Is the 100-year significant wave height increasing in the Tasman Sea?

Kevin Ewans[a], Philip Jonathan[c,b,]

[a]*MetOcean Research Ltd., New Plymouth, New Zealand*
[b]*Department of Infrastructure Engineering, University of Melbourne, Victoria 3010, Australia*
[c]*Shell Research Limited, London SE1 7NA, United Kingdom.*
[d]*Department of Mathematics and Statistics, Lancaster University LA1 4YF, United Kingdom.*

Abstract

Climate change is frequently reported to be the cause of an apparent increase in the occurrence and intensity of extreme events around the globe. Recent studies have found that increases can be expected in sea state extremes for some regions of the world's oceans. We investigate temporal trends in storm peak significant wave height for a central Tasman Sea location, using output from the FIO-ESM v2.0 CMIP6 Earth System Model of Bao et al. (2020), Song et al. (2020). These data include time-series of significant wave height for several Shared Socioeconomic Pathways (SSP) as well as 165-year historical and 700-year pre-industrial realisations. This allows estimates to be made of 100-year return values for significant wave height at the end of the 21st century, for several future atmospheric forcing scenarios, and comparison of these against estimates for present day, based on the historical data base, and estimates for pre-industrial conditions. The estimates are made with a non-stationary extreme value analysis method that allows return values of storm peak significant wave height to be estimated as a function of time and to quantify if changes in return values are statistically significant. Evidence supporting climate-related changes at the Tasman location studied is weak. Estimates of most probable changes in 100-year return value for storm peak significant wave height between 2015 and 2100, for the three SSP scenarios considered, suggest the assumption of a 2m increase in 100-year return value to be reasonable, whilst noting that this value is small compared to the inherent uncertainty (of at least ±5m) present.

1. Introduction

The effect of climate change on the reliability of marine structures is of considerable concern, yet estimating future extreme ocean environments is problematic. There is agreement on some effects, such as an increasing proportion of intense tropical cyclones (Category 4–5) and peak wind speeds of the most intense tropical cyclones globally with increasing global warming (IPCC 2021). However, uncertainties in specifying a realistic description of climate forcing, and the variability of climate model output notionally corresponding to the same climate forcing, are large. Estimating design conditions is a challenging task even when climate change effects are ignored, due generally to small sample sizes for extreme value model fitting, the effects of covariates, and the assumption of asymptotic model forms (e.g. Jonathan and Ewans 2013). In the presence of climate change, estimating design conditions is more challenging still. It is reasonable to expect that estimates of future extreme metocean conditions will be more uncertain as a result of climate change.

It is likely that typical samples of historical metocean data from measurements and hindcasts incorporate the effects of long-term (e.g. inter-annual and multi-decadal) atmospheric oscillations, and that existing metocean databases spanning several decades are in reality not stationary: 100-year return values for significant wave height (H_S) estimated from data representing different decades might themselves be different. In order to assess the strength of evidence for climate-driven changes in return value, better understanding of the inherent temporal variability of historical data is also necessary.

Estimates of future extreme sea states can be made from the output of General Circulation Models (GCMs) run under different climate scenarios at relatively coarse spatial and temporal resolutions. Until recently, these have not included estimates of sea state parameters. Predictions of wave fields involves either statistical or dynamical downscaling (e.g. Ewans and Jonathan 2020). Dynamical downscaling involves running numerical wave prediction models using the GCM data as boundary conditions, providing estimates of H_S and other sea-state variables in space and time at typical hindcast resolutions. We examine the change in 100-year return value for storm peak H_S (Section 3) over a fixed period of $P = 86$ years, using data sets produced from numerical wave models forced by GCMs. We focus on output from the CMIP6 FIO-ESM v2.0 model for a neighbourhood of five locations in the Tasman Sea, at 3 hour temporal resolution. The data considered consist of a quasi-equilibrium 700-year pre-industrial (piControl) set, a Historical set (of length 165 years), and three future scenario sets (for scenarios SSP126, SSP245, and SSP585,

each of length 86 years); see Section 2 for more information. For a given location and scenario, values of storm peak H_S (referred to henceforth as H_S^{sp}), corresponding to the most intense sea state of individual storm events (assumed independent) are isolated from time-series of H_S (Section 2). Return values for H_S^{sp} are then estimated using extreme value (EV) analysis of peaks over threshold (POT). A non-stationary EV model is assumed, with all parameters varying linearly in time. Parameter estimation is achieved using Bayesian inference (Section 3).

Estimating a non-stationary EV model over a long period of time is statistically more efficient that estimating a series of stationary models over shorter intervals. For example, Vanem (2015) demonstrates that statistically significant changes in extreme waves can be identified using a non-stationary analysis of historical and future periods together, but not when stationary models are fitted to each period separately and compared. Numerous authors have commented on the advantages of spatial aggregation of estimates to mitigate arbitrary effects in individual GCM runs. For example, McSweeney and Jones (2013) shows that spatial smoothing of model projections can provide more informative estimates for a neighbourhood than estimates for single locations in the neighbourhood. The choice of EV threshold level for analysis of peaks over threshold is also typically a large source of uncertainty in return values. For this reason, we also ensure that spatial variation of estimates of return value, and their variation with EV threshold level, are not large. Given this, we generally quantify our findings below in terms of model aggregate estimates of changes in 100-year return value over all spatial locations and EV threshold levels considered.

