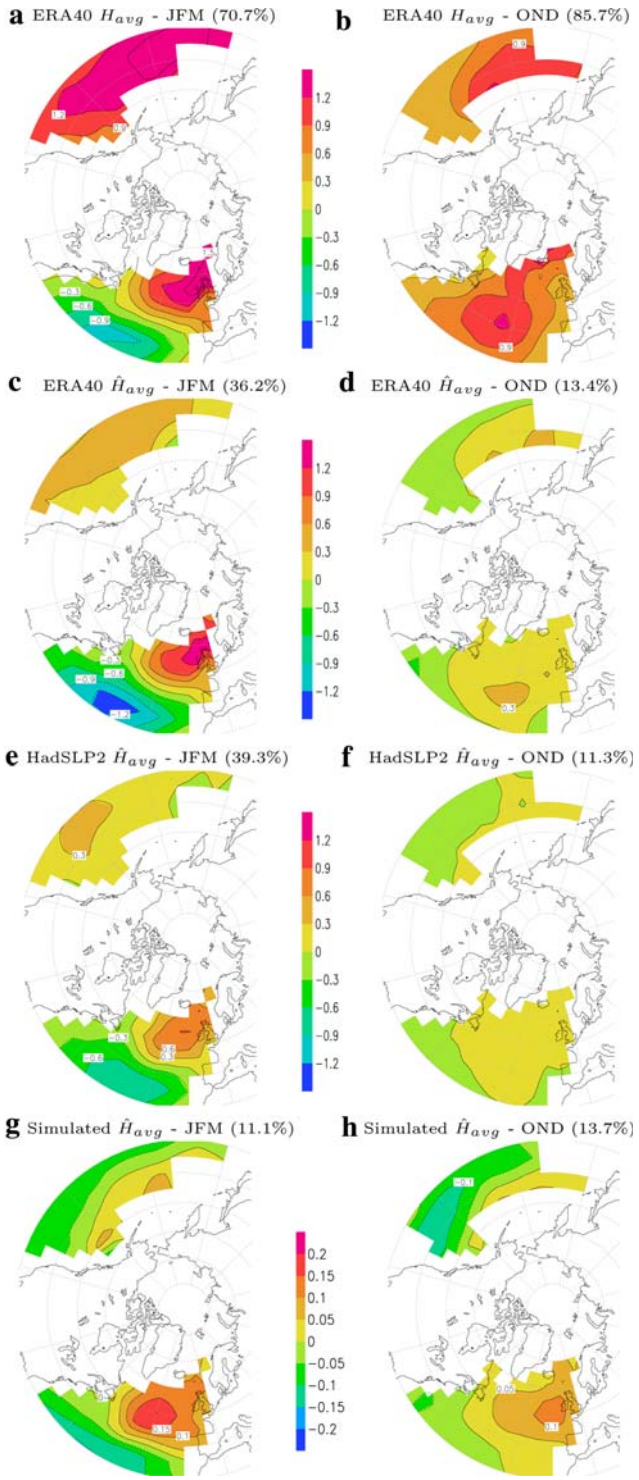

Fig. 2 The 1958–2001 linear trend patterns of the seasonal mean SWH as derived from the original ERA40 seasonal mean SWH (H_{avg}), the ERA40 and HadSLP2 hindcasts and the multi-model/ensemble mean simulations of H_{avg} (i.e., $\hat{H}_{\mathrm{avg}}$). *Yellow-red shadings* and *solid contours* indicate upward/positive trends, and *green-blue shadings* and *dashed contours* downward/negative trends (unit: cm/year). Note that the contour scale for the simulated trends (**g**, **h**) is different from the one for the observed trends, as indicated by the *contour scale bars*. The percentage of grid-boxes that were found to have a significant trend is given in the *parentheses* above

Having fit the statistical model, we feed the climate model simulated G_t and P_t into this relationship to estimate climate model simulated SWH variation $H_{\mathrm{avg},t}$ for each season at each wave gridpoint in the selected areas for each of the forced and control simulations used in this study. These estimates are also referred to as the simulated $H_{\mathrm{avg},t}$. For comparison with the original ERA40 $H_{\mathrm{avg},\,t}$ values, we also feed the ERA40 G_t and P_t values into this relationship, obtaining statistical hindcasts $\hat{H}_{\mathrm{avg},\,t}$, which are referred to as the ERA40 $H_{\mathrm{avg},t}$ hindcasts hereafter. Further, in order to have proxy SWH "observations" that are somewhat independent of the model used for ERA40, we also feed the HadSLP2 G_t and P_t values into this relationship, obtaining the HadSLP2 $H_{\mathrm{avg},\,t}$ hindcasts for the 1955–2004 period.

For our detection analysis, we estimate the linear trends in each of the $H_{\mathrm{avg},\,t}$ or constructed $\hat{H}_{\mathrm{avg},\,t}$ time series, again over the periods 1958–2001 and 1955–2004. The "observed" and multi-model simulated trend patterns for 1958–2001 are shown in Fig. 2. The corresponding 1955–2004 trend patterns (not shown) are similar to those shown in Fig. 2e–h, but with slightly smaller trend.

In general, the trend patterns in hindcast H_{avg} are reasonably well represented in the empirically downscaled climate model output, although the magnitude of trend is significantly under-estimated (see Fig. 2c–h). The trend pattern in the original ERA40 H_{avg} in the North Atlantic is also well simulated, especially in winter (see Fig. 2a, g). However, the climate models significantly under-estimate the magnitude and areal extent of the observed H_{avg} increase in the North Pacific (see Fig. 2a, b, g–h). Such under-estimation is not unexpected given that the multi-model simulated changes in the P_t and G_t fields are also smaller than observed. In addition, some variance is presumably lost in the process of converting the atmospheric circulation change to a change in H_{avg} using the regression relationship.

2.2.2 *Seasonal extreme SWH*

Seasonal maxima of SWH can also be derived from the original ERA40 6-hourly wave data, but as with seasonal mean SWH, comparable information is not directly available from climate models. Thus in this case also, we proceed to estimate extreme SWH quantities by means of a statistical downscaling approach following Wang and Swail (2006a, b). This approach uses the non-stationary Generalized Extreme Value model, GEV(μ_t, σ, ξ), where the location parameter μ_t depends on atmospheric variates G_t and P_t via the relationship

$$\mu_t = \mu_o + \gamma_1 G_t + \gamma_2 P_t + \epsilon_t, \qquad (2)$$

while the scale parameter σ and shape parameter ξ are constant over time. Here ϵ_t denotes a white noise process.

