

Xiaolan L. Wang · Val R. Swail

## Climate change signal and uncertainty in projections of ocean wave heights

Received: 10 August 2005 / Accepted: 23 September 2005 / Published online: 9 November 2005  
© Springer-Verlag 2005

**Abstract** In this study, projections of seasonal means and extremes of ocean wave heights were made using projections of sea level pressure fields conducted with three global climate models for three forcing-scenarios. For each forcing-scenario, the three climate models' projections were combined to estimate the multi-model mean projection of climate change. The relative importance of the variability in the projected wave heights that is due to the forcing prescribed in a forcing-scenario was assessed on the basis of ensemble simulations conducted with the Canadian coupled climate model CGCM2. The uncertainties in the projections of wave heights that are due to differences among the climate models and/or among the forcing-scenarios were characterized. The results show that the multi-model mean projection of climate change has patterns similar to those derived from using the CGCM2 projections alone, but the magnitudes of changes are generally smaller in the boreal oceans but larger in the region nearby the Antarctic coastal zone. The forcing-induced variance (as simulated by CGCM2) was identified to be of substantial magnitude in some areas in all seasons. The uncertainty due to differences among the forcing-scenarios is much smaller than that due to differences among the climate models, although it was identified to be statistically significant in most areas of the oceans (this indicates that different forcing conditions do make notable differences in the wave height climate change projection). The sum of the model and forcing-scenario uncertainties is smaller in the JFM and AMJ seasons than in other seasons, and it is generally small in the mid-high latitudes and large in the tropics. In particular, some areas in the northern oceans were projected to have large changes by all the three climate models.

---

### 1 Introduction

Changes in the Earth's climate occur as a result of both internal variability within the climate system and external factors. The external factors can be anthropogenic or natural. The increasing concentration of atmospheric greenhouse-gases (due to anthropogenic emissions) tends to warm the surface and lower atmosphere of the Earth, while an increase in some types of aerosols tends to cool them (Houghton et al. 2001). Natural factors, such as changes in solar output or explosive volcanic activity, can also cause radiative forcing and hence influence the Earth's climate.

Complex physically-based climate models are required to provide detailed estimates of feedbacks and regional features in the climate system. Although confidence in the ability of these models to provide useful projections of future climate has improved due to their demonstrated performance on a range of space and time-scales, the present-day climate models cannot yet simulate all aspects of climate (Houghton et al. 2001). For example, there are particular uncertainties associated with clouds and their interaction with radiation and aerosols.

Actually, there are several levels of uncertainty in the generation of "regional" climate change information [including generation of parameters/elements that are not directly available from the output of atmosphere-ocean general circulation models (AOGCMs), such as ocean wave heights]. The first level of uncertainty is associated with alternative scenarios of future emissions, their conversion to atmospheric concentrations and the radiative effects of these. This level of uncertainty is also associated with social and technological developments. The second level of uncertainty is related to the simulation of the transient climate response by AOGCMs for a given emission (forcing) scenario. The final level of uncertainty occurs when the AOGCM data are used to generate "regional" climate change information. In this regard, uncertainties are associated with imperfect knowledge and/or representation of physical processes,

---

X. L. Wang (✉) · V. R. Swail  
Climate Research Branch, Meteorological Service of Canada,  
4905 Dufferin Street, Downsview, ON M3H 5T4, Canada  
E-mail: Xiaolan.Wang@ec.gc.ca  
Tel.: +1-416-7394115  
Fax: +1-416-7395700

limitations due to the numerical approximation, simplification and assumptions in the models and/or approaches, and inter-model or inter-method differences in the simulation of climate response to given forcing conditions (Houghton et al. 2001).

Besides, climate model simulations are a combination of a forced climate change component together with internally generated natural variability. The internal variability of the global and regional climate system adds a further level of uncertainty in the evaluation of a climate change simulation.

As an important element of the climate system, ocean wave heights (among many other ocean surface characteristics) could be affected by anthropogenic forcing. However, ocean wave heights are not directly available from the output of global climate models. Useful projections of future wave height climate need to be produced through dynamical or statistical “downscaling” approaches, just like other regional climate change information. Therefore, there are various sources of uncertainty in the generation of ocean wave height climate change projections. Using projections of future climate made with the Canadian coupled climate model CGCM2 for the IPCC IS92a and the SRES (special report on emissions scenarios; Nakicenovic and Swart 2000) A2 and B2 forcing-scenarios, Wang and Swail (2005) and Wang et al. (2004a) produced ocean wave climate change scenarios for the northern hemisphere oceans for the twenty-first century. Their results show that significant changes can be anticipated in both the North Atlantic and the North Pacific under all the three forcing-scenarios. The rate and sign of the projected future wave height changes are not constant throughout the century; and in some regions, these appear to be quite dependent on the forcing conditions. The rate of change appears to have a positive relationship with the rate of increase in the greenhouse-gases forcing (Wang and Swail 2005). These among-scenarios differences lie within the first level of uncertainty mentioned above. The existence of these various sources of uncertainty requires us to characterize the uncertainty, to evaluate the level of confidence in a “regional” climate change simulation.

One of the criteria to evaluate the level of confidence in a “regional” climate change simulation can be based on how well the climate change simulations converge across models and methods (Houghton et al. 2001). Aiming to characterize the emission-scenario and AOGCM related uncertainties (i.e., the first and second levels of uncertainty), to provide a better estimate of the projected climate change of ocean wave heights, this study made and analyzed projections of ocean wave heights using projections of future climate conducted with three coupled AOGCMs, for the IS92a, A2 and B2 forcing-scenarios.

The rest of this paper is organized as follows. The data sets and methodologies are briefly described in Sects. 2 and 3, respectively. The estimates of climate changes projected by the three climate models combined are presented in Sect. 4. The relative importance of

forcing-induced variance and the emission-scenario and AOGCM related uncertainties in the wave height climate change projections are discussed in Sects. 5 and 6, subsequently. This study is completed with a summary and some discussions in Sect. 7.

## 2 Datasets

It has been shown that, on the monthly/seasonal time scale, significant wave height (SWH) variations in the northern hemisphere oceans over the past four decades are closely associated with contemporaneous mean sea level pressure (SLP) variations in the region (Wang and Swail 2001, 2002; Kushnir et al. 1997). Such relationships between wave height statistics and SLP fields have been used to make projections of wave height climate change scenarios (Wang et al. 2004a; Wang and Swail 2005; WASA Group 1998; von Storch and Reichardt 1997). Similar “downscaling” approaches as in Wang et al. (2004a) were adopted in this study. The SLP–SWH relationships in each season were represented by a pair of regression models, one for seasonal means of SWH, another for seasonal extremes of SWH. Thus, observations of seasonal SLP quantities, and of seasonal means and maxima of SWH are needed to train the regression models. Seasonal means and maxima of SWH were derived from the ERA-40 wave data for 1958–2001 (Uppala 2001; Caires et al. 2004a), which are available at six-hourly intervals on a 1.5°-by-1.5° lat/long grid over the oceans. Similarly, monthly and seasonal mean SLP fields were derived from the six-hourly SLP of the ERA-40 reanalysis for 1958–2001 (Uppala 2001; Simmons and Gibson 2000; Gibson et al. 1996), which are on a global 2.5°-by-2.5° lat/long grid. Then, the monthly squared SLP gradients (i.e., sum of the squared zonal and squared meridional SLP gradients, which is proportional to wind energy) were also calculated from monthly means of SLP, and subsequently seasonally averaged. Further, we refer the 30-year period from 1961 to 1990 as the baseline period, and the baseline period mean field, as the baseline climate. The baseline climate of the ERA-40 seasonal mean SLP (or SLP gradient index) was subtracted and the remaining quantities (i.e., anomalies) were used as predictors when training the regression models. In other words, the predictors are anomalies of seasonal mean SLP and SLP gradient index relative to their baseline climates. Thus, projections of future anomalies of the SLP quantities (i.e., AOGCM simulated SLP quantities minus its baseline climate as simulated by the same AOGCM) are needed to feed into the regression model to make projections of the predictand (i.e., seasonal means and maxima of SWH). These were obtained as follows.

First, projections of future seasonal mean SLP and SLP gradient index were obtained from three coupled AOGCMs simulations for the IS92a, A2, and B2 forcing-scenarios (downloaded from the IPCC Data Distri-

bution Centre web-site: [http://ddcweb1.cru.uea.ac.uk/asres/gcm\\_data/GCM\\_data.html](http://ddcweb1.cru.uea.ac.uk/asres/gcm_data/GCM_data.html)). The three climate models are: the Canadian CGCM2 (Flato and Boer 2001), the Hadley Centre's HadCM3 (Gordon et al. 2000), and the ECHAM4/OPYC3 of the Max-Planck-Institut (Roeckner et al. 1996a, b). These models were selected because they have similar resolutions among the global climate models that have climate change projections for the IS92a, A2 and B2 forcing-scenarios for a common period of considerable length (cf. the IPCC Data Distribution Centre web-site; the spatial resolution of the data is given below). For each of the three forcing-scenarios, three integrations with the same forcing but different initial conditions were run with CGCM2, but only one integration was run with each of the other two climate models. The IS92a scenario simulations with CGCM2 and HadCM3 cover the period 1900–2100, but with ECHAM4/OPYC3 it covers the period 1860–2049. All the A2 and B2 scenarios simulations cover the period 1990–2099 (or 1950–2099 with HadCM3).

The ERA-40 SLP fields (used to represent observations) are available on a global 2.5°-by-2.5° latitude-longitude grid. The projections of SLP fields are on a global grid of 96-by-73 grid-points (2.5°-latitude by 3.75°-longitude) for HadCM3, of 128-by-64 grid-points (approximately 2.79°-latitude by 2.8125°-longitude) for ECHAM4/OPYC3, and of 96-by-48 grid-points (approximately 3.75°-latitude by 3.75°-longitude) for CGCM2. All these SLP fields were converted to a global 96-by-48 Gaussian grid used by CGCM2.

Then, for each climate model, the simulated baseline climate of seasonal mean SLP was calculated and used to derive the simulated anomalies of seasonal mean SLP, which was done for each forcing-scenario separately. Note that the CGCM2 or ECHAM4/OPYC3 simulations for the A2 or B2 scenario do not cover the baseline period; thus, the simulated anomalies of seasonal mean SLP for the A2 or B2 scenario were derived using the baseline climate of the same variable simulated by the same model for the IS92a scenario. Similarly, the simulated anomalies of SLP gradient index were derived.

The simulated anomalies of seasonal mean SLP and SLP gradient index were then used as predictors in the SLP–SWH regression relationships to make projections of future seasonal means and extremes of SWH (see Sect. 3.1 below). The projections of ocean wave heights were then subject to analysis of variance (ANOVA) analyses to assess the relative importance of the forcing-induced variance (climate change signal), and to characterize the emission-scenario and AOGCM related uncertainties (see Sects. 3.2, Appendix).

### 3 Methodologies

#### 3.1 Regression and non-stationary GEV analyses

As mentioned before, projections of SWH are based on the SLP–SWH relationships derived from the ERA-40

data, which are represented by regression models in this study. Since we are dealing with both seasonal means and extremes of SWH (Gaussian and non-Gaussian variables), both conventional and generalized regression models were used here to represent the SLP–SWH relationships.

For the seasonal means (Gaussian variable), the following regression model was fitted:

$$h_t = a + bP_t + cG_t + \varepsilon_t, \quad (1)$$

where  $h_t$  denotes the time series of seasonal mean SWH anomaly at a wave gridpoint,  $P_t$  and  $G_t$  are the time series of seasonal mean SLP anomaly and SLP gradient index for the wave gridpoint, respectively, and  $\varepsilon_t$  denotes a Gaussian random process. Note that the SLP quantities ( $P_t$  and  $G_t$ ), which are required for each wave gridpoint, were calculated from values at four nearest SLP grid-points (all within about 500 km radius from the wave gridpoint), using weights proportional to the inverse of the distance. Here,  $P_t$  contains information about the mean state of SLP at time  $t$ , and  $G_t$  contains information about its variation in space, over the area within 500 km radius from the wave gridpoint (Wang et al. 2004a).