The long piControl output allows us to assess the inherent uncertainty of changes in 100-year return value carefully, based on EV modelling of multiple samples of length 86 years drawn from the 700 years of piControl output. The Historical output allows comparison of the change in 100-year return value over 86 years from the start to the end of the Historical period. The SSP output provides a means to estimate the change in 100-year return value of H_S^{sp} over a period of 86 years, for three scenarios of different severity, and hence potentially a basis for establishing future offshore design criteria.

Ewans and Jonathan (2023) provides a detailed analysis of trends in the 100-year return value of H_S^{sp} from FIO-ESM v2.0 including piControl, and from a number of CMIP5-driven models, for a variety of climate scenarios, for locations east of Madagascar and south of Australia. They provide a brief review of recent literature concerning the effect of climate change on the extreme ocean environment, noting the contributions of Young and Ribal (2019), Meucci et al. (2020), Morim et al. (2020), Meucci et al. (2022), Albuquerque et al. (2022), Casas-Prat et al. (2022) and Sardana et al. (2022).

Objective and outline of article

The objective of this article is to quantify the change in 100-year return value for H_S^{sp} over a period of 86 years at the specified neighbourhood in the Tasman Sea, based on FIO-ESM v2.0 piControl, Historical and SSP output. The quantitative assessment of the inherent uncertainty of change in return value corresponding to no climate forcing is essential in providing a "steady state" or "null distribution", with which to frame our assessment of changes in return value under the Historical and SSP forcing scenarios.

The layout of the article is as follows. Section 2 introduces the data, and Section 3 outlines the non-stationary EV methods used, and the method of estimation of (change in) 100-year return value. Section 4 provides illustrations of our findings, in terms of the inherent (piControl) uncertainty in change in return value (Section 4.1), the change in return value seen in the Historical output (Section 4.2), and the change in return value seen for the three SSP scenarios (Section 4.3). Section 5 provides a brief discussion and conclusions.

2. Data

The data considered are taken from output of the FIO-ESM v2.0 model (Bao et al. 2020, Song et al. 2020), and consist of values for H_S^{sp} at a star configuration of five locations in the central Tasman Sea between Australia and New Zealand, shown in Figure 1. The locations are referred to as Centre (C), North (N), South (S), East (E) and West (W). At each location, data for three CMIP6 experiments are considered (Table 2 of Song et al. 2020) : a 700-year pre-industrial period (piControl: nominal years 301-1000), a 165-year historical period (years 1850-2014), and three 86-year future scenarios (SSP126, SSP245 and SSP585, corresponding to radiative forcing of 2.5, 4.5 and 8.5 Wm^{-1} respectively in 2100, all for the years 2015-2100). Storm peak values are isolated from the time-series of 3-hourly H_S in each case, using the method of Ewans and Jonathan (2008). Storm events are identified for statistical analysis of peaks over threshold, by isolating continuous intervals of H_S between consecutive up- and down-crossings of a pre-specified threshold level, specified so that more than 20 but not exceeding 25 peaks are selected per annum, for each data set. The maximum value in a storm is referred to as the storm peak H_S (H_S^{sp}), the input for subsequent non-stationary extreme value analysis.

Our main interest is to understand changes in the tail characteristics of H_S^{sp} over a period of 86 years of output, for the different climate scenarios. Since 86 years is the longest period of data available for all scenarios, we choose this

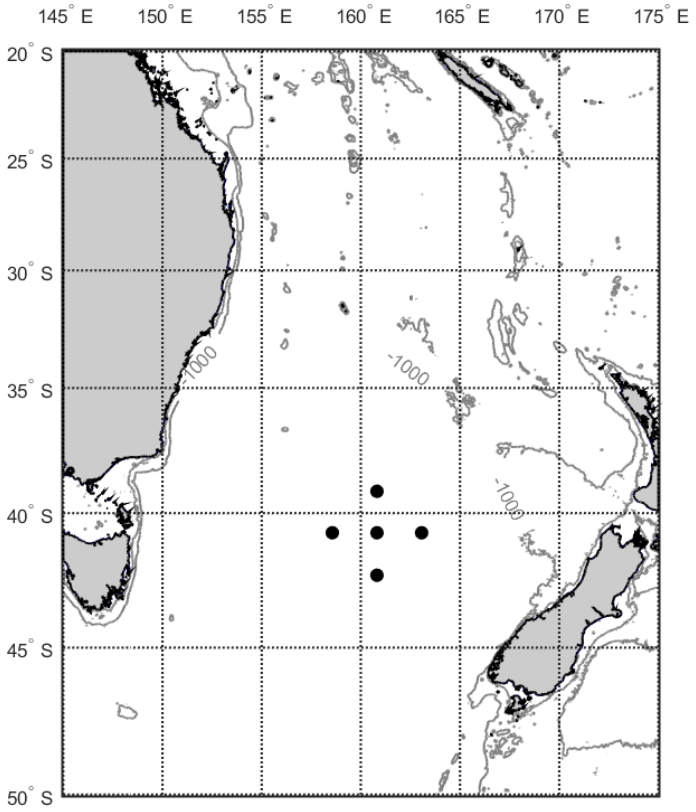

Figure 1: Locations considered in the central Tasman Sea, referred to as Centre (C), North (N), South (S), East (E) and West (W) in the obvious sense. Also shown are the land masses of Australia (left) and New Zealand (right), and bathymetry.

as the standard period for all comparisons below. As a base case, corresponding to inherent steady-state conditions, we quantify changes in H_S^{sp} tail for the piControl output over the time period 86 years. Since the piControl data is of length 700 years, we can achieve this by selecting a representative number of subintervals of piControl data, each of length 86 year, selected from 700 years of piControl data (for years 301-1000) such that the starting year for subintervals is approximately uniformly distributed on the interval of piControl data. A total of 25 subintervals are used in the current work. The spread of estimates (e.g. of return value) over the 25 subintervals gives a direct quantification of inherent piControl uncertainty. Furthermore, we can isolate Start and End periods of the Historical data (length 165 years), as the first and last 86 years of that output (recognising that there will be some overlap between the two subintervals), and then assess whether the changes in the 100-year return value for the Start and End Historical subintervals are similar.