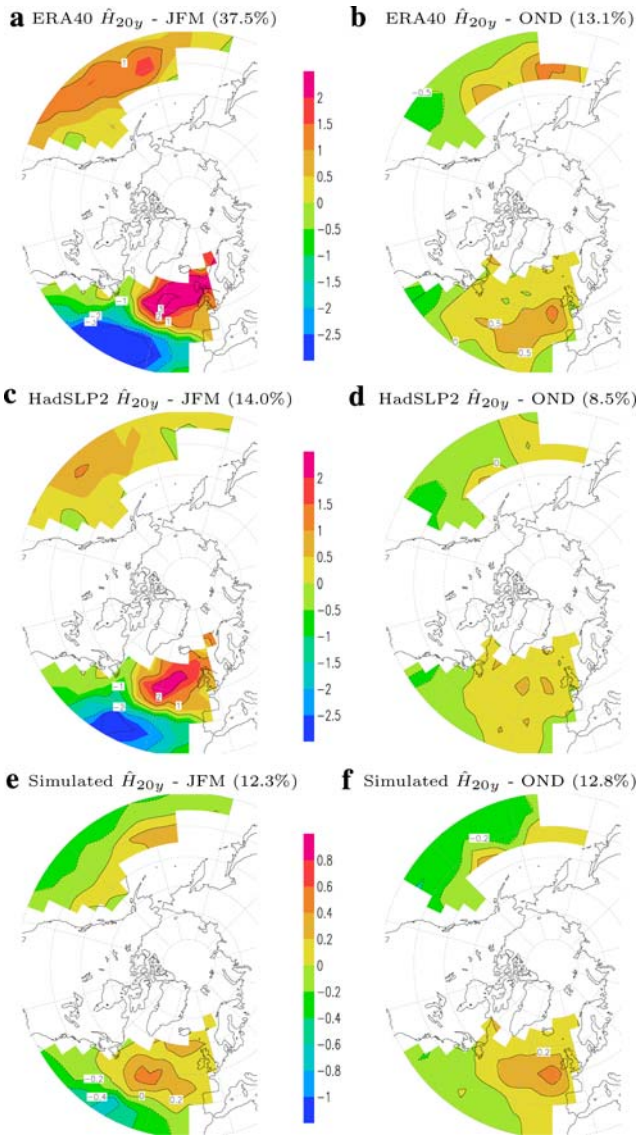

Fig. 3 The same as in Fig. 2 but for the linear trend (in cm/year) patterns of seasonal 20-year return values of SWH. Note that the contour scale for the simulated trends (**e**, **f**) is different from the one for the observed trends, as indicated by the *contour scale bars*

The time varying GEV distribution can be used in two different ways. First, by specifying a fixed size of extreme event, it is possible to estimate how the risk of that event varies as G_t and P_t vary. Alternatively, a fixed risk could be specified, such as a 1-in-20 year risk of occurrence, and the magnitude of threshold that is exceeded with that fixed risk can be estimated by finding the appropriate quantile of the GEV distribution. Since the location parameter in our statistical model depends upon G_t and P_t, that threshold then becomes a function of G_t and P_t as well. We take the latter approach in this study.

The seasonal maximal SWH and seasonal atmospheric quantities G_t and P_t derived from the ERA40 wave and

SLP data are used to estimate the parameters μ_o, γ_1, γ_2, σ, and ξ via the method of maximum likelihood. The above GEV model is fitted for each wave gridpoint analyzed and for each season, separately. The fitted model, $\mathrm{GEV}(\hat{\mu}_t = \hat{\mu}_o + \hat{\gamma}_1 G_t + \hat{\gamma}_2 P_t, \hat{\sigma}, \hat{\xi})$, is then used to estimate a series of 20-year return values of SWH, $H_{20y,\,t}$, that vary with the atmospheric circulation parameters G_t and P_t. More specifically, we use the ERA40 G_t and P_t values in the fitted model to obtain ERA40 hindcasts of $H_{20y,\,t}$, and use the HadSLP2 G_t and P_t values to obtain HadSLP2 hindcasts of $H_{20y,\,t}$. In order to estimate the $H_{20y,\,t}$ response to historical external forcing, we use climate model simulated G_t and P_t values in the fitted GEV model, obtaining the corresponding downscaled $H_{20y,\,t}$ values. This is done for each of the 41 twentieth-century simulations (Table 1), and also for each of the control simulations (Table 2). Linear trend patterns are subsequently estimated from all of the constructed return value time series and used in our detection analysis.

In general, the trend patterns in the seasonal 20-year return values of SWH (H_{20y}) in both hindcasts are reasonably well reproduced by the climate models combined with the statistical downscaling, especially in the North Atlantic (compare Fig. 3a–d with e, f). The trend patterns in H_{20y} are also similar to those of the corresponding H_{avg} trend patterns (see Figs. 2 and 3), which is not surprising given that both are functions of linear combinations of G_t and P_t.

3 Detection analysis

Climate variability refers to variations in the mean state and other statistics (such as the standard deviation, the intensity of extremes, etc.) of the climate on all spatial and temporal scales beyond that of individual weather events (Solomon et al. 2007). Variability is referred to as internal variability if it arises from natural internal processes within the climate system, and as external variability if it is due to external forcing (anthropogenic, or natural such as changes in solar radiation and volcanism). Climate change may be due to internal climate system processes and/or external forcings on the climate system. The objective of climate change detection analysis is to understand how climate changes that result from anthropogenic and natural external forcings may be distinguished from changes and variability that result from internal climate system processes. The spatial and temporal scales used to analyze climate change are carefully chosen so as to focus on the spatio-temporal scale of the response, filter out as much internal variability as possible and enable the separation of the responses to different forcings (Hegerl et al. 2007). In this study, we compare the pattern of observed trends with a multi-model estimate of the trend that is expected to arise from external forcing to determine whether or not external influence is detectable in the pattern of observed trends. We do this by means of the optimal detection approach (Allen and Stott 2003; Hegerl et al. 1997; Hasselmann 1993).

Let Y_o denote an estimate of the observed linear trend pattern, and let Y_m be the multi-model mean linear trend pattern (i.e., the average of the linear trends in the individual simulations that make up the combined ensemble of 41 simulations from the 9 climate models used in this study; see Table 1). The observed trend Y_o consists of a response to historical external forcing and internal variability η with variance-covariance matrix C_η. Similarly, Y_m consists of the simulated response to the historical external forcing and some internal variability ζ with variance-covariance matrix C_ζ. Note that C_ζ is small because Y_m is a multi-model mean trend pattern and thus much of the internal variability that affects trends in individual simulations has been averaged out. The objective here is to compare the response to forcing in the observations ($Y_o - \eta$) with the simulated response ($Y_m - \zeta$). This is accomplished by means of the optimal detection approach, i.e., by fitting with optimization the simulated pattern of trends to the observed pattern of trends as follows:

$$Y_o = \beta(Y_m - \zeta) + \eta. \tag{3}$$

Here, optimization means to maximize signal-to-noise ratio by rotating the coordinate space (Hasselmann 1993). Because the data required to estimate C_η and C_ζ are limited, and because models do not simulate internal variability well on all scales, the analysis is typically performed in a dimension reduced space spanned by some number of leading EOFs (see Table 3; in this study we consider EOF truncations between 2 and 10). The regression coefficient β is estimated by means of the total least squares (TLS) algorithm (Allen and Stott 2003). [Note that the ordinary least squares algorithm is not quite suitable here because it assumes no error term in the explanatory variable, while both the response variable Y_o and the explanatory variable Y_m in model (3) contain a noise term, η or ζ.]