For each wave gridpoint, model (Eq. 1) was trained using the time series  $h_t$ ,  $P_t$ , and  $G_t$  derived from the ERA-40 wave and SLP fields for 1958–2001 (44 years). The statistical significance of the regression parameter  $b$  or  $c$  (which represents the relationship between the predictand  $h_t$  and the predictor  $P_t$  or  $G_t$ ) was determined by performing likelihood ratio tests, in which the sum of squared errors (SSE) of the full model (Eq. 1) was compared with the SSE of the one-predictor model that excludes the predictor being tested (i.e.,  $h_t = a + c G_t + \varepsilon_t$  for the significance of  $b$ ,  $h_t = a + b P_t + \varepsilon_t$  for that of  $c$ , and  $h_t = a + \varepsilon_t$  for the significance of  $b$  and  $c$  combined; Johnson and Wichern 1982). Note that skillful predictions of the predictand are impossible without a close relationship between the predictors and the predictand. Fortunately, the results (cf. Table 1, Fig. 1) show that variation of seasonal mean SWH is closely related to both predictors  $P_t$  and  $G_t$ . Thus, cross-validation was used to determine the skill of model (Eq. 1). Specifically, we withheld all data at times  $t-1$ ,  $t$ , and  $t+1$  when training model (Eq. 1) to make a cross-validated prediction for time  $t$  ( $t=1,2,\dots,44$  for the 44 year period from 1958 to 2001). Then, the predictive skill of model (Eq. 1) is represented by the serial correlation between the ERA-40 predictand time series (i.e., the ERA-40 seasonal mean SWH) and its cross-validated prediction time series. The results (cf. Fig. 2) show that model (Eq. 1) has highly significant (at least at 5% level) predictive skill at the vast majority (about 60–70%) of the area of wave data. Low predictive skill is mainly seen in the narrow zone along 30° S and along the equator, as well as in the 20°N–30°N zone. Since model (Eq. 1) is significantly skillful at about two-thirds of the area of wave data, time series of  $P_t$  and  $G_t$  derived from the climate model simulations were substituted into the fit-



**Table 1** The percentages of area in which the indicated regression parameters ( $b$ ,  $c$ ,  $r_1$ , and  $r_2$ ; see Sect. 3.1) were found to be significantly different from zero at the 5% level, which measure the field significance of the regression parameters

Predictand	Parameters	JFM	AMJ	JAS	OND
Mean SWH	$b$	51.9	41.6	43.5	42.7
	$c$	70.3	63.9	65.4	67.7
	$b$ and $c$	80.5	73.1	74.9	75.0
Maximum SWH	$r_1$ (GEV <sub>1</sub> vs. GEV <sub>0</sub> )	27.2	19.1	22.0	22.9
	$r_2$ (GEV <sub>2</sub> vs. GEV <sub>1</sub> )	43.8	39.4	39.8	41.3
	$r_1$ and $r_2$ (GEV <sub>2</sub> vs. GEV <sub>0</sub> )	51.4	43.4	46.7	47.4

According to Livezey and Cheng (1983), a rejection rate of 19% or higher would indicate a field significance at the 5% level in a field of 18 or more spatial degrees of freedom. The ERA-40 significant wave height data is very likely to have more than 18 spatial degrees of freedom, because the number of its leading EOFs (Empirical Orthogonal Functions) required to represent even 80% of its total variance is 25 (and 37 for 95% of the total variance). Thus, these

rejection rates indicate a field significance at the 5% level or higher. Note that leading EOFs are EOFs ordered by their associated percentages of the total variance, from the largest to the smallest. A dataset is unlikely to have a large number of spatial degrees of freedom if a small number of its leading EOFs can explain most of its total variance

ted model (Eq. 1) to make projections of seasonal mean SWH anomalies for the period of 1990–2099 (1990–2049 for the ECHAM4/OPYC3 IS92a scenario), i.e.,

$$\hat{h}_t = \hat{a} + \hat{b}P_t + \hat{c}G_t, \quad (2)$$

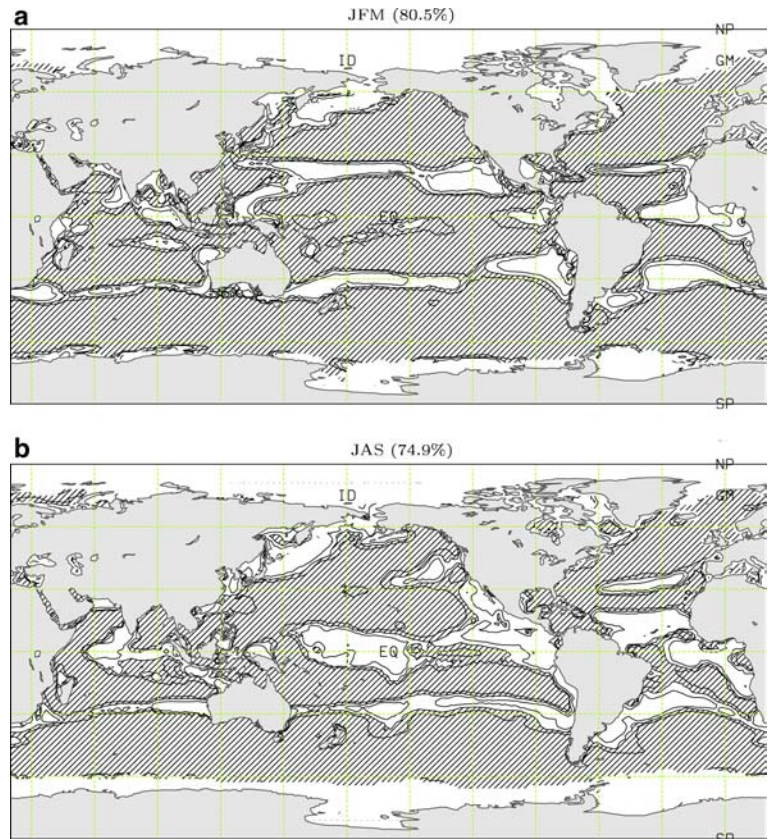
where (and throughout this paper)  $\hat{y}$  denotes an estimate of  $y$ . The projected time series  $\hat{h}_t$  was then subject to a trend analysis (see Sect. 4 below). The projections were also subject to an ANOVA (see Sects. 3.2, Appendix).

Note that non-Gaussian behaviour is a particular concern for extremes, and extremes from the changing climate system are most likely those of a non-stationary process and thus of time-dependent characteristics.

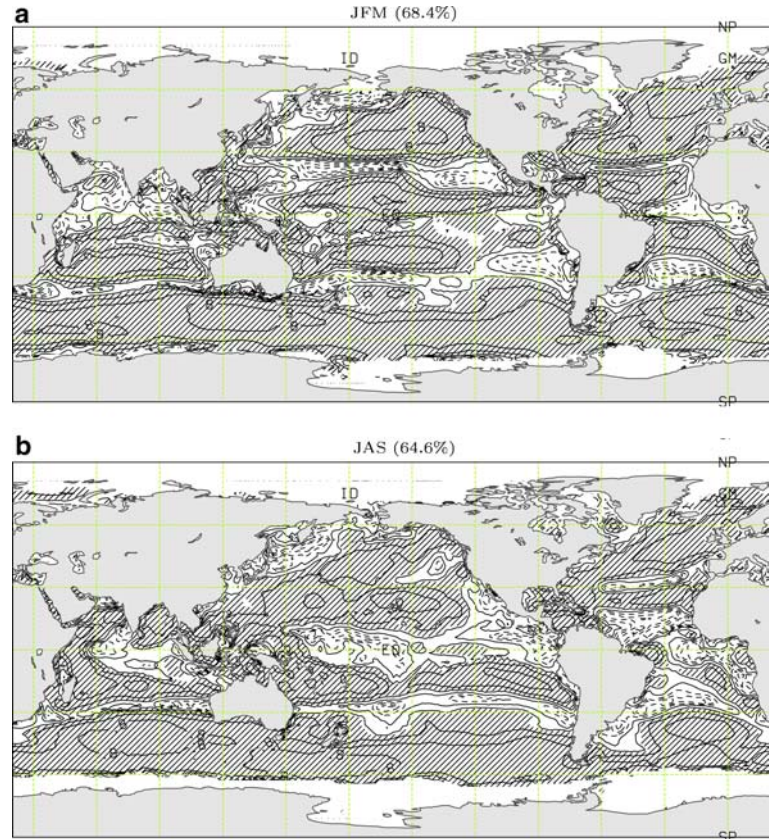
Therefore, the non-stationary generalized extreme value (GEV) analysis as described in Wang et al. (2004a) was carried out in this study to represent the relationship between seasonal extremes of SWH and the predictors  $P_t$  and  $G_t$ . Specifically, the following five nested GEV models were fitted to the ERA-40 seasonal maxima of SWH at each wave gridpoint:

- GEV<sub>0</sub>( $\mu, \sigma, \xi$ );
- GEV<sub>1</sub>( $\mu_t = \mu_o + r_1 P_t, \sigma, \xi$ );
- GEV<sub>2</sub>( $\mu_t = \mu_o + r_1 P_t + r_2 G_t, \sigma, \xi$ );
- GEV<sub>3</sub>( $\mu_t = \mu_o + r_1 P_t + r_2 G_t, \log(\sigma_t) = b_o + q_1 P_t, \xi$ );
- GEV<sub>4</sub>( $\mu_t = \mu_o + r_1 P_t + r_2 G_t, \log(\sigma_t) = b_o + q_1 P_t + q_2 G_t, \xi$ ),

**Fig. 1** The  $p$  values of the regression parameters  $b$  and  $c$  (combined), representing the significance level ( $1-p$ ) of the relationships between seasonal mean SWH and the predictors  $P_t$  and  $G_t$ , respectively. The contour values are  $p = 0.80$ , 0.95, and 0.99. Areas of  $p \geq 0.95$  are hatched; and their percentages are given in parentheses above



**Fig. 2** Serial correlations between the ERA-40 seasonal mean significant wave heights (i.e., the predictand time series) and its cross-validated prediction time series, which measure the predictive skill of model (Eq. 1) (see Sect. 3.1). The contour interval is 0.2. Dashed and solid lines indicate negative and non-negative contours, respectively. Hatching indicates areas of significant (at 5% level) predictive skill. The percentages of these areas are given in parentheses about



where  $\mu$  (or  $\mu_t$ ),  $\sigma$  (or  $\sigma_t$ ), and  $\xi$  are the location, scale, and shape parameters of the GEV distribution, respectively. The statistical significance of the linear relationships built into the GEV models (and the goodness of fit of the GEV models themselves) was assessed by performing likelihood ratio tests (see Wang et al. 2004a for the details). Consistent with what was reported by Wang et al. (2004a) and Wang and Swail (2005), results of the tests show that the location parameter is significantly correlated with both predictors  $P_t$  and  $G_t$  (cf. Table 1), but the scale parameter appears to be independent of either  $P_t$  or  $G_t$ . In other words, the above  $GEV_2$  is the model of best fit. Therefore, the fitted  $GEV_2$  was used to make projections of SWH extremes. Specifically, time series of  $P_t$  and  $G_t$  derived from the climate model simulations were substituted into the fitted expression for the location parameter to produce time series of the location parameter estimates for the period of 1990–2099 (1990–2049 for the ECHAM4/OPYC3 IS92a scenario):

$$\hat{\mu}_t = \hat{\mu}_0 + \hat{r}_1 P_t + \hat{r}_2 G_t. \quad (3)$$

Such time series were then substituted into the fitted  $GEV_2(\hat{\mu}_t, \hat{\sigma}, \hat{\xi})$  to estimate possible future 20-year return values of SWH, which were then subject to ANOVA analyses (see Sects. 3.2, Appendix). Besides, as in Wang et al. (2004a), the time series of the projected location parameter values  $\hat{\mu}_t$  were subject to a linear/quadratic trend analysis (see Sect. 4 below).

Note that there are a total of 15 simulations: one for each of the three forcing-scenarios with each of the three climate models (that is  $3 \times 3$ ), plus two extra integrations for each of the three forcing-scenarios with the CGCM2 (that is  $2 \times 3$ ). For each of the 15 simulations, the above projections were done for each of the four seasons separately, with the four seasons being defined as January–March (JFM), April–June (AMJ), July–September (JAS), and October–December (OND).

Data discontinuity is always a concern, especially for the SH, because there were generally much fewer observation data available in the pre-satellite era and there were very few data over the SH oceans in that period (Sterl 2004; Caires et al. 2004b; Wang et al. 2005). The possible discontinuity could be a mean-shift or an abrupt change in the variance or both (Wang et al. 2005). This kind of discontinuity may or may not have an effect on the above regression relationships (between predictand  $h_t$  or  $\mu_t$  and the predictors  $P_t$  and  $G_t$ ). In order to assess the effects of possible data discontinuity around year 1979 (when satellite data started to exist), we also carried out the above regression analyses using normalized predictand and predictors time series.