Figure 2 illustrates the tail characteristics of all data considered in terms of scatter plots of the ordered 40 largest values of H_S^{sp} from a particular climate scenarios against the piControl "Standard" or "expected" tail. This is estimated as the mean ordered sequence of top 40 values of H_S^{sp} over the 25 random subintervals of piControl data. That is, if $x_{(i),j}$ is the i^{th} largest value of H_S^{sp} from random subinterval j, $j = 1, 2, ..., 25$, then $\bar{x}_{(i)} = \sum_j x_{(i),j}/25$ is the i^{th} largest value of piControl Standard tail. Lines falling above the $y = x$ guide indicate tails "longer" than the piControl Standard tail in each panel. From the top left panel, we see that there is some difference between the tails of H_S^{sp} for the Start and End Historical data, and the piControl Standard. It appears that the largest values in the Start Historical period are larger than their counterparts in the piControl Standard, but that the largest values in the End Historical period are smaller than their piControl counterparts. The remaining panels indicate little evidence of disparity between SSP and piControl Standard for scenarios SSP126 and SSP245. For SSP585, it appears that the largest values of H_S^{sp} are larger than their piControl Standard counterparts. Nevertheless, there is considerable variability in the position of lines for different locations.

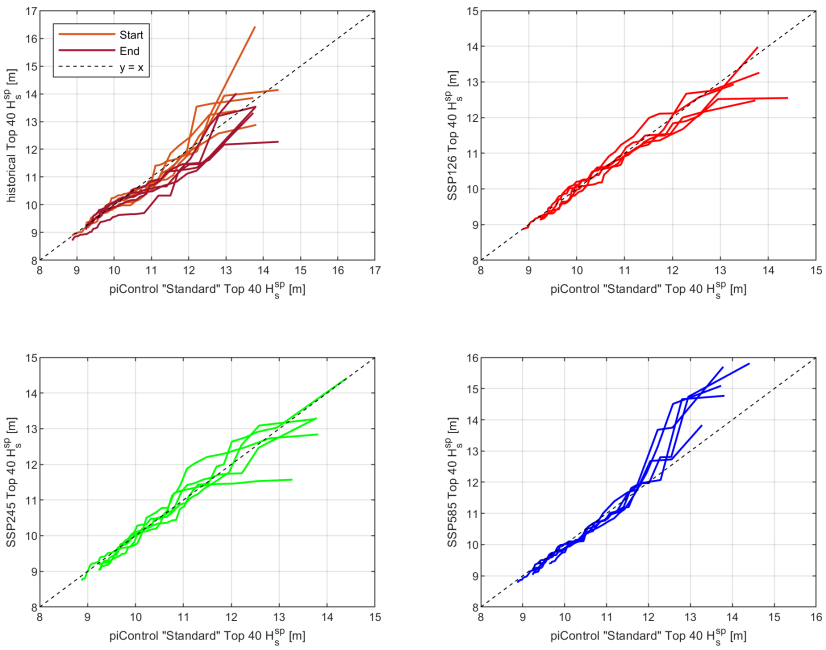

Figure 2: Comparison of the ordered 40 largest values of H_S^{sp} from different climate scenarios with the piControl "Standard" tail: Start and End Historical period (top left), SSP126 (top right), SSP245 (bottom left) and SSP585 (bottom right). In each panel, individual lines illustrate the relationship for each of five locations (C, N, S, E, W). The line $y = x$ is added for guidance.

3. Methodology

The statistical methodology used is described in Ewans and Jonathan (2023) and is summarised here. We use generalised Pareto (GP) regression to estimate a model for a set of observations $\{x_{t_i}, t_i\}_{i=1}^n$ of peaks over threshold of storm peak H_S events X_t at times $t_i \in (0, P)$, within which all model parameters are assumed to vary linearly over the period of observation, unless stated otherwise. That is, for any model parameter η, we assume that

$$\eta_t = \eta(t) = \eta^S + \frac{t}{P}(\eta^E - \eta^S), \text{ for } t \in (0, P] \tag{1}$$

in year t, where η^S and η^E are the parameter values at the start year (e.g. 2015) and end year (e.g. 2100) of P (e.g.86) years of data to be estimated. We assume that $X_t|X_t > \psi_t$ follows the GP distribution with threshold parameter $\psi_t \in \mathbb{R}$, scale $\sigma_t > 0$ and shape $\xi_t \in \mathbb{R}$ for $t \in (0, P)$ with distribution function

$$F_{\text{GP}}(x|X_t > \psi_t, \psi_t, \sigma_t, \xi_t) = 1 - \left[1 + \frac{\xi_t}{\sigma_t}(x - \psi_t)\right]^{-1/\xi_t} \tag{2}$$