Optimal detection analysis requires knowledge of the internal climate variability. Two independent estimates of the internal variability variance-covariance matrix C_η are required in this study: $\hat{C}_{\eta_1}$ and $\hat{C}_{\eta_2}$. These estimates are obtained by pooling control simulation variability with the inter-integration variability found in ensembles of twentieth-century simulations. Details are given in Appendix B. $\hat{C}_{\eta_1}$ is used in the optimization, while $\hat{C}_{\eta_2}$ is used in estimating the scaling factor β and to obtain a confidence interval for $\hat{\beta}$. Once the regression model has been fitted, the residuals $\hat{\eta} = Y_o - \hat{Y}_o = Y_o - \hat{\beta}(Y_m - \hat{\zeta})$ can be calculated , so that their variance can be compared with the

Table 3 The estimated scaling factor β, its 5–95% uncertainty range (β_l, β_u), and the result of the related residual consistency test for each of the listed detection variables as derived from the ERA40 or HadSLP2 dataset

Season	Observations	Variable	Domain	$\hat{\beta}$	$\hat{\beta}_l$	$\hat{\beta}_u$	#EOFs	Internal variability	Detected?
			GL (31)	4.22	0.62	11.68	6 (2–6)	OK	Yes
		G_t	NH (14)	7.05	2.16	27.45	3 (2–5)	OK	Yes, $\beta > 1$
			NA (102)	4.58	1.51	10.19	6 (2–10)	OK	Yes, $\beta > 1$
			GL (31)	0.65	−0.82	2.20	6 (4)	OK	No
	HadSLP2	P_t	NH (14)	3.31	−0.55	10.93	3 (5–6)	Under-estimated	No
			NA (102)	5.07	1.83	11.46	6 (2–9)	OK	Yes, $\beta > 1$
			NH (175)	4.24	0.63	11.77	7 (2–8)	Under-estimated	Yes
		$\hat{H}_{\mathrm{avg}}$	NA (102)	5.03	1.47	12.79	5 (2–10)	OK	Yes, $\beta > 1$
			NH (175)	3.45	0.15	9.24	7 (4–10)	Under-estimated	Yes
		$\hat{H}_{20y}$	NA (102)	4.86	1.58	11.33	5 (2–10)	OK	Yes, $\beta > 1$
			GL (31)	12.38	−99.64	3.94	6 (2, 4)	Under-estimated	No
JFM		G_t	NH (14)	12.00	−280.7	4.08	3 (2, 4)	Under-estimated	No
			NA (102)	10.58	3.89	100.1	6 (2, 6–10)	Under-estimated	Yes, $\beta > 1$
			GL (31)	3.27	1.47	5.60	6 (2–10)	Under-estimated	Yes, $\beta > 1$
		P_t	NH (14)	9.83	3.53	72.55	3 (2–3, 5–10)	Under-estimated	Yes, $\beta > 1$
			NA (102)	6.99	3.16	16.48	6 (2–10)	OK	Yes, $\beta > 1$
	ERA40		NH (175)	10.86	4.57	55.22	7 (2, 6–10)	Under-estimated	Yes, $\beta > 1$
		$\hat{H}_{\mathrm{avg}}$	NA (102)	6.57	2.77	15.96	5 (2–10)	Under-estimated	Yes, $\beta > 1$
			NH (175)	8.21	3.62	23.25	7 (2–10)	Under-estimated	Yes, $\beta > 1$
		$\hat{H}_{20y}$	NA (102)	7.09	3.11	17.52	5 (2–10)	Under-estimated	Yes, $\beta > 1$
			NH (175)	22.13	11.02	172.77	7 (6–7)	Under-estimated	Yes, $\beta > 1$
		H_{avg}	NA (102)	8.02	4.16	17.34	5 (2–10)	Under-estimated	Yes, $\beta > 1$
			GL (32)	0.26	−2.16	2.75	6	OK	No
		G_t	NH (15)	0.23	−3.77	4.43	3	OK	No
			NA (114)	0.49	−2.63	3.87	6	Over-estimated	No
			GL (32)	0.69	−0.66	2.10	6	Over-estimated	No
	HadSLP2	P_t	NH (15)	0.36	−1.77	2.28	3	Over-estimated	No
			NA (114)	0.87	−2.33	4.58	6	Over-estimated	No
			NH (179)	1.21	−0.98	3.73	7	OK	No
		$\hat{H}_{\mathrm{avg}}$	NA (114)	0.68	−2.33	4.01	5	Over-estimated	No
			NH (179)	1.05	−1.13	3.50	7	OK	No
		$\hat{H}_{20y}$	NA (114)	0.48	−2.74	3.97	5	Over-estimated	No
			GL (32)	0.62	−2.27	3.79	6	OK	No
OND		G_t	NH (15)	1.26	−6.96	17.24	3 (7–8)	OK	No
			NA (114)	2.04	−3.67	13.82	6	OK	No
			GL (32)	1.51	−0.17	3.40	6 (7–10)	Under-estimated	No
		P_t	NH (15)	1.31	−1.43	4.59	3	OK	No
			NA (114)	2.39	−2.52	12.27	6	OK	No
	ERA40		NH (179)	2.70	0.22	6.13	7 (6–9)	OK	Yes
		$\hat{H}_{\mathrm{avg}}$	NA (114)	3.37	−1.26	14.82	5	OK	No
			NH (179)	2.70	0.16	6.35	7 (7–9)	OK	Yes
		$\hat{H}_{20y}$	NA (114)	3.51	−0.53	12.18	5 (10)	OK	No
			NH (179)	33.43	16.28	1704.9	7 (6–10)	Under-estimated	Yes, $\beta > 1$
		H_{avg}	NA (114)	31.29	15.25	1034.7	5 (2–3, 5–9)	Under-estimated	Yes, $\beta > 1$
			GL (32)	−2.69	−4.64	−1.09	6	Under-estimated	No
	HadSLP2	G_t	NH (15)	−0.48	−2.10	1.09	3	Over-estimated	No
			SH (11)	−28.78	−8.20	27.96	3	Under-estimated	No
			GL (32)	0.53	−0.54	1.63	6	Under-estimated	No