More specifically, we normalized the time series  $P_t$  and  $G_t$ , and the wave height data (seasonal means and maxima) series with respect to their means and variances in the first and second 22-year periods (1958–1979 and 1980–2001), separately. Say, let  $\bar{x}_1$  and  $\hat{\sigma}_1^2$  denote respectively the mean and variance of time series  $x_t$  in



the first 22-year period, and  $\bar{x}_2$  and  $\hat{\sigma}_2^2$ , in the second 22-year period. Then, we normalized time series  $x_t$  as follows:  $y_t = (x_t - \bar{x}_1)/\hat{\sigma}_1$  for  $t \leq 1979$  and  $y_t = (x_t - \bar{x}_2)/\hat{\sigma}_2$  for  $t > 1979$ . The possible abrupt changes in the mean/variance around year 1979 should be greatly diminished from the normalized time series  $y_t$ . Thus, we also carried out the above regression analyses using the normalized predictand and predictors time series. The climate model projections of the predictors time series  $P_t$  and  $G_t$  were also normalized with respect to their means and variances in the whole projection period (mostly 1990–2099) and used to make projections of the normalized predictand value. The mean and standard deviation of the corresponding projections made without the normalization procedure were calculated and used to scale the projections of the normalized predictand value (e.g.,  $\hat{x}_t = \hat{y}_t * \hat{\sigma}_0 + \bar{x}_0$ , where  $\bar{x}_0$  and  $\hat{\sigma}_0$  are calculated from the projections of  $x_t$  made without the normalization procedure), so that projections made with and without the normalization procedure have the same mean and variance and can be compared with each other to assess the possible effects of data discontinuity around year 1979. In general, the results show good agreement in terms of the patterns of projected change (see Sect. 4 below for more details).

### 3.2 Analysis of variance

Since there are various sources of uncertainty in the generation of projections of wave height climate change information, it is necessary to characterize the various sources of uncertainty and to assess the relative importance of the forcing-induced variance. To this end, both the one- and two-factor ANOVA models (Huitson 1966; von Storch and Zwiers 1999) were used in this study.

Note that ensembles of the CGCM2 simulations have been produced with the same projected forcing but with different initial conditions that lead to different evolutions of the internal natural variability. Thus, an assessment of the relative importance of the CGCM2 projected climate change can be done with a one-factor ANOVA-based test.

Let  $Y_{ts}$  denote time series of SWH (means or extremes) projected by the CGCM2 with one of the three forcing-scenarios (time  $t=1,2,\dots,n$ , where  $n=110$  for period 1990–2099, simulation  $s=1,2,\dots,S$  where  $S=3$  for a 3-member ensemble; the word “ensemble” is used to indicate  $S>1$ ). Then, the one-factor ANOVA model for  $Y_{ts}$  (Huitson 1966; von Storch and Zwiers 1999) has the form:

$$Y_{ts} = v + \beta_t + \epsilon_{ts}, \quad (4)$$

where  $v$  is the grand mean of  $Y_{ts}$ ,  $\beta_t$  are deviations from the grand mean that arise from the effects of the prescribed forcing (i.e., the forcing effects, which are common to all simulations with the same forcing-scenario), and  $\epsilon_{ts}$  represent internally generated variations specific to simulation  $s$ . Here, it is assumed that (1) the forcing effects  $\beta_t$  are deterministic and are treated as fixed

because the forcing conditions are prescribed identically in each simulation, (2)  $\epsilon_{ts}$  is an iid Gaussian random variable with mean zero and variance  $\sigma_\epsilon^2$ . Thus, model (Eq. 4) is a one-way fixed effects ANOVA (von Storch and Zwiers 1999).

Then, the null hypothesis that the prescribed forcing conditions do not influence the variability in  $Y_{ts}$  (i.e.,  $H_\beta: \beta_t=0$  for  $t=1,2,\dots,n$ ) can be tested using the statistic  $F_\beta$  defined in the Appendix. The proportion of the total variance in  $Y_{ts}$  that is due to the forcing effects can also be estimated (see estimator  $P_\beta$  in Appendix).

Mathematically/statistically, Zwiers et al. (2000) have shown that the above ANOVA-based test is much more powerful than the test that compares the ensemble variance with the variance of a “control” simulation (i.e., a simulation with zero forcing; see Sect. 5 of their paper). The power of test is proportional to  $S$ , the number of simulations in the ensemble (see Fig. 1 of Zwiers et al. 2000). This ANOVA-based test is also more powerful than a regression trend analysis approach, because it does not need to specify the form of the temporal variation (e.g., linear or quadratic trends...) in  $Y_{ts}$ . For example, if a straight line (linear trend) is fitted while the temporal variation is really quadratic, then the test on the linear time coefficient becomes less efficient because there is still a systematic time component that remains in the residuals and inflates the estimate of the residual variance. [Nevertheless, for the northern hemisphere oceans, Wang et al. (2004a, b) and Wang and Swail (2005) have estimated the systematic trends in  $Y_{ts}$  using a nested regression models approach. A similar regression approach is also used later in Sect. 4 below.]

Similarly, let  $X_{ijt}$  denote time series of seasonal SWH quantities (means or extremes) projected by climate model  $i$  with forcing-scenario  $j$  ( $i=1,2,\dots,m$ ;  $j=1,2,\dots,q$ ; time  $t=1,2,\dots,n$ . Here,  $m=3$ ,  $q=3$ , and  $n=60$  for period 1990–2049, which is the common period for which all the three climate models have simulations for the three forcing-scenarios). Then, we used the following two-factor ANOVA model (Huitson 1966; Wang and Zwiers 1999; von Storch and Zwiers 1999) to partition the total variance of  $X_{ijt}$  into four components:

$$X_{ijt} = \omega + \gamma_i + \theta_j + \delta_{ij} + \epsilon_{ijt}, \quad (5)$$

where  $\omega$  is the grand mean of  $X_{ijt}$ ,  $\gamma_i$  represent a bias in  $X_{ijt}$  that varies from one climate model to the next but is constant for a given model regardless of the kind of forcing that is used (and hence is referred to as “inter-model variability”);  $\theta_j$  represent the component of the response to forcing that varies from one forcing scenario to another but is common among all the three climate models (thus referred to as “inter-scenario variability”);  $\delta_{ij}$  represent variations in response to forcing  $j$  that vary from one model to the next (i.e., deviations from the mean-model response  $\theta_j$ ); and  $\epsilon_{ijt}$  are the effects of internal variability and the common forcing (forcing conditions that are common to all the three scenarios) and are assumed to have zero-mean and variance  $\sigma_\epsilon^2$ .

The model factor  $\gamma_i$  is treated as deterministic, because it is presumably characteristic of the group of climate models used. Similarly, the forcing-scenario factor  $\theta_j$  is also treated as deterministic, representing the characteristic of the group of forcing scenarios used. The interaction term  $\delta_{ij}$  (also referred to as “model-scenario uncertainty”) is needed here because not all models respond to the same forcing (say, the IS92a forcing) in the same way. Also, this term would be deterministic rather than random because it is presumably characteristic of the model (i.e., the same deviation from the common response would be produced each time that the same experiment is run). Therefore, model (Eq. 5) is a two-factor fixed effects ANOVA model.

Then, the null hypothesis that there are no inter-model variability effects in  $X_{ijt}$  (i.e.,  $H_\gamma: \gamma_i = 0$  for  $i = 1, 2, \dots, m$ ) can be tested using the statistic  $F_\gamma$  defined in the Appendix. As a result, the relative importance of the inter-model variability in  $X_{ijt}$  (relative to the effects of internal variability and the common forcing) is assessed.

Similarly, as described in the Appendix, the null hypothesis that (1) there are no inter-scenario variability effects in  $X_{ijt}$  (i.e.,  $H_\theta: \theta_j = 0$  for  $j = 1, 2, \dots, q$ ) can be tested using the statistic  $F_\theta$ ; (2) there are no interaction effects (i.e.,  $H_\delta: \delta_{ij} = 0$  for  $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, q$ ) can be tested using the statistic  $F_\delta$ ; and (3) there are no effects arising from the model and/or forcing-scenario uncertainties (i.e.,  $H_{\gamma+\theta+\delta}: \gamma_i = \theta_j = \delta_{ij} = 0$  for  $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, q$ ) can be tested using the statistic  $F_{\gamma+\theta+\delta}$  (see Appendix).

The proportion of the total variance in variable  $X_{ijt}$  that is due to the effects of inter-model variability, inter-scenario variability, and the interaction between them, as well as the total effects of the model and forcing-scenario uncertainties can also be estimated (see estimators  $P_\gamma$ ,  $P_\theta$ ,  $P_\delta$ , and  $P_{\gamma+\theta+\delta}$  in the Appendix).

Note that in the above two-factor ANOVA, only one of the three integrations with CGCM2 was used, because only one integration was available from the other two climate models (i.e., HadCM3 and ECHAM4/OPYC3). Also, for a given variance proportion of the factor being tested, the power of the above ANOVA-based tests is proportional to the  $S$  and  $n$  (one-way ANOVA), or to the  $m$ ,  $q$ , and  $n$  (two-way ANOVA). For example, with ensemble size  $S=3$  and time series length  $n=110$ , the forcing-induced variance has about 97% likelihood to be identified as statistically significant at 5% level when its variance proportion is 20%, and the likelihood reduces to about 55% when the variance proportion is 10% (courtesy of Dr. Sofia Caires). Similarly, the small  $S$ ,  $m$ , and  $q$  (all equal 3) limit the power of the ANOVA-based tests, but the large  $n$  ( $n=110$  or  $n=60$ ) more or less alleviates this problem.

projected for period 1990–2099 (110 years) with the three climate models were combined into a single time series, to estimate the mean climate change projected by the three climate models. More specifically, let  $x_i$  denote the projected value ( $\hat{h}_{t_i}$  or  $\hat{\mu}_{t_i}$ ) for year  $t_i$ , then  $t_i = \{1, 2, \dots, 110, 1, 2, \dots, 110, 1, 2, \dots, 110\}$  for  $i = \{1, 2, \dots, 110, 111, 112, \dots, 220, 221, 222, \dots, 330\}$ , that is, the first, second, and third group of 110 values in time series  $x_i$  were projected respectively by the first, second, and third climate models for the 110 years. Then, the following regression models were fitted to time series  $x_i$  at each grid point:

- $RM_0: x_i = \alpha_0 + \varepsilon_i$  (i.e., no trend in  $x_i$ );
- $RM_1: x_i = \alpha_0 + \alpha_1 t_i + \varepsilon_i$  (i.e., a linear trend in  $x_i$ );
- $RM_2: x_i = \alpha_0 + \alpha_1 t_i + \alpha_2 t_i^2 + \varepsilon_i$  (i.e., a quadratic trend in  $x_i$ ),

where  $\alpha_0$ ,  $\alpha_1$  and  $\alpha_2$  are the regression parameters, and  $\varepsilon_i$  denotes a zero-mean AR(1) process (i.e., red noise). Here, the use of red (instead of white) noise is to diminish the effect of positive autocorrelation on the power of test for the significance of trends (i.e., regression parameters  $\alpha_1$  and  $\alpha_2$ ; von Storch and Navarra 1995; Zhang and Zwiers 2004; Wang and Swail 2001). A higher order autoregressive process is not necessary here, since few partial correlations at lag  $> 1$  are significant at the (approximately) 5% level. Zhang and Zwiers (2004) show that the iteration scheme proposed by Wang and Swail (2001) for estimating the lag-1 autocorrelation ( $\rho$ ) and trend in a time series works better than a non-iteration scheme. Thus, the lag-1 autocorrelation ( $\rho$ ) of time series  $x_i$ , and the above regression parameters, are estimated by applying the iteration scheme of Wang and Swail (2001) on the pre-whitened time series  $w_i = (x_i - \rho x_{i-1})$  (for  $t_i = \{1, 2, \dots, 109, 1, 2, \dots, 109, 1, 2, \dots, 109\}$  and  $i = \{1, 2, \dots, 109, 110, 111, \dots, 218, 219, 220, \dots, 327\}$ ). The pre-whitening process is necessary only when  $\rho \geq 0.05$ , because only large positive autocorrelations would notably increase the apparent level of significance of trends (cf. von Storch and Navarra 1995). In other words, a white noise is assumed in the above regression analysis for all locations where the final estimate of  $\rho < 0.05$ .