when $\xi_t \neq 0$ and $1 - \exp(-(x - \mu_t)/\sigma_t)$ otherwise. Model parameters $\eta_t \in \{\sigma_t, \xi_t\}$ vary with t as described in Equation 1. To use Equation 2 in practice also requires a model for EV threshold ψ_t. Here we use quantile regression with lack-of-fit criterion

$$\ell_\psi = \tau \sum_{i, r_i \geq 0}^{n} |r_i| + (1 - \tau) \sum_{i, r_i < 0}^{n} |r_i| \tag{3}$$

for residuals $r_i = x_{t_i} - \psi_{t_i}$, and some fixed quantile non-exceedance probability $\tau \in [0, 1]$. Equation 3 can be interpreted as a Laplace likelihood for estimation. Since estimation of an optimal τ is problematic in general, models are estimated for a wide range of values of τ exceeding the mode of the empirical distribution of X_t, and sensitivities of inferences to τ assessed. In this work, we examine the performance of GP models over four choices of EV threshold. These are specified in terms of the non-exceedance probability (NEP) to which they correspond, and referred to henceforth as NEP1-4. NEP1 corresponds to $\tau = 0.5$, and NEP4 to the non-exceedance probability which leaves 30 threshold exceedances remaining for EV analysis. Values of τ for intermediate NEP2 and NEP3 are then equally spaced (on log scale) between those of NEP1 and NEP4. The asymptotic GP model form is likely to be more appropriate for NEP4 in general, and hence the bias of inference smallest; but the uncertainty in parameter estimates is likely to be lowest for NEP1: a classic bias-variance trade-off. To use Equation 2 for return value estimation, we also need to estimate

the annual rate of occurrence ρ_t of threshold exceedances in time for given τ. We achieve this using Poisson regression (e.g. Chavez-Demoulin and Davison 2005, Ross et al. 2017) with density

$$f(\{c_t\} \mid \rho_t) = \exp\left(-\sum_{t=1}^{P} \rho_t\right) \prod_{t=1}^{P} \rho_t^{c_t} \tag{4}$$

where $\{c_t\}_{t=1}^{P}$ are empirical annual counts of threshold exceedances, and ρ_t is also parameterised as in Equation 1.

For GP-distributed threshold exceedances, and Poisson-distributed rate of threshold exceedance, the distribution of the annual maximum is known to be GEV-distributed (e.g. Jonathan and Ewans 2013). Hence, in the absence of parameter uncertainty, the T-year return value Q_t at year t (for $T = 100$ years) is estimated as the $p = 1 - 1/T$ quantile of this distribution. Specifically

$$Q_t = \frac{\sigma_t}{\xi_t}\left[\left(-\frac{\log p}{\rho_t}\right)^{-\xi_t} - 1\right] + \mu_t \tag{5}$$

when $\xi_t \neq 0$ and $\mu_t - \sigma_t \log[-(1/\rho_t)\log p]$ otherwise. Note that since all model parameters change in time, then so does the value of Q_t.

Parameter estimation is undertaken in a sequential manner using Bayesian inference. First we perform quantile regression, generating a sample $\{\hat{\psi}_k^S, \hat{\psi}_k^E\}_{k=1}^{n_I}$ of size n_I from the joint posterior distribution of EV threshold parameters. We then use the non-stationary threshold corresponding to posterior mean parameter estimates from the quantile regression, (a) in Poisson regression, to generate a sample $\{\hat{\rho}_k^S, \hat{\rho}_k^E\}_{k=1}^{n_I}$ from the joint posterior of ρ; and (b) in GP regression to sample $\{\hat{\sigma}_k^S, \hat{\sigma}_k^E, \hat{\xi}_k^S, \hat{\xi}_k^E\}_{k=1}^{n_I}$ from the joint posterior of GP parameters, where $n_I > 10000$. These sets of posterior parameters are used to estimate the distribution of T-year return value, and in particular to compare the estimates Q_1 (for the first year) with Q_P (for the last year, $P = 86$) with $T = 100$. Further details are provided in Ewans and Jonathan (2023).

4. Analysis

Here we quantify climate-related trends in the 100-year return value for H_S^{sp} for the neighbourhood of interest, under the various forcing scenarios. We use the piControl data as a base case for comparison, corresponding to inherent steady-state behaviour.

Comparisons are made in terms of two quantities, ΔH_{S100}^{sp} and $\text{med}\Delta H_{S100}^{sp}$. For a given location, EV threshold level, climate scenario and potentially subinterval of the climate output, we define ΔH_{S100}^{sp} as the difference between the estimates of 100-year return value at the end (Q_P, see Equation 5) and the start (Q_1) of each sample of $P = 86$ years of data considered. EV analysis of this sample using MCMC provides 1000 estimates from the posterior distribution of ΔH_{S100}^{sp} (namely the final 1000 iterations from the MCMC analysis post burn-in). The median value of ΔH_{S100}^{sp} over MCMC iterations, referred to as $\text{med}\Delta H_{S100}^{sp}$ is a useful central summary statistic of these 1000 estimates.