Table 3 continued

Season	Observations	Variable	Domain	$\hat{\beta}$	$\hat{\beta}_l$	$\hat{\beta}_u$	#EOFs	Internal variability	Detected?
AMJ		P_t	NH (15)	−0.42	−1.66	0.79	3	OK	No
			SH (11)	1.56	0.23	3.04	3 (2–6)	Under-estimated	Yes
			GL (32)	6.52	1.83	24.08	6 (5–8)	Under-estimated	Yes, $\beta > 1$
	ERA40	G_t	NH (15)	0.47	−1.87	2.95	3	Under-estimated	No
			SH (11)	17.60	−44.92	5.75	3 (6–10)	Under-estimated	No
			GL (32)	1.35	0.08	2.73	6 (4, 6–7, 9)	Under-estimated	Yes
		P_t	NH (15)	−1.68	−2.84	−0.61	3	Under-estimated	No
			SH (11)	3.37	1.69	5.50	3 (2–6, 8–9)	Under-estimated	Yes
			GL (33)	3.57	1.71	6.05	6 (2–10)	Under-estimated	Yes, $\beta > 1$
	HadSLP2	G_t	NH (16)	−1.61	−17.20	5.62	3	OK	No
			SH (11)	5.95	3.34	10.38	3 (2–10)	Under-estimated	Yes, $\beta > 1$
			GL (33)	−0.69	−2.29	0.83	6	OK	No
JAS		P_t	NH (16)	−0.89	−4.13	1.97	3	Over-estimated	No
			SH (11)	−0.77	−2.62	0.95	3	OK	No
			GL (33)	7.82	4.19	16.00	6 (3, 5–10)	Under-estimated	Yes, $\beta > 1$
	ERA40	G_t	NH (16)	−2.03	−23.41	5.28	3	Over-estimated	No
			SH (11)	6.24	2.31	16.44	3 (2–3, 5–6)	Under-estimated	Yes, $\beta > 1$
			GL (33)	6.44	1.96	21.34	6 (6)	Under-estimated	Yes, $\beta > 1$
		P_t	NH (16)	0.47	−2.87	4.07	3	OK	No
			SH (11)	−3.15	−12.24	1.12	3 (9–10)	Under-estimated	No

The numbers in the Domain column are the number of gridboxes retained in the domain for the analysis. The "#EOFs" column shows the number of retained leading EOFs corresponding to these estimates; and the numbers in parentheses, if given, mean that a detection would be possible (i.e., $\hat{\beta}_l > 0$) if any number of leading EOFs in this range were retained

corresponding simulated internal variability using a standard residual consistency test (Allen and Stott 2003). Inconsistency could occur either because (1) the climate models estimated the forced signal correctly, but under- or over-estimated internal variability; (2) the climate models did not respond correctly to forcing, but did simulate internal variability correctly; or (3) a combination of (1) and (2).

The main results of the above optimal detection analysis can be interpreted as follows: The external influence is detected if the scaling factor β in (3) is estimated to be significantly greater than zero, but not detected if $\hat{\beta}$ is not significantly different from zero or negative. An estimate $\hat{\beta}$ that is significantly greater than zero and consistent with unity indicates that the observed and simulated responses are considered to be comparable with each other, in which case it may be possible to attribute the observed trends to the historical external forcing if other plausible explanations (causes) can be ruled out.

This optimal detection approach was applied to determine whether or not external influence on the observed trends in the following variables is detectable: (1) the geostrophic wind energy index G_t, (2) the mean pressure field P_t, (3) seasonal mean SWH H_{avg}, and (4) the 20-year return value of seasonal extreme SWH H_{20y}. For each of the four variables, at least two sets of observationally based datasets were available, derived from either ERA40 or HadSLP2, as described earlier in Sect. 2. The detection analysis is carried out for each of these sets of proxy observations in the following detection domains: (a) global (GL: 60°S–80°N), (b) northern hemisphere (NH: 0°N–80°N), (c) North Atlantic (NA: the ocean domain in 20°N–75°N) or (d) southern hemisphere (SH:0°S–60°S).

For GL, NH and SH, the analysis is performed on a 20°-by-60° lat-long grid, using grid-box-area weighted averages of the values at the 48 5°-by-5° gridpoints in each 20°-by-60° grid-box [the longitudes of the 20°-by-60° grid-box-centers (grid-points) are 30°E, 90°E,..., 240°E, 300°E, while the latitudes are 70°S, 50°S,..., 50°N, 70°N]. For the North Atlantic (NA: 20°N–75°N), we carried out the detection analysis on the 5°-by-5° grid over the ocean only (gridpoints over land are excluded). For the analyses on SWH statistics, the detection domain only covers the selected ocean areas in 30°N–70°N (see Sect. 2.3 and Figs. 2, 3).

In order to exclude gridpoints with frequent missing observations from our detection analysis, we use the missing data information obtained from HadSLP2.0, the un-interpolated HadSLP2 product (Allan and Ansell 2006).

For any year, a missing value is assigned to all 20°-by-60° grid-boxes that have missing data at 50% or more of their 48 5°-by-5° gridpoints; and a linear trend is estimated only when there are at least 30 years of valid data for a 20°-by-60° gridbox during the period of 1955–2004. The excluded grid-boxes are not shaded in Fig. 1. Note that the same retained grid-boxes are analyzed regardless of whether HadSLP2 or ERA40 is analyzed. The same missing data mask was also applied to the model simulations, so that the observed and simulated quantities are analyzed on the same grid.

The results of our detection analysis are presented and discussed in the next section. Since $H_{\mathrm{avg},\ t}$ is a linear combination of G_t and P_t, and $H_{20y,\ t}$ is a monotonic function of a linear combination of G_t and P_t, detection results for $H_{\mathrm{avg},\ t}$ and $H_{20y,\ t}$ are not independent of detection results for G_t and P_t. Nevertheless, detection analyses on $H_{\mathrm{avg},\ t}$ and $H_{20y,\ t}$ are useful because they allow a synthesis of results on G_t and P_t into results concerning SWHs.

4 Results of detection analysis

First of all, in order to compare with Gillett et al. (2005), a detection study on the decadal mean DJF SLP fields over the globe (derived from HadSLP2.0), we also carried out our detection analysis using the global decadal means of JFM P_t as derived from HadSLP2. We obtained similar results, that is, a scaling factor of 1.21 that is significantly greater than zero but not significantly different from unity [the estimated 5–95% uncertainty range on β is (0.17, 2.33)], and that also passes the residual consistency check. Note that these results are not listed in Table 3, which contains only the results of detection analyses on trends. This corroborates the finding of Gillett et al. (2005) that boreal winter "sea level pressure trends may be attributed to external influence". The difference in the definition of the boreal winter season (i.e., DJF vs. JFM) between Gillett et al. (2005) and this study does not affect the detection conclusion; neither does the slight difference between using HadSLP2.0 and HadSLP2 (un-interpolated and interpolated versions).

We now turn to detection analysis on the linear trend patterns of geostrophic wind energy index G_t, of seasonal mean SLP anomalies P_t, and of seasonal mean and extreme SWH. The results, which are summarized in Table 3 and shown in Figs. 4, 5, 6, indicate that there exist detectable external influences on the observed trends of atmospheric storminess and ocean wave heights in boreal winter (JFM) in the past half century, especially for the North Atlantic region. These results are discussed in detail in the subsections below.

4.1 Detection results for boreal winter trends

As shown in Fig. 4, in boreal winter, the scaling factor β is estimated to be significantly greater than zero (i.e., inconsistent with zero) for the NA domain in each case, being consistent across the different variables (G_t, P_t, H_{avg}, and H_{20y}) and across the different observationally based data sets (ERA40 or HadSLP2, original or hindcast wave data; see also Table 3). The detection results are generally also consistent across a range of EOF truncations (see the "#EOFs" column in Table 3), although scaling factors, confidence intervals, and the results of residual consistency test are reported for fixed levels of EOF truncation that were chosen according to the domain and variable of interest. The chosen levels of EOF truncation (Table 3) explain 65–95% of the variance in $\hat{C}_{\eta_1}$.