Standard  $F$ -tests were used to intercompare pairs of regression models  $RM_j$  and  $RM_k$  ( $j > k$ ; using the pre-whitened time series when necessary), to determine the statistical significance of the estimated trend component (see Wang et al. 2004a for more details). The percentages of areas where  $RM_0$  was rejected in favor of  $RM_2$  are 42–62% for the A2 scenario, and 22–42% for the B2 scenario (cf. Figs. 3, 4, 5, 6, 7). According to Livezey and Cheng (1983), a rejection rate of 22% or higher would indicate a field significance at the 5% level in a field of 15 or more spatial degrees of freedom. The global wave height field in question is very likely to have more than 15 spatial degrees of freedom, because the number of its leading EOFs required to represent even 80% of its total variance is greater than 15 (it ranges between 19 and 41, depending on season). These indicate

#### 4 Projected changes in ocean wave heights

First of all, for either the A2 or the B2 forcing-scenario, the three time series of seasonal mean (or extreme) SWH



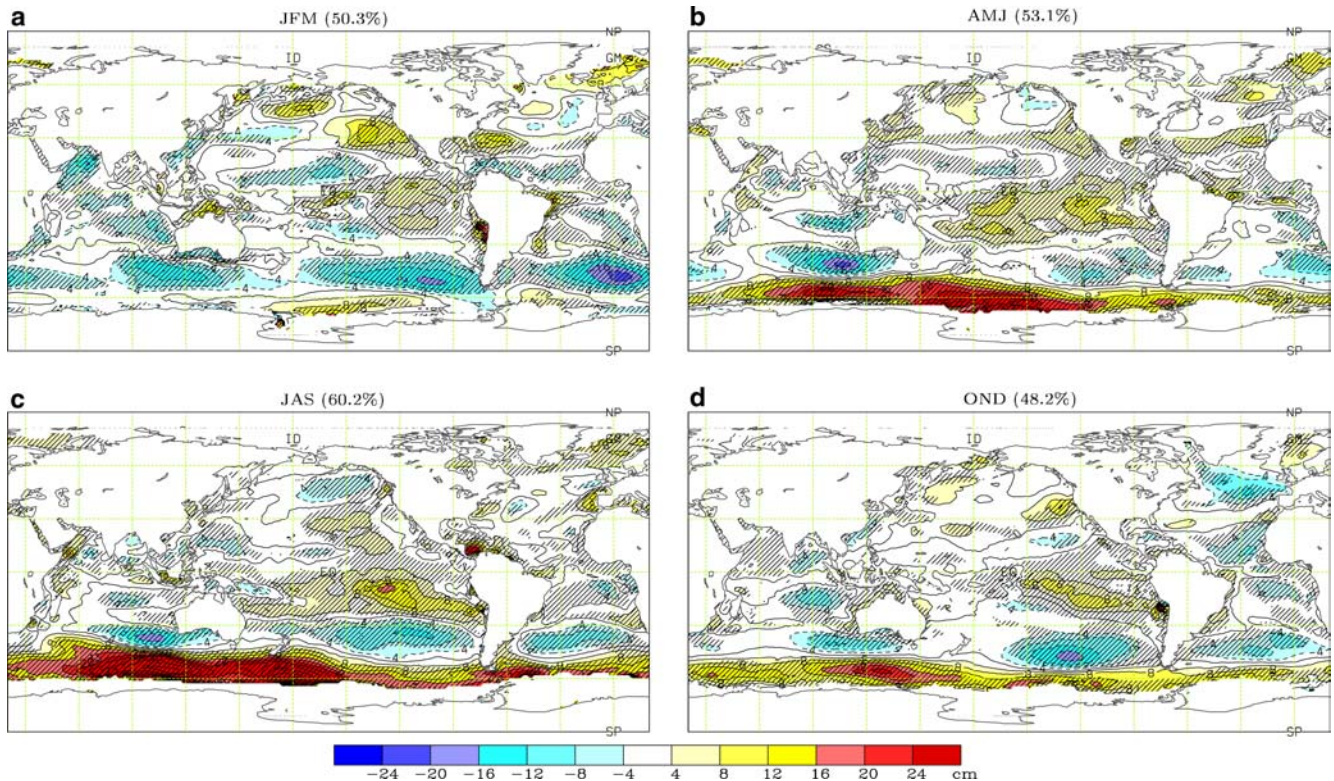
that both  $\hat{h}_{t_i}$  and  $\hat{\mu}_{t_i}$  have a linear/quadratic trend component that is of field significance at the 5% level, for both the A2 and B2 scenarios. Thus, the following trend curves were estimated:  $\hat{h}_{tr}(t_i) = \hat{\alpha}_0 + \hat{\alpha}_1 t_i + \hat{\alpha}_2 t_i^2$ , and  $\hat{\mu}_{tr}(t_i) = \hat{\alpha}_0 + \hat{\alpha}_1 t_i + \hat{\alpha}_2 t_i^2$ , where  $\hat{\alpha}_0$ ,  $\hat{\alpha}_1$ , and  $\hat{\alpha}_2$  denote the estimated parameters. Since the changes are non-linear, the differences  $[\hat{h}_{tr}(t_{91}) - \hat{h}_{tr}(t_1)]$ , or the differences between the two 20-year return values derived respectively from the fitted models  $\text{GEV}[\hat{\mu}_{tr}(t_{91}), \hat{\sigma}, \hat{\xi}]$  and  $\text{GEV}[\hat{\mu}_{tr}(t_1), \hat{\sigma}, \hat{\xi}]$ , were calculated and used to show the total changes in the period from time  $t_1$  to time  $t_{91}$ . Here,  $t_{91}=91$  and  $t_1=1$ , which corresponds respectively to year 2080 and year 1990. The total changes in period 1990–2080 are shown in Figs. 3, 4, 5, 6, 7.

For the seasonal means of SWH projected with the A2 forcing-scenario, as shown in Fig. 3, for boreal winter and fall (JFM and OND), the areas of large increases in the North Pacific (NP) and the North Atlantic (NA) are similar to what was reported by Wang and Swail (2005) using only the CGCM2 projections. In boreal winter, the projected changes are characterized by significant increases in the mid-latitudes of the eastern NP and high-latitudes of western NP, as well as in the northeast Atlantic and in the southwest NA (off the southeastern coast of the United States). The total increase during the period from 1990 to 2080 is up to

about 12 cm (cf. Fig. 3a; or equivalently about 6% of the climate value for year 1990). In boreal spring and summer (AMJ and JAS), significant increases in seasonal mean wave heights were projected for the northeast Atlantic, with some increases in the subtropical eastern NA and in the mid-latitudes of the eastern NP (Fig. 3b, c; up to about 10 cm or 5%). In boreal fall, the pattern of projected changes (Fig. 3d) is also similar to that shown in Wang and Swail (2005) for the northern oceans, however, changes projected by the three models combined are much smaller than those projected by CGCM2 alone.

In the southern hemisphere, the changes projected for the JFM mean SWH are characterized by significant decreases in the region between 40°S and 60°S and some increases nearby the Antarctic coastal zone (Fig. 3a). In the other three seasons, the projected changes are characterized by large increases nearby the Antarctic coastal zone and some increases in the subtropical South Pacific (SP), with decreases in the zone between 40°S and 60°S (Fig. 3b–d).

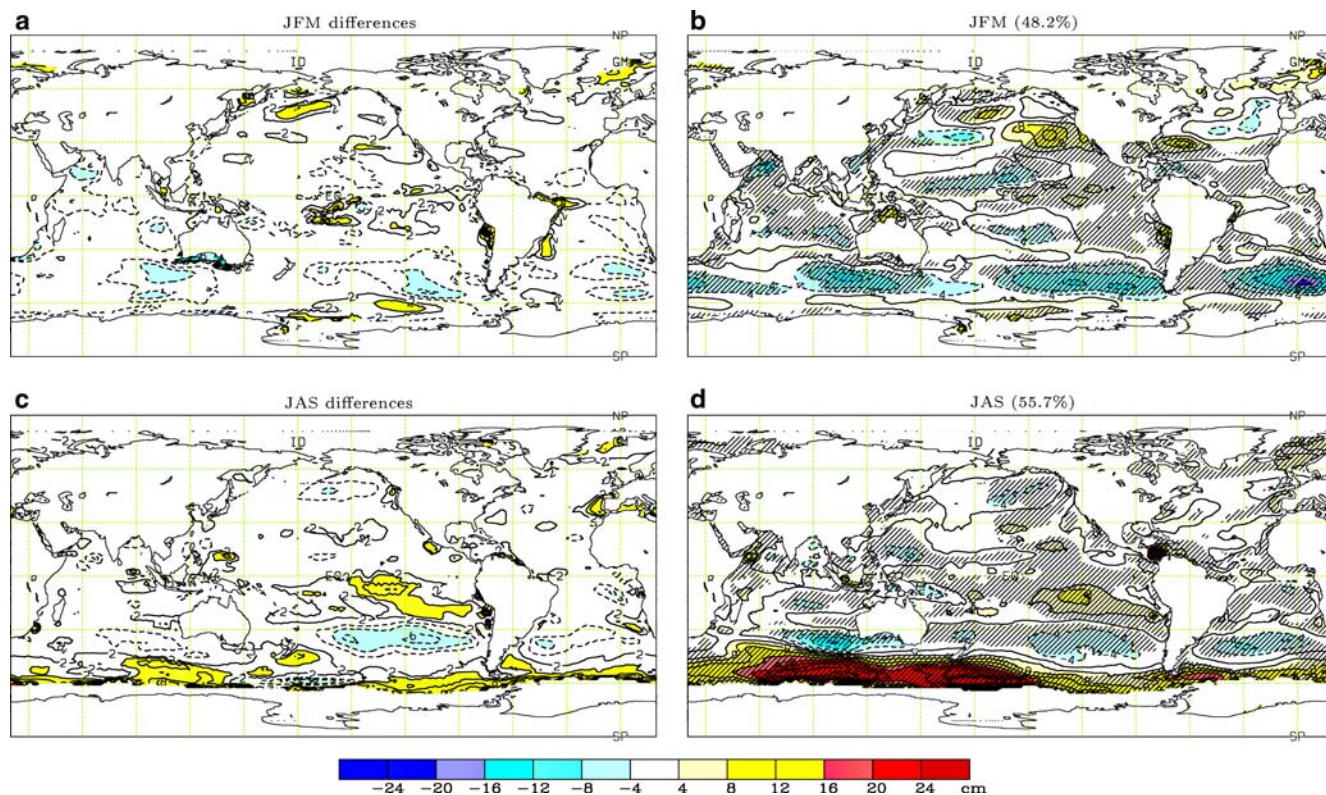
Note that changes similar to those shown in Fig. 3 were also estimated from projections made with the normalization procedure described in Sect. 3.1. The differences in absolute value are smaller than 4 cm in most areas, being largest (up to 12 cm) in the region



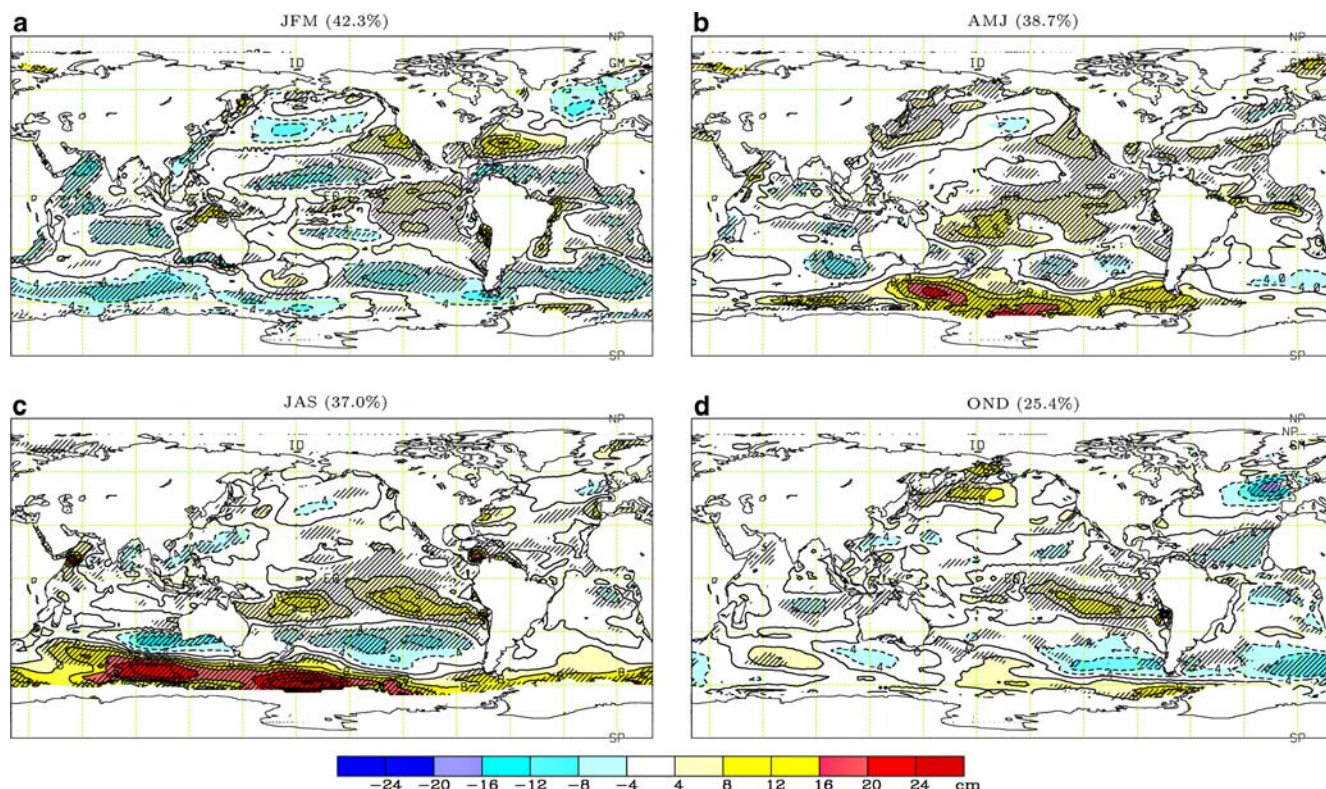
**Fig. 3** Changes (cm) in the indicated seasonal mean SWH from 1990 to 2080 (values of 2080 – values of 1990), as estimated from combining the three climate models' projections with the SRES A2 forcing-scenario. *Dashed and solid lines* indicate negative and non-negative contours, respectively. *Hatching* indicates areas of

significant linear/quadratic trends in the projected seasonal means of SWH (i.e., areas where the null hypothesis of no trend was rejected in favor of model RM<sub>2</sub> at 5% level). The percentages of these areas (i.e., rejection rates) are given in *parentheses* above