Since we believe that the H_S^{sp} tail for the five locations under consideration will be very similar (for given EV threshold level, climate scenario and potentially subinterval of the climate output), and in particular that differences between locations are more likely due to arbitrary GCM-wave model run effects rather than long-term physics, it is reasonable to model average estimates for the change in 100-year return value over locations (providing location differences are not implausibly large, indicating more fundamental issues with GCM-wave model output, or the EV analysis performed). Similarly, selection of a single best EV threshold level for POT analysis is problematic; therefore it is reasonable to model average estimates over a set of plausible threshold choices. Final estimates for the distribution of the change in 100-year return value over 86 years (or summary statistics of that distribution, such as the median) for a given climate scenario will therefore be model averages over five locations, and four EV thresholds and potentially 25 subintervals of the piControl output.

In Section 4.1, we estimate the distribution of ΔH_{S100}^{sp} and $\text{med}\Delta H_{S100}^{sp}$ for subintervals of length 86 years of the piControl output. In Section 4.2, we compare estimates for the distribution of ΔH_{S100}^{sp} and $\text{med}\Delta H_{S100}^{sp}$ for the Start and End Historical output with that for the piControl output. In Section 4.3, we compare estimates for the distribution of ΔH_{S100}^{sp} and $\text{med}\Delta H_{S100}^{sp}$ for the SSP126, SSP245 and SSP585 output with that for piControl.

4.1. Inherent variability in return value over 86 years

We begin by considering the inherent uncertainty of ΔH_{S100}^{sp} and $\text{med}\Delta H_{S100}^{sp}$. The black dashed curve in Figure 3 represents the empirical density of ΔH_{S100}^{sp} over all locations, EV threshold levels and piControl subintervals (therefore estimated using a set of $5 \times 4 \times 25 \times 1000 = 500,000$ values). The black dotted line in the figure is the corresponding

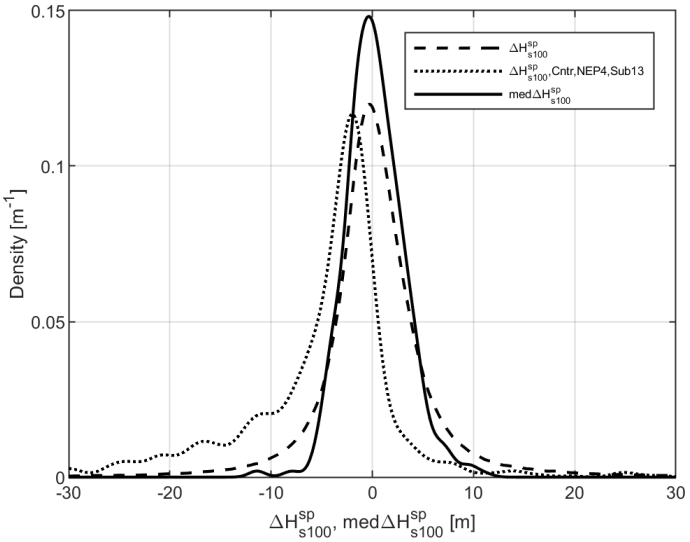

Figure 3: Empirical density (black dashed) of ΔH_{S100}^{sp} over all locations, EV threshold levels and piControl subintervals. Empirical density (black dotted) of ΔH_{S100}^{sp} for the Centre location, NEP4 and the 13$^{\text{th}}$ piControl subinterval (with starting value at year approximately 650). Empirical density (black solid) of medΔH_{S100}^{sp} over all locations (C, N, S, E, W), EV threshold levels (NEP1-NEP4) and piControl subintevals (with 25 starting years uniformly spread on the interval of piControl data).

density for the Centre location, EV threshold level NEP4 and the 13$^{\text{th}}$ piControl subinterval (with starting year approximately midway between years 301 and 1000, at around year 650). Comparing the dashed and dotted curves therefore gives some impression of the size of inherent uncertainty, with likely contributing factors from genuine temporal non-stationarity combined with epistemic uncertainty from a limited sample for EV analysis. The black solid curve is the empirical density of medΔH_{S100}^{sp} over locations, EV threshold levels and piControl subintervals (therefore estimated using a set of $5 \times 4 \times 25 = 500$ values).

From the figure we see that the most probable value of the change ΔH_{S100}^{sp} in the 100-year return value over a period of 86 years (black dashed) is approximately zero, as would be expected given that the piControl output represent a pre-industrial climate with no climate forcing. Further, the most probable medΔH_{S100}^{sp} (black solid) is also approximately zero. However, empirical densities for ΔH_{S100}^{sp} and medΔH_{S100}^{sp} are very broad, approximately $\pm$5m at half height. The density of ΔH_{S100}^{sp} for the specific combination of location, EV threshold level and piControl subinterval indicated, is similar in general shape to that of ΔH_{S100}^{sp} over all combinations, except that the density is displaced to the left by approximately 2m. That is, estimates of medΔH_{S100}^{sp} and medΔH_{S100}^{sp} from analysis of H_S^{sp} from an SSP output, with magnitudes up to 5m are entirely consistent with the inherent uncertainty of the piControl data.

4.2. Variability in return value over early and late historical periods of 86 years

Now we consider the variability of ΔH_{S100}^{sp} and medΔH_{S100}^{sp} for the Start and End Historical intervals introduced in Section 2. Figures 4 and 5 show estimated empirical distributions for ΔH_{S100}^{sp} and medΔH_{S100}^{sp} over locations and EV threshold levels for the Start and End Historical samples in orange dashed and brown dashed respectively. The figures also show corresponding "null distribution" estimate from the piControl data, reproduced from Figure 3.