For the NH domain, the detection results are consistent across both SWH variates for both observationally based data sets; but they are not consistent between G_t and P_t, nor between HadSLP2 and ERA40. For the NH and GL domains, external influence is detectable in the G_t trends only when HadSLP2 is used, and in the P_t trends only when ERA40 is used. As mentioned before, results of detection on SWH are a synthesis of the detection results on G_t and P_t. Thus, external influence on SWH could be detected if it is detected on G_t or P_t or both.

In general, the results of detection on the wave heights are physically consistent with those of detection on the geostrophic wind energy field and/or sea level pressure field. Increases in geostrophic wind energy and/or decreases in sea level pressure are associated with increases in ocean wave heights (both seasonal means and seasonal extremes). The detection results are also consistent with the relative magnitudes of the responses shown in the left panels of Figs. 1, 2, 3.

Nevertheless, the associated 5–95% uncertainty ranges on β do not include unity in most cases, especially when ERA40 SLP or waves are used as observations (see Table 3 and Fig. 4). In these cases, the scaling factor is estimated to be significantly greater than unity, which suggests that the climate models, or the combination of climate models and empirical downscaling models, significantly under-estimate the magnitude of the response of atmospheric storminess or ocean wave heights to the observed changes in external forcing (assuming the observed trends are not systematically over-estimated in the HadSLP2 or ERA40 data). Note that the estimated scaling factor is particularly large when the original ERA40 waves data are used (see the right most dashed bars in Fig. 4b). This is because the climate model simulated SWH is derived statistically using an observed SLP-SWH relationship that has previously been found to under-estimate the variability and trend magnitude of ocean wave heights (see Sect. 2.2). Under-

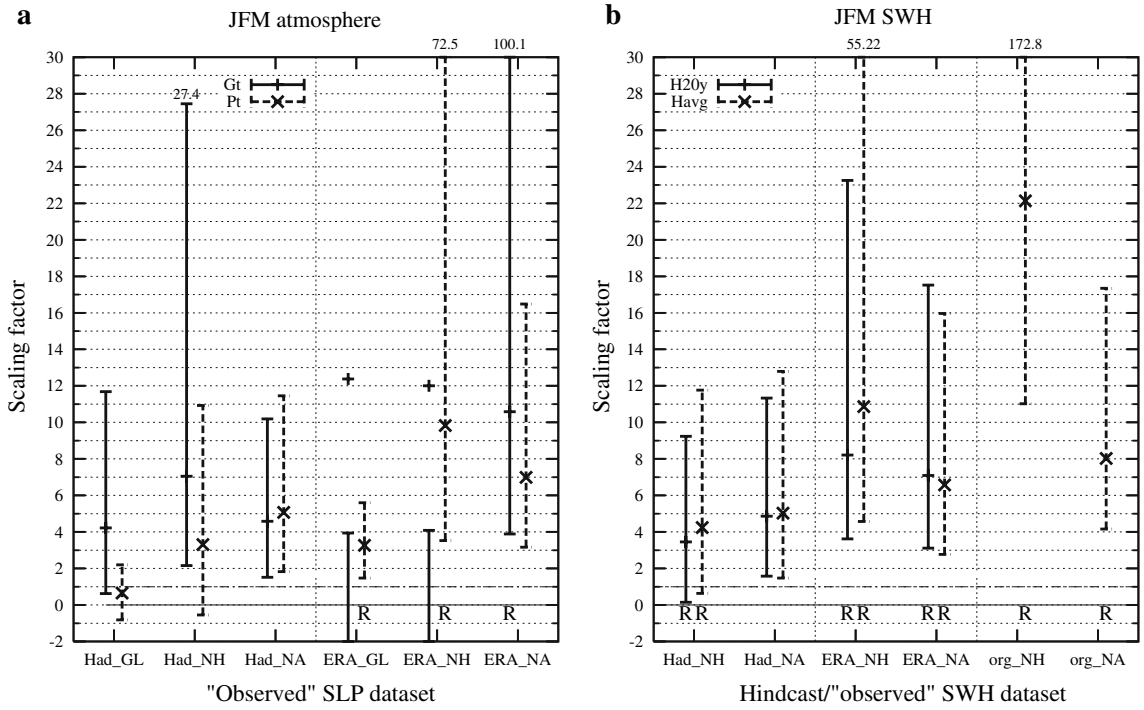

Fig. 4 The scaling factors (β) in the regressions of the winter (JFM) trend patterns of P_t (seasonal mean SLP anomalies), G_t (seasonal anomalies of squared SLP gradients), H_{avg} (seasonal mean SWH), and H_{20y} (seasonal 20-year return values of SWH), separately, on the relevant multi-model mean of simulated trend patterns. The period for calculating the trend is 1955–2004 when the HadSLP2 is used, and 1958–2001 when ERA40 SLP is used. In the dataset labels, "Had" stands for HadSLP2, and "ERA" for ERA40 SLP; the domain of detection is denoted by two letters: GL for global, NH for the Northern Hemisphere, and NA for the North Atlantic; and "org" is used to denote that the original ERA40 wave data (seasonal mean SWH) is used as the proxy of SWH observations. An "R" just under the zero-β line denotes that the climate models under-estimate the internal variability in the trend of the particular variable in the corresponding observed dataset