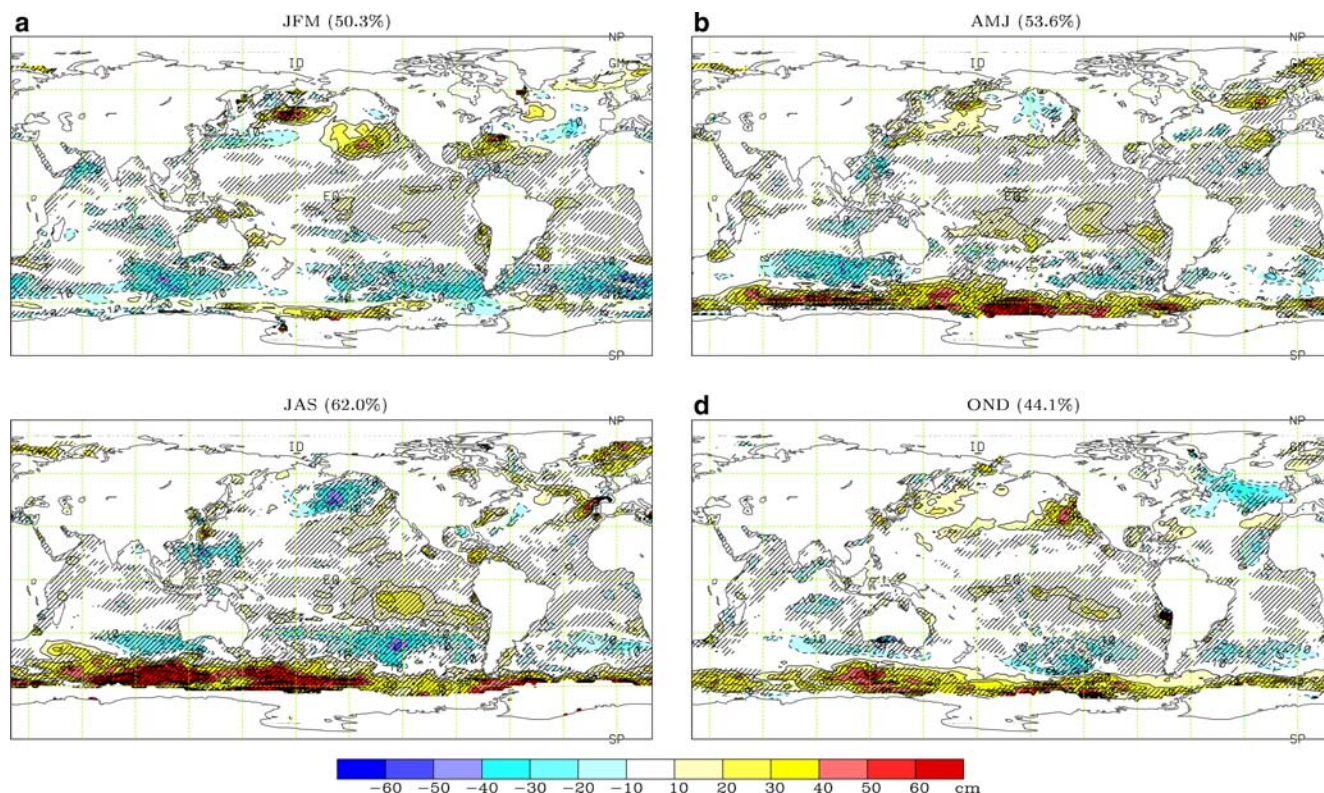


**Fig. 4** a–b Differences between the changes estimated from projections made without and with the normalization procedure described at the end of Sect. 3.1. c–d The same as in Fig. 3 but for changes estimated from projections made with the normalization procedure



**Fig. 5** The same as in Fig. 3 but for the three climate models' B2 scenario projections





**Fig. 6** Changes (cm) in the indicated seasonal 20-year return values of SWH from 1990 to 2080 (values of 2080 – values of 1990), as estimated from combining the three climate models' projections with the SRES A2 forcing-scenario. Zero-contours are not drawn. Solid and dashed lines indicate positive and negative contours,

respectively. Hatching indicates areas of significant linear/quadratic trends in the location parameter of seasonal SWH extremes (i.e., areas where the null hypothesis of no trend was rejected in favor of model RM<sub>2</sub> at 5% level). The percentages of these areas (i.e., rejection rates) are given in parentheses above

nearby the Antarctic coastal zone in all seasons other than JFM (cf. Fig. 4a–b; the other seasons are similar to Fig. 4b and hence not shown). In general, there are few notable differences in the patterns of projected change, although the rejection rates are slightly lower with the normalization procedure (cf. Figs. 3a, c, 4b, d; the other seasons are similar and hence not shown). That is, the possible data discontinuity around year 1979 does not significantly affect the patterns of projected change. This is also the case for the projections of wave height extremes and hence will not be discussed further hereafter.

With the weaker B2 forcing-scenario, as shown in Fig. 5, the projected changes are generally smaller but have patterns similar to those projected with the A2 scenario. The biggest differences between the A2 and the B2 scenarios are seen in the southern hemisphere in the OND season (see Figs. 3d, 5d). We do not know why the A2 and B2 scenarios show different patterns of projected change. However, there is little difference (in terms of pattern of change) between the two scenarios in the regions off both the Atlantic and the Pacific coasts of the United States in boreal winter (see Figs. 3a, 5a).

For the seasonal extremes, as shown in Figs. 6 and 7, the patterns of projected changes are also similar to, but not as smooth as their seasonal mean counterparts. Again, the weaker B2 forcing-scenario is generally

associated with smaller changes than the A2 scenario. The JFM changes are characterized by increases (up to about 50 cm or 7%) in the regions off both the Atlantic and the Pacific coasts of the United States, with the B2 forcing-scenario being associated with slightly larger increases in the region off the Atlantic coast of United States (Fig. 7a). With the A2 forcing-scenario, increases were also projected in the northeast Atlantic in all seasons, being largest in the JAS season (up to about 50 cm or 9% cf. Fig. 6c). In the southern hemisphere, significant increases in seasonal extremes of SWH were projected for the region nearby the Antarctic coastal zone in the AMJ and JAS seasons, with decreases in the zone between 40°S and 60°S (cf. Figs. 6b, c, 7b, c). Such a pattern of change was also projected for the OND season with the A2 forcing-scenario (Fig. 6d).

Overall, the projected changes in wave heights are consistent with the projected changes in extra-tropical storm tracks and cyclone activity. For example, the NP storm track was projected by CGCM2 to “rotate” clock-wise (i.e., to shift southward over the eastern NP and northward over western NP) in winter, and to shift northward in spring (Wang et al. 2004b). These changes in the mean position of the storm track are associated with changes in the occurrence location and frequency of strong cyclones (Wang et al. 2004b). The



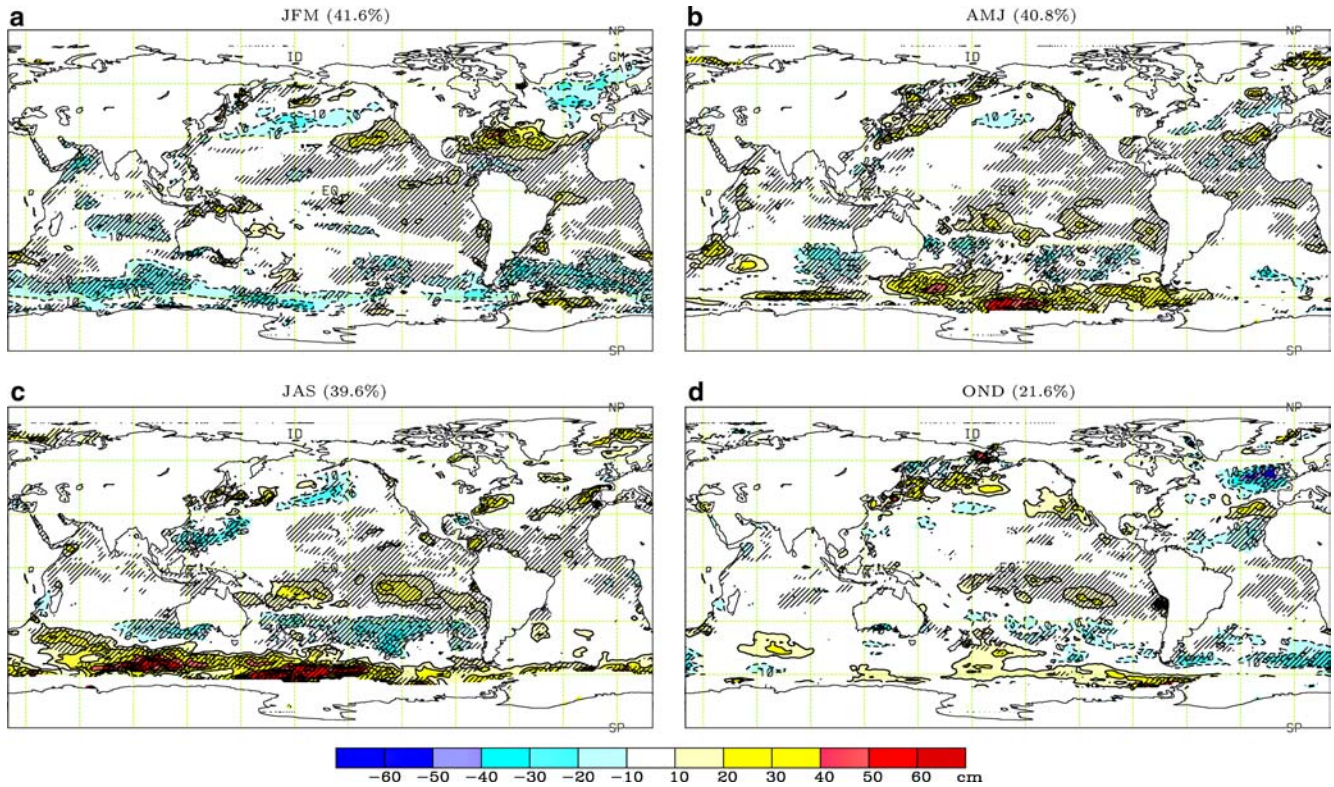


Fig. 7 The same as in Fig. 6 but for the three climate models' B2 scenario projections

areas of more frequent occurrence of strong cyclones were identified to have significant increases in ocean wave heights, and those of less frequent strong cyclone activity, decreases in wave heights. This connection makes sense physically.