For the Start Historical period, the figures indicate negative mode estimates of approximately -1m for ΔH_{S100}^{sp} (Figure 4) and approximately -1.5m for medΔH_{S100}^{sp} (Figure 5). Otherwise, differences between empirical densities for ΔH_{S100}^{sp} and medΔH_{S100}^{sp} for the Start and End Historical periods, and the corresponding "null densities" estimated using the piControl data are small. There is some evidence, e.g. a left-hand shoulder on the empirical densities for ΔH_{S100}^{sp} and medΔH_{S100}^{sp} for the End Historical period, suggesting a further reduction in return value, but this evidence is very weak.

4.3. Variability in return value over 86 years under SSP climate scenarios

Finally, we consider the variability of ΔH_{S100}^{sp} and medΔH_{S100}^{sp} for the SSP126, SSP245 and SSP585 output. Figures 6 and 7 give empirical densities for ΔH_{S100}^{sp} and medΔH_{S100}^{sp} (over locations and EV threshold levels) for each scenario, and compare these with the corresponding "null distribution" estimates from the piControl data, again reproduced from Figure 3.

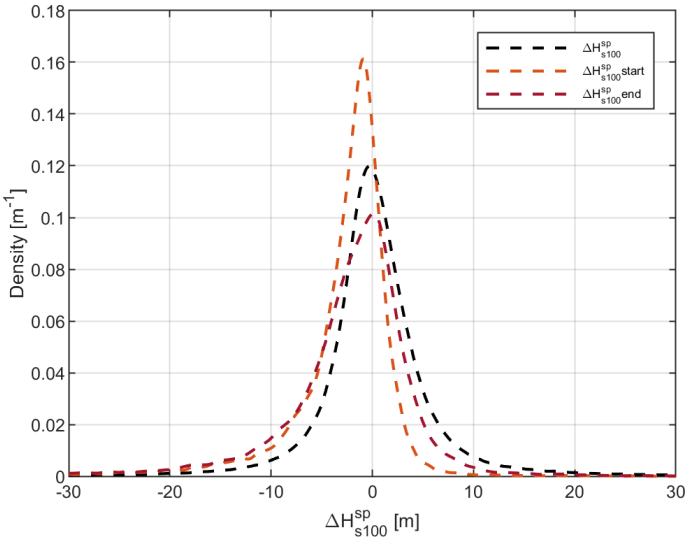

Figure 4: Empirical densities for ΔH^{sp}_{S100} estimated over all locations (C, N, S, E, W) and EV threshold levels (NEP1-NEP4), for the Start (orange dashed) and End (brown dashed) Historical periods. Also shown for comparison (black dashed) is the corresponding empirical density for ΔH^{sp}_{S100} from the piControl data (over all locations, NEP levels and piControl subintervals), reproduced from Figure 3.

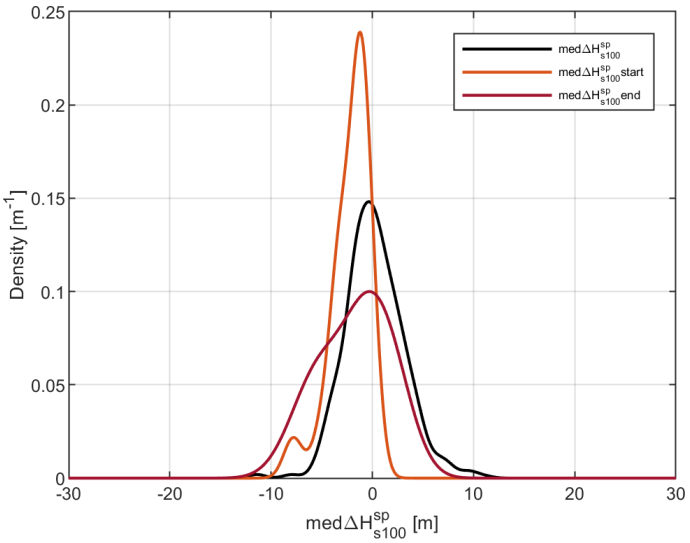

Figure 5: Empirical densities for $\text{med}\Delta H^{sp}_{S100}$ estimated over all locations (C, N, S, E, W) and EV threshold levels (NEP1-NEP4), for the Start (orange solid) and End (brown solid) Historical periods. Also shown for comparison (black solid) is the corresponding empirical density for $\text{med}\Delta H^{sp}_{S100}$ from the piControl data (over all locations, NEP levels and piControl subintervals), reproduced from Figure 3.

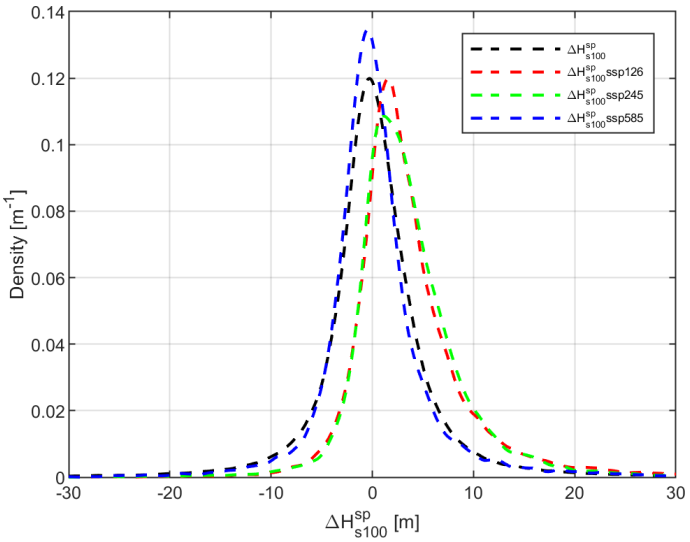

Figure 6: Empirical densities for ΔH^{sp}_{S100} estimates over all locations (C, N, S, E, W) and EV threshold levels (NEP1-NEP4) for the SSP126 (red dashed), SSP245 (green dashed) and SSP585 (blue dashed) climate scenarios. Also shown for comparison (black dashed) is the corresponding empirical density for ΔH^{sp}_{S100} from the piControl data (over all locations, NEP levels and piControl subintervals), reproduced from Figure 3.