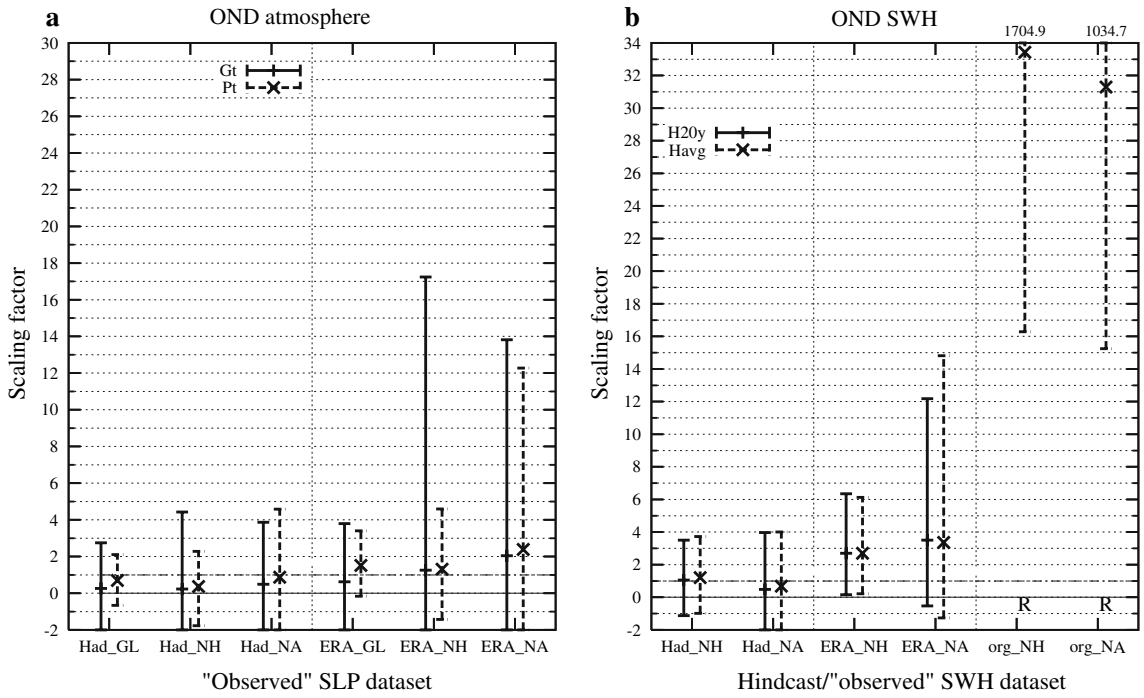

Fig. 5 The same as in Fig. 4 but for the boreal fall (OND) season

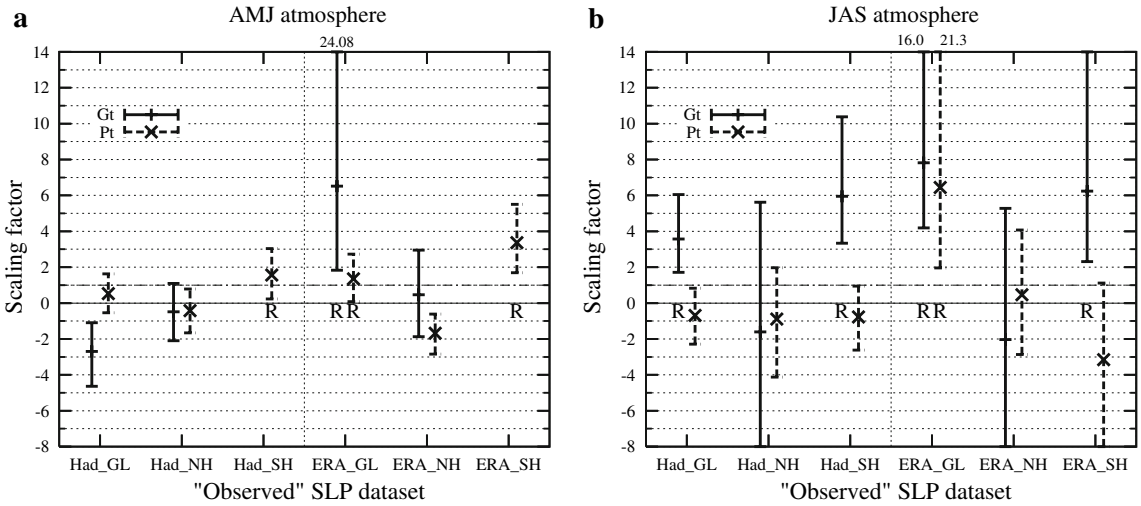

Fig. 6 As in Fig. 4a but for the austral fall (AMJ) and austral winter (JAS) trends. North Atlantic (NA) results are not shown, but Southern Hemisphere (SH) results are given

estimation of the observed trend magnitude by climate models would bias high the estimate of the scaling factor β. However, the effect of such under-estimation is reduced when statistical SWH hindcasts are used as observations, because both "observed" and simulated SWH quantities were derived from the same statistical SLP-SWH relationship in such cases. Thus, the scaling factor is somewhat closer to unity when the ERA40 hindcast SWH quantities are used as observations (see Fig. 4b). In general, the estimated β values are still greater than unity in these cases, which arises from under-estimation of the observed trend magnitude of seasonal mean SLP anomalies P_t and of geostrophic wind energy index G_t by the climate models (especially the latter; see Fig. 4a).

The residual variance was found to be consistent with the corresponding simulated internal variability for all cases of detection for the NA domain when HadSLP2 is used (see Table 3 and Fig. 4a). In contrast, the climate models, or the combinations of climate models and statistical downscaling, were found to under-estimate the observed internal variability in almost all cases of detected influence (except that of P_t in NA) when ERA40 is used (Table 3). Such under-estimation weakens the robustness of the detection results.

4.2 Detection results for the transition seasons

External influence is basically not detectable in the boreal fall season (OND). It was not detected for any of the four variates (G_t, P_t, H_{avg}, and H_{20y}) when HadSLP2 is used (see Table 3 and Fig. 5). When ERA40 is used, external influence was not detected in either G_t or P_t, no matter in which domain; but it was detected in both H_{avg} and H_{20y} for the NH domain (and in the original ERA40 H_{avg} trends over the NA domain only).

Note that the detection results in boreal fall season differ considerably between the original and hindcast ERA40 wave heights (see Table 3). This is because the observed SWH trend patterns are less well reproduced by the statistical SLP-SWH relationship in this season (see Fig. 2).

In the austral fall season (AMJ), external influence was detected in P_t for the SH domain, no matter whether HadSLP2 or ERA40 is used (see Table 3 and Fig. 6a). It was also detected in both G_t and P_t globally, but only when ERA40 is used. When HadSLP2 is used, external influence is not detectable in either G_t or P_t, globally or for the NH domain. Thus, in this season, external influence appears to be stronger in the SH than in the NH, and also stronger on P_t than on G_t. It is detectable only in P_t over the southern hemisphere.

4.3 Detection results for austral winter trends

In austral winter (JAS), as shown in Fig. 6b (see also Table 3), external influence on the observed G_t trends was detected globally and for the SH domain, no matter whether HadSLP2 or ERA40 is used; but it was not detected for the NH domain. External influence on the observed P_t trends is much weaker than on the observed G_t trends in this season. It appears to be detectable only when ERA40 is used (and only for the GL domain; see Table 3 and Fig. 6b).

In summary, in this season, external influence is mainly detectable on the geostrophic wind energy field, and is detectable mainly in the southern hemisphere. External influence is not detectable in the northern hemisphere in austral winter.

Table 4 As in Table 3 but for the results of detection analysis on the 1900–1949 trend patterns

Season	Observations	Variable	Domain	$\hat{\beta}$	$\hat{\beta}_l$	$\hat{\beta}_u$	#EOFs	Internal variability	Detected?
JFM	HadSLP2	G_t	NH (8)	34.09	−640.3	640.3	3	Under-estimated	No
		P_t	NH (8)	−9.58	−640.3	640.3	3	OK	No
OND	HadSLP2	G_t	NH (9)	−1.40	−14.71	5.73	3	OK	No
		P_t	NH (9)	0.39	−4.08	5.26	3	OK	No
AMJ	HadSLP2	G_t	NH (7)	−0.89	−640.3	640.3	3	OK	No
		P_t	NH (7)	0.61	−5.94	8.73	3	OK	No
JAS	HadSLP2	G_t	NH (9)	0.52	−640.3	640.3	3	OK	No
		P_t	NH (9)	2.20	−640.3	640.3	3	OK	No

5 Concluding remarks

In this study, work on the detection of external influence on atmospheric storminess and ocean wave heights is carried out, based on twentieth-century simulations from multiple climate models with combined natural and anthropogenic forcing and statistical downscaling of the corresponding changes in ocean wave heights for the periods 1955–2004 and 1958–2001. The observational data used in this study was obtained from ERA40 (SLP, geostrophic wind energy, and SWHs; Uppala et al. 2005; Caires et al. 2004a) and HadSLP2 (Allan and Ansell 2006).