### 5 Relative importance of the forcing-induced variance

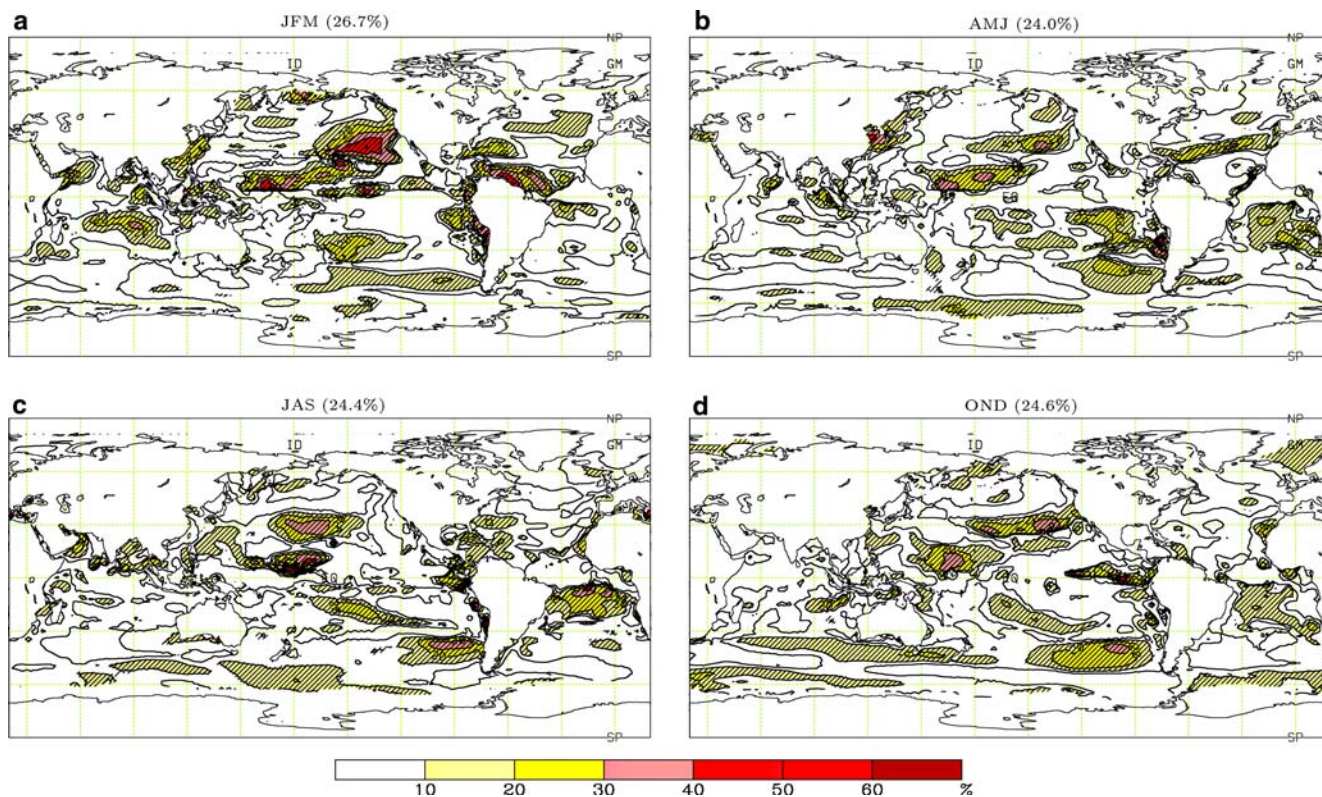
As mention before, for each of the three forcing-scenarios, three integrations were conducted with CGCM2 with different initial conditions. These 3-member ensemble simulations allow us to assess the relative importance of the variability that is due to the prescribed forcing (i.e., forcing-induced variability or climate change signal), by means of a one-factor ANOVA, as described in Sect. 3.2. The one-factor ANOVA was applied to the A2 and B2 ensemble projections of seasonal SWH quantities (means or extremes) for each season, separately. The results are selectively shown in Figs. 8 and 9, in which hatching indicates areas in which the null hypothesis of zero forcing-induced variability (i.e.,  $H_0: \beta_t = 0$  for  $t = 1, 2, \dots, n$ ) was rejected at 5% significance level. The percentages of areas in which  $H_0$  was rejected (rejection rates) are also listed in Table 2.

Figure 8 shows the proportion of the total variance in the CGCM2 projected seasonal mean SWH that is due to the forcing prescribed in the A2 scenario. In boreal winter, the forcing-induced variance is largest and most

significant in the mid-latitudes of eastern NP, and in the tropics of NP and NA (Fig. 8a). In the other three seasons, the largest proportions of forcing-induced variance are also seen in the NP, with the center of high variance-proportions shifting westward to the central NP (Figs. 8b–d). The forcing-induced variance was also identified to be statistically significant in the zone between 40°S and 60°S and in the tropical South Atlantic (SA) in all seasons, and also in the northeast Atlantic in the OND season (Fig. 8d). The rejection rates range between 24% and 27% (cf. Table 2), which are very likely of 5% field significance because the projected wave height fields very likely have 15 or more spatial degrees of freedom (see the ad hoc argument given in the second paragraph of Sect. 4).

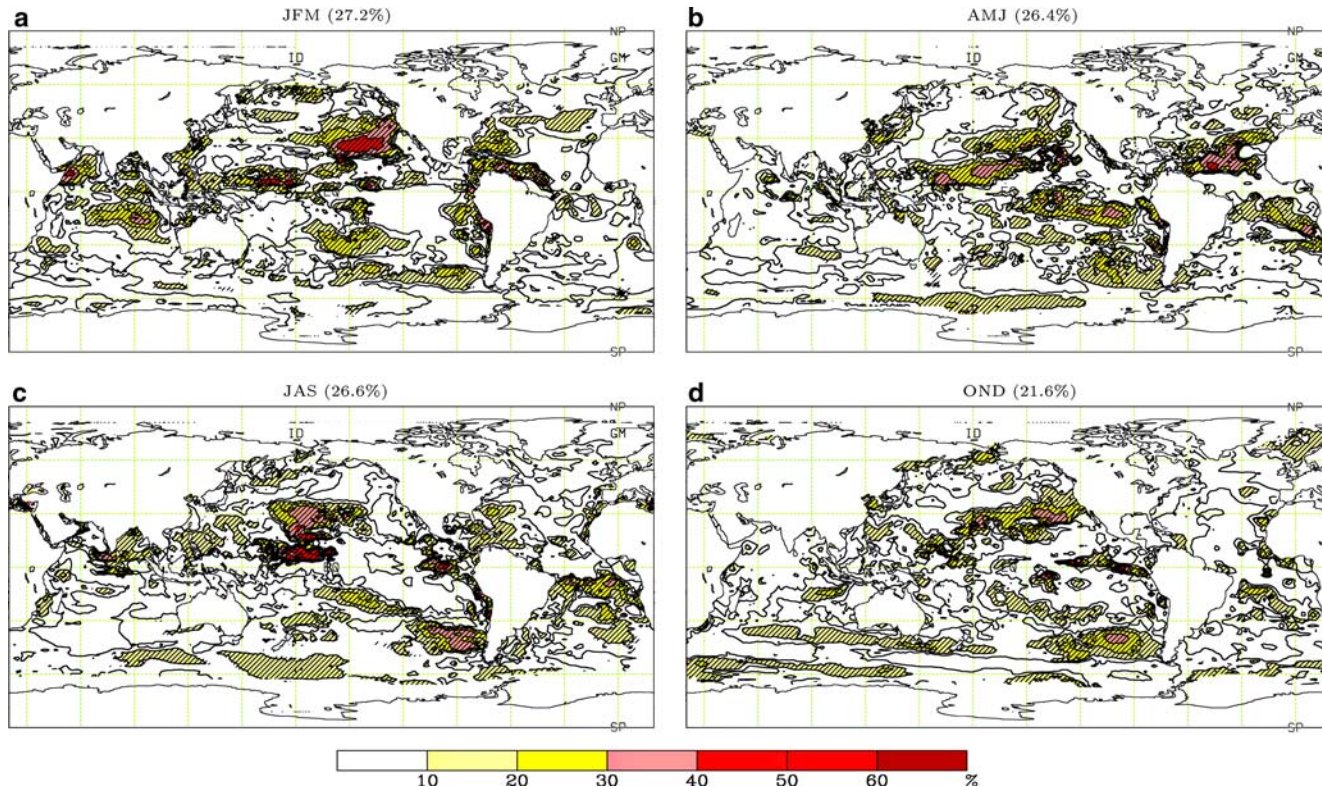
Note that, in the region nearby the Antarctic coastal zone, the area of significant forcing-induced variance does not correspond well with the area of large increases shown in Fig. 3. The discrepancy is largely due to the inter-model variability, because we see much better correspondence if we replace Fig. 3 with the corresponding changes estimated from CGCM2 simulations (instead of combining the three climate models' projections, as is the case in Fig. 3). In this region, CGCM2 projects significant increases of up to about 10, 25, and 16 cm in JFM, AMJ, and JAS, respectively, with no significant change in OND (not shown). These changes are much smaller than the three model mean projections of change shown in Fig. 3.





**Fig. 8** The proportion of the total variance in the CGCM2 projected seasonal mean SWH that is due to the forcing prescribed in the A2 scenario. The contour interval is 10%. Hatching indicates areas of statistically significant forcing-induced variance (i.e., areas

where the null hypothesis of zero forcing-induced variance is rejected at 5% level). The percentages of these areas (i.e., rejection rates) are given in *parentheses* above



**Fig. 9** The same as in Fig. 8 but for the proportion of the total variance in the CGCM2 projected seasonal 20-year return values of SWH that is due to the forcing prescribed in the A2 scenario



**Table 2** The percentages of the area (i.e., rejection rates) in which the indicated null hypothesis ( $H_\beta$  or  $H_\gamma$  or  $H_\theta$  or  $H_\delta$  or  $H_{\gamma+\theta+\delta}$ ; see Sect. 3.2) was rejected at 5% significance level

	Seasonal means				20-year return values			
	JFM	AMJ	JAS	OND	JFM	AMJ	JAS	OND
$H_\beta$ (A2)	26.7	24.0	24.4	24.6	27.2	26.4	26.6	21.6
$H_\beta$ (B2)	16.4	9.9	9.4	8.2	16.4	12.2	10.4	7.9
$H_\gamma$	87.3	81.9	89.7	86.5	87.8	91.2	93.9	90.1
$H_\theta$	46.7	42.9	57.2	66.2	49.0	61.0	72.2	77.6
$H_\delta$	63.2	57.6	66.2	77.0	63.3	76.0	82.3	84.2
$H_{\gamma+\theta+\delta}$	90.1	86.4	92.8	93.9	91.2	94.5	96.7	97.1

These measure the field significance (Livezey and Chen 1983) of the related model and/or scenario uncertainties (see the last paragraph of Sect. 5 for more details)

For the weaker B2 scenario, patterns of the forcing-induced variance proportions are similar to those shown in Fig. 8 and hence are not shown in this paper, but the proportions are generally smaller (e.g., about a half of those shown in Fig. 8a in the mid-latitudes of eastern NP), and the areas of significant forcing-induced variance are much less extensive (see Table 2).

For the projections of seasonal extremes, the patterns of forcing-induced variance and variance-proportions are similar to those identified in the projections of seasonal mean SWH (cf. Figs. 8, 9). The largest differences between the seasonal means and extremes are seen in the AMJ season in the subtropics of NP and NA, where the forcing appears to induce remarkably larger variance in the seasonal extremes than in the seasonal means (see Figs. 8b, 9b). At the 5% significance level, the rates of rejecting the null hypothesis  $H_\beta$  are also similar: 22–27% for the A2 scenario, and 8–16% for the B2 scenario (cf. Table 2). According to Livezey and Cheng (1983), a rejection rate of 8.2% (10%) would indicate a field significance at the 5% level in a field of 160 (80) or more spatial degrees of freedom. The CGCM2 projected global wave height data is very likely to have more than 80 but less than 160 spatial degrees of freedom, because its first leading 80 (120) EOFs represent about 90% (95%) of its total variance. Thus, rejection rates of 10% or higher very likely indicate a field significance at the 5% level, and rejection rates of 8.2% or lower, no field significance. Very likely, the changes projected for the B2 scenario are field significant only in the JFM and AMJ seasons, but insignificant in the OND season; while those projected for the A2 scenario are of field significance in all seasons (Table 2).