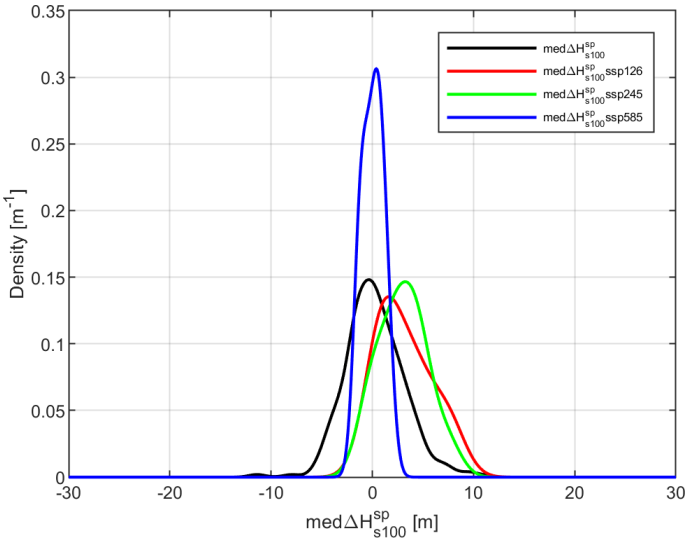

Figure 7: Empirical densities for medΔH^{sp}_{S100} estimates over all locations (C, N, S, E, W) and EV threshold levels (NEP1-NEP4) for the SSP126 (red solid), SSP245 (green solid) and SSP585 (blue solid) climate scenarios. Also shown for comparison (black solid) is the corresponding empirical density for medΔH^{sp}_{S100} from the piControl data (over all locations, NEP levels and piControl subintervals), reproduced from Figure 3.

From the figures we see that the modes of the distributions of ΔH^{sp}_{S100} and medΔH^{sp}_{S100} lie at approximately +2m and +1.5m respectively for SSP126, and at approximately +2m and +3m for SSP245, suggesting an increase in return value for these scenarios. However, for scenario SSP585, the most probable values for both ΔH^{sp}_{S100} and medΔH^{sp}_{S100} are both less than 0.5m in magnitude. Once again however, we note that all these differences are small compared to the inherent uncertainties present.

5. Discussion and conclusions

In this work, we examine data for storm peak significant wave height (H^{sp}_S) from the FIO-ESM v2.0 Earth System Model, for five locations in a neighbourhood in the central Tasman Sea between Australia and New Zealand, for a number of different climate scenarios.

The piControl output provides a useful base case to assess the size of changes in 100-year return value for H^{sp}_S. We find that the most probable value of the change ΔH^{sp}_{S100} in the 100-year return value over a period of 86 years is approximately zero, and that the most probable change medΔH^{sp}_{S100} in the median 100-year return value over 86 years is also approximately zero. However, empirical densities for ΔH^{sp}_{S100} and medΔH^{sp}_{S100} are very broad: we would not be particularly surprised by a change in the 100-year return value for H^{sp}_S over 86 years, or its median value, of $\pm$5m.

Exploratory analysis of Historical H^{sp}_S (for years 1850-2014) indicates that the largest values of H^{sp}_S in the first 86 years (1850-1935, referred to as "Start Historical") are somewhat larger than the typical largest values observed in a contiguous period of 86 years from the pre-industrial piControl data (for nominal years 301-1000); further, the largest values of H^{sp}_S in the last 86 years (1929-2014, referred to as "End Historical") are somewhat smaller than the typical largest values observed in a contiguous period of 86 years from the pre-industrial piControl data. Overall, this represents weak evidence in support of a slight reduction in storm severity over the course of the Historical period. The typical reduction in maximum H^{sp}_S is approximately 1m; however, this change cannot be regarded as "significant" in a statistical sense. Non-stationary extreme value analysis for the Start Historical period provides weak evidence in favour of a reduction in the 100-year return value (e.g. a negative mode estimate of approximately -1m for ΔH^{sp}_{S100} and approximately -1.5m for medΔH^{sp}_{S100}). The differences between the empirical densities for ΔH^{sp}_{S100} and medΔH^{sp}_{S100} and the corresponding "null densities" estimated using the piControl data are small. There is some evidence, e.g. a left-hand shoulder on the empirical densities for ΔH^{sp}_{S100} and medΔH^{sp}_{S100} for the End Historical period, suggesting a further reduction in return value, but this evidence is very weak. Moreover, the magnitudes of these effects are small relative to the inherent uncertainty ($\pm$5m) observed in the piControl output.