It has been shown that the observed trend patterns in atmospheric storminess are reasonably well reproduced by the climate models, especially for the North Atlantic in boreal winter. Observed NH ocean wave height trends are also reasonably well reproduced when climate model output is statistically downscaled, although the magnitude of the trends is under-estimated.

In boreal winter, the observed 1955–2004 trend patterns in atmospheric storminess and ocean wave heights are characterized by an upward trend in the high-latitudes (especially the northeast North Atlantic) with a downward trend in the mid-latitudes, which were found to contain a detectable response to a given combination of natural and anthropogenic external forcing.

In general, the results of our detection analysis suggest that, in the past half century, the external forcing has had a detectable influence on trends in atmospheric circulation (including storminess) in the winter hemisphere (i.e., northern hemisphere in JFM and southern hemisphere in JAS), and on trends of NH ocean wave heights in boreal winter. The signal of external influence is weaker in the transition seasons, and is hardly detectable in the northern hemisphere in boreal summer. Climate models generally simulate smaller changes than observed and also appear to under-estimate the internal variability, reducing the robustness of our detection results.

Analyses of triangles of long surface pressure records (e.g., Matulla et al. 2008; Alexandersson et al. 1998, 2000), which are generally located in the coastal region of northern Europe, suggest that there have been earlier, long term variations in atmospheric storminess that are comparable to changes seen at these same locations during the past half century. Whether these earlier changes were associated with a similar large scale pattern of change as detected here remains an open question. It would therefore be of great interest to carry out the detection analysis on storminess trends over a longer period, such as the entire past century.

The HadSLP2 does have data for the entire past century, although the data coverage for the first half of the twentieth century is not as good as during the latter half of the century. However, the number of available climate model simulations is insufficient to confidently estimate the internal variability on the century time scale.

Alternatively, we carried out the optimal detection analysis on the observed 1900–49 G_t and P_t trend patterns (not shown), comparing the HadSLP2 G_t and P_t trend patterns with the corresponding multi-model mean trend patterns for this period. Data coverage, however, is limited during this period. In fact, only 9–11 60° × 20° gridboxes satisfy our data coverage criterion (see Sect. 3), with all but two of these gridboxes located in the NH (mostly between 10°N and 50°N, as well as one gridbox between 50°N and 70°N in the North Atlantic in JAS). Thus, the detection domain here is NH only. In boreal winter, the observed 1900–1949 G_t trend pattern is characterized by decreases in the mid-latitudes (north of 40°N) of the North Atlantic.

The detection results for our 1900–1949 analysis are summarized in Table 4. In contrast to the results for the latter half of the twentieth century (Table 3), external influence on the observed G_t and P_t trends over the early half of the twentieth century is not detectable, regardless of season and the number of leading EOFs retained. This maybe because limited data coverage and a weaker signal have conspired to reduce the signal-to-noise ratio in our detection analysis. However, it also suggests that external forcing is less likely to have been an important factor in surface pressure and atmospheric storminess change during the first half of the twentieth century.

Acknowledgments The authors are grateful to Dr. Jiafeng Wang for his help in compiling the climate model outputs, to Dr. Tara Ansell and Dr. Nathan P. Gillett for their help in answering our questions about the HadSLP2 data, and to Dr. Myles Allen for his help in originating the idea of a detection work on ocean wave heights. The authors also wish to thank Dr. Seung-Ki Min and Dr. Bin Yu for their helpful comments on an earlier version of this manuscript.

Appendix A Calculation of the atmospheric storminess index G_t

The atmospheric storminess index G_t is calculated as follows. First, the squared SLP gradient at gridpoint (i, j) is computed as

$$G_t^0(i,j)=\frac{\left[\frac{P_t^0(i,j)-P_t^0(i-1,j)}{\cos(\phi_j)}\right]^2+\left[\frac{P_t^0(i+1,j)-P_t^0(i,j)}{\cos(\phi_j)}\right]^2}{2} + \frac{\left[P_t^0(i,j)-P_t^0(i,j-1)\right]^2+\left[P_t^0(i,j+1)-P_t^0(i,j)\right]^2}{2} \quad (4)$$

where $P^0{}_t(i, j)$ denotes the seasonal mean SLP for gridpoint (i, j) in year t, ϕ_j is the latitude of the gridpoint (i, j) and where it is assumed that the grid spacing, in degrees, is equal in both latitude and longitude. Then, the 1961–1990 mean field, $\bar{G}^0(i,j)$, is subsequently calculated for each season and subtracted to obtain anomalies of squared seasonal mean SLP gradients:

$$G_t(i,\ j) = G_t^0(i,\ j) - \bar{G}^0(i,\ j). \quad (5)$$

The use of the $\cos(\phi_j)$ weighting above accounts for the dependence of the spherical distance between two neighboring gridpoints on the latitude at which both gridpoints are located (for gridpoints on the equator: $\phi_j = 0$ and $\cos(\phi_j) = 1$). This computation is performed for each of the $5° \times 5°$ grid boxes analyzed, and for each season and year. The unit for the resulting squared SLP gradient is $(\text{hPa})^2$ per 5° spherical distance.

Appendix B Estimation of internal climate variability

The optimal detection analysis requires knowledge of the internal climate variability. It involves two estimates of the internal variability in this study: $\hat{C}_{\eta_1}$ and $\hat{C}_{\eta_2}$. Both $\hat{C}_{\eta_1}$ and $\hat{C}_{\eta_2}$ are climate model based estimates of internal climate variability; both are obtained by pooling the variability of the control simulations together with the inter-integration variability of the twentieth-century simulations (after removing the ensemble mean field from each integration in each of the nine ensembles; see Table 1). However, they are based on two non-overlapping periods of the twentieth-century simulations and two non-overlapping ensembles of control simulations.

More specifically, when analyzing the 1955–2004 (or 1958–2001) trends, $\hat{C}_{\eta_1}$ is based on the 41 twentieth-century simulations for the period 1955–2004 (or 1958–2001), while $\hat{C}_{\eta_2}$ is based on the 41 twentieth-century simulations for the period 1900–1949 (or 1900–1943). Accordingly, when analyzing the 1900–1949 trends (see Section 5), $\hat{C}_{\eta_1}$ is based on the 41 twentieth-century simulations for the period 1900–49, while $\hat{C}_{\eta_2}$ is based on the 41 twentieth-century simulations for the period 1955–2004. In the mean time, each of the available control simulations (Table 2) is also divided into non-overlapping 50 or 44-year simulation segments, depending upon whether the detection analysis was of a 50-year period (1955–2004 or 1900–1949) or a 44-year period (1958–2001). For example, the ECHO-G 340-year control simulation can be divided into six non-overlapping 50-year simulation segments, corresponding to years 1–50, 51–100,..., and 251–300, respectively (or seven non-overlapping 44-year segments corresponding to 1–44, 45–88,..., and 265–308; see Table 2). As a result, a total of 86 50-year control simulation segments (or 96 44-year segments) were obtained (Table 2). Half of these control simulation segments are used to obtain $\hat{C}_{\eta_1}$, and the other half, $\hat{C}_{\eta_2}$. Thus, combining these with information derived from the 41 twentieth-century simulations, we use a total of 84 (= 86/2 + 41) 50-year simulation segments for the detection analysis on the 1955–2004 or 1900–1949 trend pattern. The degree of freedom here is 74 (= 84−9−1), since nine twentieth-century ensemble-mean trend fields and one control ensemble-mean field (here we pool all control simulation segments used into a single large ensemble because some models have short control simulations) were subtracted from the estimated trend patterns before they are used in the detection analysis. Similarly, for the detection analysis on the 1958–2001 trend pattern, a total of 96 non-overlapping 44-year control simulation segments are used (see Table 2); thus the degree of freedom is 79 (= 96/2 + 41−1−9).