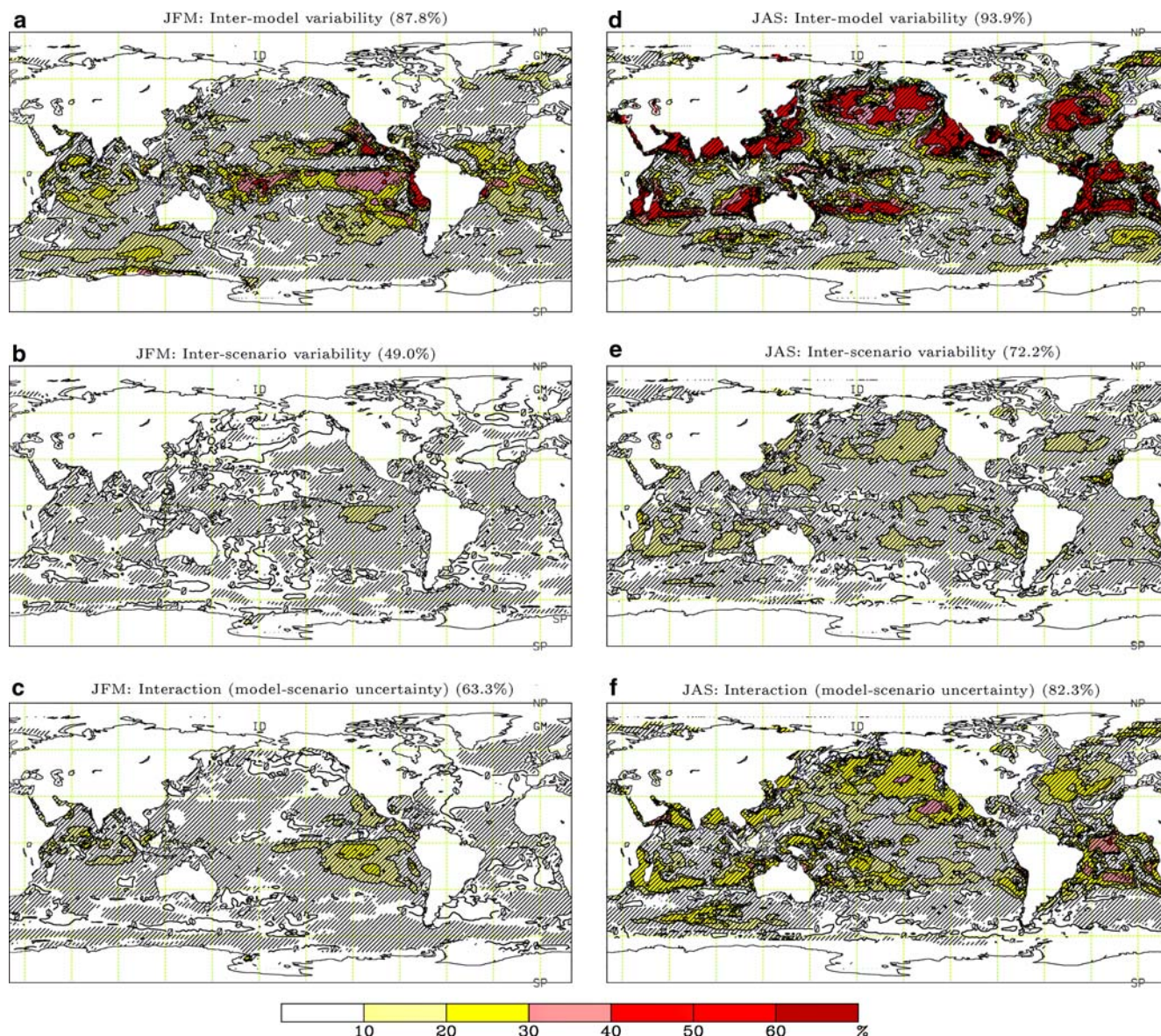
## 6 Characteristics of the model and scenario uncertainties

Remember that, for each of the three forcing-scenarios, one simulation was performed with each of the three climate models. Thus, there are a total of 9 projections of seasonal mean (or extreme) SWH for the period of 1990–2049 (because the ECHAM4/OPYC3 IS92a projection covers only up to 2049). These projections can be classified into a 3-by-3 table, according to the three cli-

mate models (the “model factor”) and the three forcing-scenarios (the “forcing or scenario factor”). This allows for the two-factor fixed effects ANOVA described in Sect. 3.2, which was carried out for each of the four seasons, and for the seasonal means and extremes of SWH, separately. As a result, the null hypothesis that there are no effects of inter-model variability or inter-scenario variability or model-scenario uncertainties (i.e.,  $H_\gamma$  or  $H_\theta$  or  $H_\delta$  or  $H_{\gamma+\theta+\delta}$ ; see Sects. 3.2, Appendix) are tested, with the rejection rates shown in Table 2. The proportions of the total variance in the projected SWH quantities (means or extremes) that is due to either the inter-model or the inter-scenario variability or both of them, as well as their statistical significance, were also estimated and selectively shown in Figs. 10, 11.

Figure 10 shows the proportion of the total variance in the projected JFM and JAS seasonal 20-year return values of SWH that is due to the three sources of uncertainty: (1) uncertainty due to differences among the climate models (i.e., inter-model variability); (2) uncertainty due to the differences among the forcing-scenarios (i.e., inter-scenario variability); (3) uncertainty due to different model sensitivities to differences in forcing conditions (i.e., interaction or model-scenario uncertainty). Clearly, among the three sources of uncertainty, the inter-model variability is the largest, accounting for up to 50% of the total variance in the tropical Pacific; it dominates the pattern of the “total” uncertainty (i.e., the sum of the above three sources of uncertainty; see Fig. 11a). The inter-scenario variability is much smaller. Its variance proportions rarely exceed 10%. The relatively small inter-scenario variability is not surprising, because the differences among the three forcing-scenarios used here are not big (the IS92a scenario is similar to the A2 scenario, only is the B2 scenario weaker; see Flato and Boer 2001 or Nakicenovic and Swart 2000). It should be pointed out that the inter-scenario variability is statistically significant in 49–78% of the area of wave data (Table 2), although it is relatively small in general. This means that different forcing conditions do make significant differences in the projections of ocean wave heights.

In the other two seasons, the situation is similar. The uncertainty due to differences among the three climate models (the model uncertainty) is much larger than the uncertainty due to differences among the three forcing-



**Fig. 10** The proportion of the total variance in the projected JFM and JAS seasonal 20-year return values of SWH that is due to the effect of the indicated sources of uncertainty. The contour interval is 10%. Hatching indicates areas of statistically significant

uncertainty (i.e., areas where the null hypothesis of zero inter-model or inter-scenario or interaction effects is rejected at 5% level). The percentages of these areas (i.e., rejection rates) are given in *parentheses* above

scenarios (the scenario uncertainty). Thus, only the “total” uncertainty is shown in Fig. 11. Clearly, among the four seasons, the uncertainty is smaller in boreal winter and spring than in the other two seasons. And it is generally small in the middle-high latitudes, especially in boreal winter and spring; but it is large in the tropics (Fig. 11; here small or large is relative to the effects of internal variability and the common forcing).

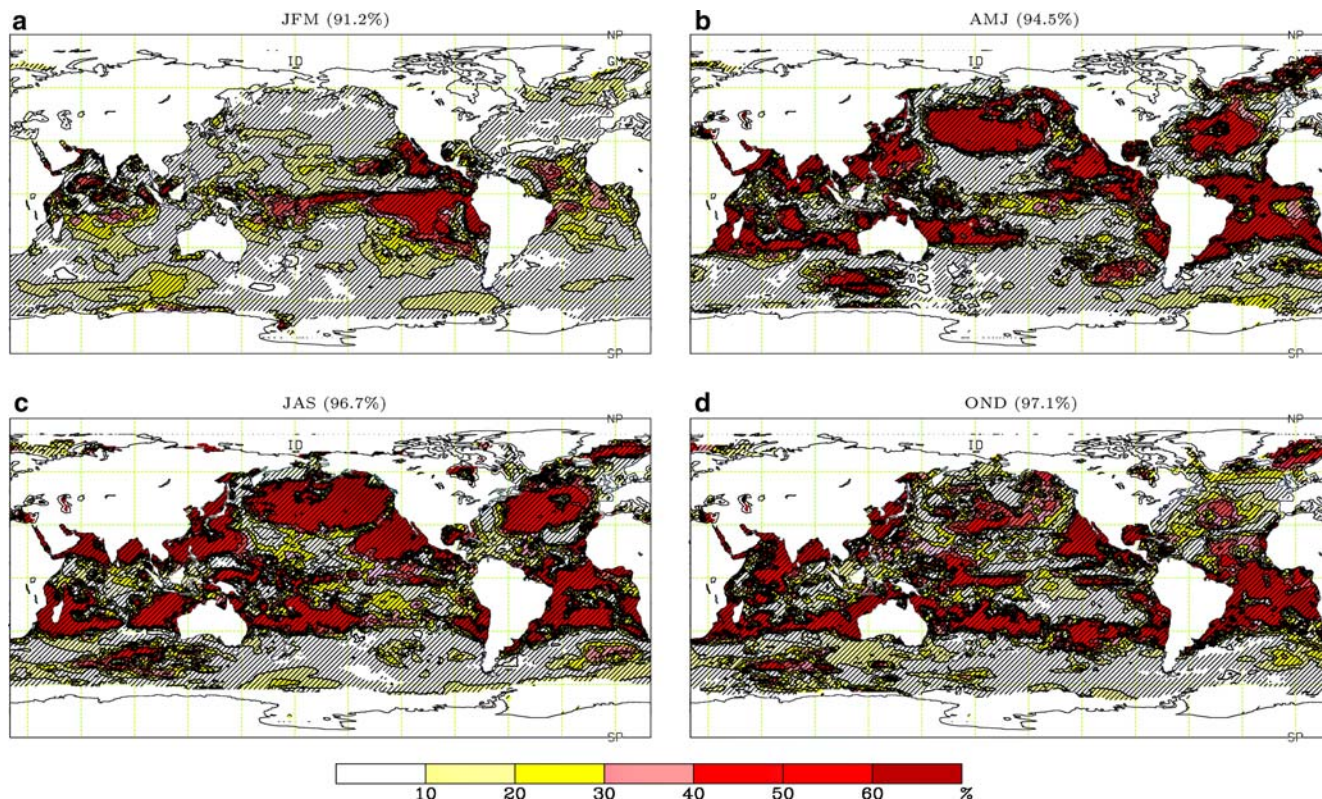
For the projections of seasonal mean SWH, the three sources of uncertainty, and the sum of them, have characteristics that are quite similar to those shown in Figs. 10 and 11, in terms of both the pattern and the magnitude (and hence are not shown in this paper). Again, the model uncertainty is larger than the scenario

uncertainty. The inter-scenario variability is also statistically significant, although it is much smaller than the inter-model variability (especially in the tropics).

## 7 Summary and discussions

In this study, we have made and analyzed projections of seasonal means and extremes of ocean wave heights using projections of possible future climates conducted with three global climate models for three forcing-scenarios. We have estimated the multi-model mean climate change by combining the three climate models’ projections for the same forcing-scenario. The relative





**Fig. 11** The proportion of the total variance in the projected seasonal 20-year return values of SWH that is due to the effects of forcing and model uncertainties. The contour interval is 10%.

Hatching indicates areas of statistically significant (at 5% level) model and/or scenario uncertainties. The percentages of these areas (i.e., rejection rates) are given in *parentheses* above

importance of the variability in the projected wave heights that is due to the forcing prescribed in the A2 or B2 scenario was assessed on the basis of the CGCM2 ensemble simulations. We have also characterized the uncertainty in the wave height climate change projections that is due to differences among the three climate models and/or among the three forcing-scenarios.

The results show that the multi-model mean projections of climate change has patterns similar to those derived from using the CGCM2 projections alone, but the magnitudes of changes are generally smaller in boreal oceans and larger in the region nearby the Antarctic coastal zone. The forcing-induced variance was identified to be statistically significant in some areas in all seasons, being largest in the NP in boreal winter with the A2 (or IS92a; not shown) forcing-scenario.

It has also been shown that, in the projections of wave height climate change, the uncertainty due to differences among the three climate models is much larger than that due to differences among the three forcing-scenarios. The small scenario uncertainty is not surprising because the IS92a scenario is similar to the A2 scenario. The sum of the model and forcing-scenario related uncertainties is smaller in the JFM and AMJ seasons than in the JAS and OND seasons, and it is generally small in the mid-high latitudes and large in the tropics (relative to the effects of internal variability and the common forcing). In particular, the areas of large

projected changes (e.g. the mid-latitudes of eastern NP) were identified to have inter-model variability that is small in terms of variance proportion (cf. Figs. 8a, 11a). In other words, all the three climate models projected large changes in these same areas, so the confidence of the projections for these areas should be higher. Also, the inter-scenario variability was identified to be statistically significant in most areas of the oceans, although it is small relative to the inter-model variability. This indicates that different forcing conditions do make significant differences in the wave height climate change projection.