References

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

Alexandersson H, Tuomenvirta H, Schmith T, Iden K (2000) Trends of storms in NW Europe derived form an updated pressure data set. Clim Res 14:71–73

Allan R, Ansell TL (2006) A new globally complete monthly historical mean sea level pressure data set (HadSLP2): 1850–2004. J Clim 19(22):5816–5842

Allen MR, Stott PA (2003) Estimating signal amplitudes in optimal fingerprinting, part I: theory. Clim Dyn 21:477–491

Bacon S, Carter DJT (1991) Wave climate changes in the North Atlantic and North Sea. Int J Climatol 11:545–558

Barnett TP, Pierce D, AchutaRao K, Santer B, Gleicker P (2005) Penetration of human-induced warming into the world's oceans. Science 309:284–287

Blackmon ML (1976) A climatological spectral study of the 500-mb geopotential height of the Northern Hemisphere. J Atmos Sci 33:1607–1623

Caires S, Sterl A, Komen G, Swail V (2004a) The Web-based KNMI/ERA-40 global wave climatology atlas. WMO Bull 53(2):142–146

Caires S, Sterl A, Bidlot J-R, Graham N, Swail V (2004b) Intercomparison of different wind-wave reanalyses. J Clim 17(10):1893–1913

Cox AT, Swail VR (2001) A global wave hindcast over the period 1958–1997: validation and climate assessment. J Geophys Res 106(C2):2313–2329

Gillett NP, Zwiers FW, Weaver AJ, Stott PA (2003) Detection of human influence on sea-level pressure. Nature 422:292–294

Gillett NP, Allan RJ, Ansell TJ (2005) Detection of external influence on sea level pressure with multi-model ensemble. Geophys Res Lett 32:L19714. doi:10.1029/2005GL023640

Gulev SK, Zolina O, Grigoriev S (2001) Extratropical cyclone variability in the Northern Hemisphere winter from the NCEP/NCAR reanalysis data. Clim Dyn 17:795–809

Hasselmann K (1993) Optimal fingerprint for detection of time dependent climate change. J Clim 6:1957–1971

Hegerl G, Hasselmann K, Cubasch U, Mitchell JFB, Roeckner E, Voss R, Waszkewitz J (1997) On multi-fingerprint detection and attribution of greenhouse gas and aerosol forced climate change. Clim Dyn 13:613–634

Hegerl GC et al (2007) Understanding and attributing climate change. In: Solomon S et al (eds) Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. Cambridge University Press, New York, p 944

Hoskins BJ, Hodges KI (2002) New Perspectives on the Northern Hemisphere Winter Storm Tracks. J Atmos Sci 59:1041–1061

IDAG (International ad hoc Dectection and Attribution Group) (2005) Detecting and attributing external influences on the climate system: A review of recent advances. J Clim 18:1291–1314

Jones GS, Tett SFB, Stott PA (2003) Causes of atmospheric temperature change 1960–2000: A combined attribution analysis. Geophys Res Lett 30:1228. doi:10.1029/2002GL016377

Kalnay E et al (1996) The NCEP/NCAR 40-year reanalysis project. Bull Amer Meteor Soc 77:437–471

Kistler R et al (2001) The NCEP-NCAR 50-year reanalysis: monthly means CD-Rom and documentation. Bull Am Meteor Soc 82:247–267

Matulla C, Schoener W, Alexandersson H, von Stroch H, Wang XL (2008) European Storminess: late 19th century to present. Clim Dyn 31:125–130. doi:10.1007/s00382-007-0333-y

Meehl GA et al (2007) Global climate projections. In: Solomon SC et al (eds) limate change 2007: The physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. Cambridge University Press, New York, p 944

Nakicenovic N, Swart R (eds) (2000) Special report on emissions scenarios (SRES). Cambridge University Press, New York, p 599

Pettersen S (1956) Weather analysis and forecasting, vol 1, 2nd edn. McGraw-Hill, NY, p 422

Santer BD et al (2003) Contributions of anthropogenic and natural forcing to recent tropopause height changes. Science 301:479–483

Solomon S et al (2007) Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. Cambridge University Press, New York, p 944

Stott PA (2003) Attribution of regional-scale temperature changes to anthropogenic and natural causes. Geophys Res Lett 30:1724. doi:10.1029/2003GL017324

Tett SFB, Stott PA, Aleen MR, Ingram W, Mitchell J (1999) Causes of twentieth-century temperature change near the Earth's surface. Nature 339:569–572

Uppala SM et al (2005) The ERA-40 re-analysis. Q J Roy Meteor Soc 131:2961–3012

Wang XLL, Swail VR (2001) Changes of extreme wave heights in northern hemisphere oceans and related atmospheric circulation regimes. J Clim 14:2204–2221

Wang XLL, Swail VR (2002) Trends of Atlantic wave extremes as simulated in a 40-year wave hindcast using kinematically reanalyzed wind fields. J Clim 15:1020–1035

Wang XLL, Swail VR (2006a) Historical and possible future changes of wave heights in northern hemisphere oceans. Atmosphere Ocean Interactions, vol 2. In: Perrie W (ed) Advances in fluid mechanics Series vol 39, Wessex Institute of Technology Press, Southampton, p 240

Wang XLL, Swail VR (2006b) Climate change signal and uncertainty in projections of ocean wave heights. Clim Dyn 26:109–126. doi:10.1007/s00382-005-0080-x

Wang XLL, Swail VR, Zwiers FW (2006) Climatology and changes of extra-tropical cyclone activity: comparison of ERA-40 with NCEP/NCAR Reanalysis for 1958–2001. J Clim 19:3145–3166. doi:10.1175/JCLI3781.1

WASA Group (1998) Changing waves and storms in the Northeast Atlantic? Bull Am Meteorol Soc 79:741–760

Whitaker LM, Horn LH (1984) Northern Hemisphere extra-tropical cyclone activity for four midseason months. J Climotol 4:297–310

Zhang X, Zwiers FW, Stott PA (2006) Multi-model multi-signal climate change detection at regional scale. J Clim 19:4294–4307

Zhang X, Zwiers FW, Hegerl GC, Lambert FH, Gillett NP, Solomon S, Stott PA, Nozawa T (2007) Detection of human influence on twentieth-century precipitation trends. Nature 448:461–465

Zwiers FW, Zhang X (2003) Towards regional scale climate change detection. J Clim 16:793–797