It should be pointed out that the model uncertainty presented in this study is limited to the three climate models used. Its characteristics could change if other global climate models were used or added in this study. Similarly, the inter-scenario variability is also limited to the three forcing scenarios analyzed.

Also, there are other sources of uncertainty that are not discussed in the present study. For example, the statistical downscaling completed in this study takes into account only projected changes in SLP and thus might well be associated with the problem of imperfect knowledge and/or representation of the physical processes (although variations in atmospheric pressure may somehow reflect changes in other variables such as sea ice cover). This problem is common to all statistical downscaling approaches, in addition to their common

assumption that the statistical relationship will hold under the climate model projected future climate conditions. It adds uncertainty in the projections, which belongs to the third level of uncertainty. However, we did not discuss the third level of uncertainty, i.e., the uncertainty due to different approaches taken to generate “regional” scale climate change information from global climate model simulations, such as the use of different regional climate models (RCMs), or dynamical versus statistical downscaling approaches, or the use of different predictor variables in a statistical downscaling model, or the use of GEV versus generalized pareto distribution (GPD) models for making projections of extremes, and so on. In this regard, the GEV approach has been compared with a non-stationary GPD approach (Caires et al. 2005).

In addition, the quality of the simulated SLP and wave heights (especially extremes) in ERA-40 could also have an impact on the resulting projections. According to Caires and Sterl (2005), the two main limitations of the ERA-40 significant wave height data are the existences of inhomogeneities and underestimation of the wave height values that discourages the use of the data in design studies. They have thus attempted to improve the ERA-40 wave height fields. The use of their corrected ERA-40 wave data could probably improve the accuracy of wave height climate change projections. Differences between the projections made using the raw ERA-40 wave data (as in this study) and those made using the corrected ERA-40 data represent another source of uncertainty, which also belongs to the third level of uncertainty and should be investigated in the future.

**Acknowledgments** The authors are greatly indebted to Qiuqi Wen and Yang Feng for their computing support. The authors thank Dr. Francis Zwiers for his help in clarifying questions about ANOVA techniques. Drs. Sofia Caires and Vichoslav Kharin are acknowledged for their helpful comments on an earlier version of this manuscript. All anonymous reviewers are also acknowledged for their helpful comments.

## 8 Appendix: analysis of variance

According to the one-factor ANOVA model (Eq. 4) (see Sect. 3.2), the total sum of squares of variable  $Y_{ts}$ :

$$SS_Y = \sum_t \sum_s (Y_{ts} - Y_{oo})^2 \quad (6)$$

can be partitioned into two statistically independent variance components (Huitson 1966):

$$SS_Y = SS_\beta + SS_\epsilon, \quad (7)$$

where

$$SS_\beta = S \sum_t (Y_{to} - Y_{oo})^2 = S \sum_t \beta_t^2 + (n-1)\hat{\sigma}_\epsilon^2,$$

$$SS_\epsilon = \sum_t \sum_s (Y_{ts} - Y_{to})^2 = n(S-1)\hat{\sigma}_\epsilon^2,$$

and a “o” is used to replace a subscript when an arithmetic average is taken over that index (this notation is used throughout this section). The null hypothesis

$$H_\beta : \beta_t = 0 \text{ for } t = 1, 2, \dots, n$$

can be tested by comparing

$$F_\beta = \frac{SS_\beta / (n-1)}{SS_\epsilon / [n(S-1)]} \quad (8)$$

with the critical value  $F_p[(n-1), n(S-1)]$  where  $F_p[v_1, v_2]$  denotes the  $p$ -quantile of the  $F$  distribution with  $v_1$  and  $v_2$  degrees of freedom. And the proportion of the total variance in  $Y_{ts}$  that is due to the effects of factor  $\beta_t$  is estimated as

$$P_\beta = \frac{SS_\beta - \frac{n-1}{n(S-1)} SS_\epsilon}{SS_Y}. \quad (9)$$

Similarly, according to the two-factor fixed effect ANOVA model (Eq. 5) (see Sect. 3.2), the total sum of squares of variable  $X_{ijt}$ :

$$SS_X = \sum_i \sum_j \sum_t (X_{ijt} - X_{ooo})^2 \quad (10)$$

can be partitioned into four statistically independent variance components (Huitson 1966):

$$SS_X = SS_\gamma + SS_\theta + SS_\delta + SS_\epsilon, \quad (11)$$

where

$$SS_\gamma = nq \sum_i (X_{ioo} - X_{ooo})^2 = nq \sum_i \hat{\gamma}_i^2 + (m-1)\hat{\sigma}_\epsilon^2,$$

$$SS_\theta = nm \sum_j (X_{ojo} - X_{ooo})^2 = nm \sum_j \hat{\theta}_j^2 + (q-1)\hat{\sigma}_\epsilon^2,$$

$$\begin{aligned} SS_\delta &= n \sum_i \sum_j (X_{ijo} - X_{ojo} - X_{ioo} + X_{ooo})^2 \\ &= n \sum_i \sum_j \hat{\delta}_{ij}^2 + (m-1)(q-1)\hat{\sigma}_\epsilon^2, \end{aligned}$$

$$SS_\epsilon = \sum_i \sum_j \sum_t (X_{ijt} - X_{ijo})^2 = mq(n-1)\hat{\sigma}_\epsilon^2.$$

Therefore, the null hypothesis

$$H_\gamma : \gamma_i = 0 \text{ for } i = 1, 2, \dots, m$$

can be tested by comparing

$$F_\gamma = \frac{SS_\gamma / (m-1)}{SS_\epsilon / [mq(n-1)]} \quad (12)$$

with the critical value  $F_p[(m-1), mq(n-1)]$ . And an unbiased estimator of the proportion of the total variance in  $X_{ijt}$  that is due to the effects of factor  $\gamma_i$  is

$$P_\gamma = \frac{SS_\gamma - \frac{m-1}{mq(n-1)} SS_\epsilon}{SS_X}. \quad (13)$$



Similarly, the null hypothesis

$$H_0 : \theta_j = 0 \text{ for } j = 1, 2, \dots, q$$

can be tested by comparing

$$F_\theta = \frac{SS_\theta / (q - 1)}{SS_\varepsilon / [mq(n - 1)]} \quad (14)$$

with the critical value  $F_p[(q-1), mq(n-1)]$ . And the proportion of the total variance in  $X_{ijt}$  that is due to the effects of factor  $\theta_j$  is estimated as

$$P_\theta = \frac{SS_\theta - \frac{q-1}{mq(n-1)} SS_\varepsilon}{SS_X}. \quad (15)$$

The null hypothesis

$$H_\delta : \delta_{ij} = 0 \text{ for } i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, q$$

can be tested by comparing

$$F_\delta = \frac{SS_\delta / [(m - 1)(q - 1)]}{SS_\varepsilon / [mq(n - 1)]} \quad (16)$$

with the critical value  $F_p[(m-1)(q-1), mq(n-1)]$ . And the proportion of the total variance in  $X_{ijt}$  that is due to the interaction effects  $\delta_{ij}$  is estimated as

$$P_\delta = \frac{SS_\delta - \frac{(m-1)(q-1)}{mq(n-1)} SS_\varepsilon}{SS_X}. \quad (17)$$

Also, the null hypothesis

$$H_{\gamma+\theta+\delta} : \gamma_i = \theta_j = \delta_{ij} = 0 \text{ for } i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, q$$

can be tested by comparing

$$F_{\gamma+\theta+\delta} = \frac{(SS_\gamma + SS_\theta + SS_\delta) / (mq - 1)}{SS_\varepsilon / [mq(n - 1)]} \quad (18)$$

with the critical value  $F_p[(mq-1), mq(n-1)]$ . And the proportion of the total variance in quantity  $X_{ijt}$  that is due to the total effects of all factors (i.e.,  $\gamma_i$ ,  $\theta_j$ , and  $\delta_{ij}$ ) is estimated as

$$P_{\gamma+\theta+\delta} = \frac{(SS_\gamma + SS_\theta + SS_\delta) - \frac{(mq-1)}{mq(n-1)} SS_\varepsilon}{SS_X}. \quad (19)$$

## References

- 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) Inter-comparison of different wind-wave reanalyses. J Clim 17:1893–1913
- Caires S, Swail VR, Wang XLL (2005) Projection and analysis of extreme wave climate. J Clim (accepted)
- Flato GM, Boer GJ (2001) Warming asymmetry in climate change simulations. Geophys Res Lett 28:195–198
- Gibson R, Kallberg P, Uppala S (1996) The ECMWF reanalysis (ERA) project. ECMWF Newlett 73:7–17
- Gordon C, Cooper C, Senior CA, Banks H, Gregory JM, Johns TC, Mitchell JFB, Wood RA (2000) The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. Clim Dyn 16:147–168
- Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds) (2001) Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, p 881
- Huitson A (1966) The analysis of variance. Charles Griffin & Company Ltd., London, p 83
- Johnson RA, Wichern DW (1982) Applied multivariate statistical analysis. Prentice-Hall Inc, Englewood Cliffs, p 594
- Kushnir Y, Cardone VJ, Greenwood JG, Cane MA (1997) The recent increase in North Atlantic wave heights. J Clim 10:2107–2113
- Livezey RE, Chen WY (1983) Statistical field significance and its determination by Monte Carlo techniques. Mon Weather Rev 111:46–59
- Nakicenovic N, Swart R (eds) (2000) Special report on emissions scenarios (SRES). Cambridge University Press, Cambridge, p 599
- Roeckner E, Arpe K, Bengtsson L, Christoph M, Claussen M, Dümenil M, Esch M, Giorgetta M, Schlese U, Schulzweida U (1996a) The atmospheric general circulation model ECHAM-4: model description and simulation of present-day climate. Technical report 218, Max-Planck-Institut für Meteorologie, Hamburg, Germany, p 90
- Roeckner E, Oberhuber JM, Bacher A, Christoph M, Kirchner I (1996b) ENSO variability and atmospheric response in a global coupled atmosphere-ocean GCM. Clim Dyn 12:737–754
- Simmons AJ, Gibson JK (2000) The ERA-40 project plan. ERA-40 Project Report Series No 1 (Available online at <http://www.ecmwf.int/research/era/>)
- Sterl A (2004) On the (In)homogeneity of reanalysis products. J Clim 17:3866–3873
- Uppala S (2001) ECMWF ReAnalysis, 1957–2001. In: Proceedings of the ECMWF workshop on reanalysis, ERA-40 Project report series No 3: 1–10, Reading, UK, pp 5–9
- von Storch H, Navarra A (1995) Analysis of climate variability—applications of statistical techniques. Springer, Berlin Heidelberg New York, p 334
- von Storch H, Reichardt H (1997) A scenario of storm surge statistics for the German Bight at the expected time of doubled atmospheric carbon dioxide concentration. J Clim 10:2653–2662
- von Storch H, Zwiers FW (1999) Statistical analysis in climate research. Cambridge University Press, Cambridge, p 484
- 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 (2005) Historical and possible future changes of wave heights in northern hemisphere oceans. In: Perrie W (eds) Atmosphere ocean interactions, Vol 2. (Advances in Fluid Mechanics Series Vol 39) Wessex Institute of Technology Press (in press)
- Wang XLL, Zwiers FW (1999) Interannual variability of precipitation in an ensemble of AMIP climate simulations conducted with the CCC GCM2. J Clim 12:1322–1335
- Wang XLL, Zwiers FW, Swail VR (2004a) North Atlantic ocean wave climate change scenarios for the twenty-first century. J Clim 17:2368–2383
- Wang XLL, Swail VR, Zwiers FW (2004b) Climatology and changes of extra-tropical storm tracks and cyclone activities as derived from two global reanalyses and the Canadian CGCM2

- projections of future climate. In: Preprints of the eighth international workshop on wave forecast and hindcast, 14–19 November 2004, North Shore, Hawaii
- Wang XLL, Swail VR, Zwiers FW (2005) Climatology and changes of extra-tropical cyclone activity: Comparison of ERA-40 with NCEP/NCAR Reanalysis for 1958–2001. *J Clim* (accepted)
- WASA Group (1998) Changing waves and storms in the Northeast Atlantic? *Bull Am Meteorol Soc* 79:741–760
- Zhang XB, Zwiers FW (2004) Comments on “Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test” by Sheng Yue and Chun Yuan Wang. *Water Resources Res* 40: W03805, DOI:10.1029/2003WR002073
- Zwiers FW, Wang XLL, Sheng J (2000) Effects of specifying bottom boundary conditions in an ensemble of atmospheric GCM simulations. *J Geophys Res* 105:7295–7315