There is no evidence that the largest values of H^{sp}_S from model output corresponding to forcing scenario SSP126 and SSP 245 are different to typical largest values from the piControl output. For scenario SSP585 however, the very largest values of H^{sp}_S are approximately 1.5m larger than typical largest values from piControl. Estimates for the distributions of ΔH^{sp}_{S100} and medΔH^{sp}_{S100} from non-stationary extreme value analysis suggest that the 100-year return value for H^{sp}_S increases under scenarios SSP126; the modes of the distributions of ΔH^{sp}_{S100} and medΔH^{sp}_{S100} lie at approximately +2m and +1.5m respectively. The story is similar for scenario SSP245; the modes of the distributions of ΔH^{sp}_{S100} and medΔH^{sp}_{S100} lie at approximately +2m and +3m. However, for scenario SSP585, the most probable values for both ΔH^{sp}_{S100} and medΔH^{sp}_{S100} are both less than 0.5m in magnitude.

In order to investigate the different characteristics of the empirical densities for ΔH^{sp}_{S100} corresponding to piControl and SSP scenarios shown in Figure 6, Figure 8 provides plots of densities estimated with specific choices of location and EV threshold level. The figure is comprised on four panels, one for each of piControl, SSP126, SSP245 and SSP585. Within each panel, dashed lines represent densities estimated using either EV threshold levels NEP1 or NEP2, for each of the five spatial locations. Similarly, solid lines represent densities estimated using either EV threshold levels NEP3 or NEP4. Hence there are 10 lines, 5 dashed and 5 solid, per panel. Estimates for piControl and SSP585 show considerable consistency; we cannot see obvious differences associated with change of location or EV threshold. However, for SS126 and SSP245, a right-hand shoulder is present. Further investigation indicates that the shoulder appears at least in part to be associated with zonal variation in ΔH^{sp}_{S100} for SSP126 and meridional variation for SSP245. However, there appears not to be a clear effect with EV threshold level for either scenario. We speculate that the source of zonal and meridional effects may be the arbitrary GCM-wave model run effects discussed in Section 4, since in reality we do not expect to find systematic variation in return value over the relatively small extent of the spatial neighbour examined. More generally, we believe these results provide a warning against over-interpretation of specific EV analyses which may be sensitive to arbitrary modelling effects, and the prudence of careful model averaging to reveal persistent features. Accordingly, we believe that model averaging over analysis of output from multiple GCM-wave models, and ensembles from given GCM-wave model, will prove advantageous.

From a design perspective for year 2100 at this location, with no preference regarding climate scenario and notwithstanding large uncertainties, we might conclude that planning for an increase in 100-year return value for H^{sp}_S of around 2m relative to 2015 would be wise. However, the evidence in favour of a climate-driven change in H^{sp}_S is weak. More

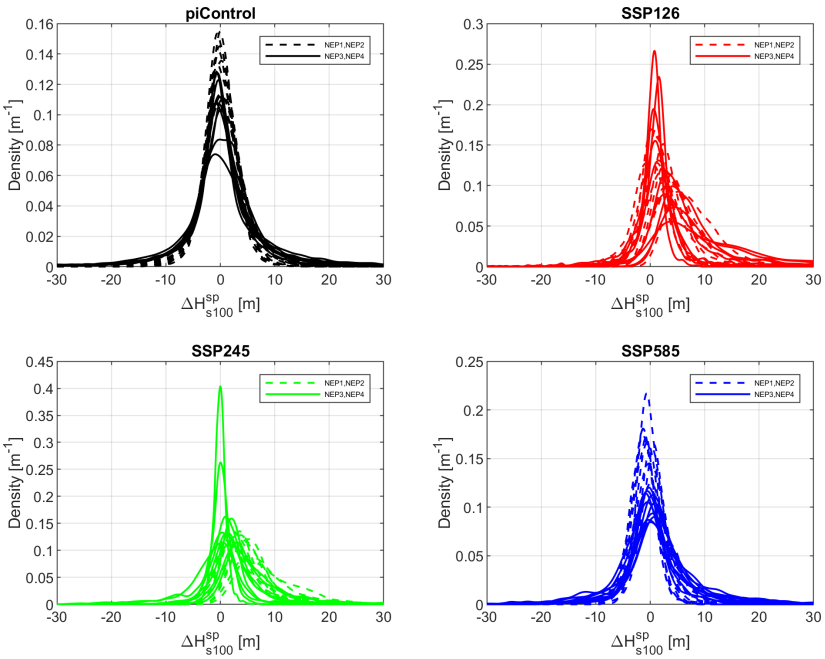

Figure 8: Empirical densities of ΔH^{sp}_{S100} estimates for individual combinations of each location (C, N, S, E, W) with EV threshold levels NEP1-NEP2 (dashed) and with NEP3-NEP4 (solid), for piControl (top left, black), SSP126 (top right, red), SSP245 (, bottom left, green) and SSP585 (bottom right, blue) climate scenarios. Empirical densities for piControl are model averages over estimates for the 25 piControl subintervals.

generally, we hope that the approach adopted in this work, particularly (1) non-stationary extreme value analysis and the uncertainty quantification it allows, (2) exploitation of piControl data to estimate inherent uncertainty, and (3) appropriate model averaging over space and other sources of model uncertainty, will prove generally useful to metocean practitioners.

Acknowledgments

The authors would like to thank Zhenya Song (First Institute of Oceanography, Qingdao, China), and Alberto Meucci and Ian Young (U. Melbourne, Australia), for discussions. Software for the non-stationary extreme value analysis is provided at Jonathan (2021).

References

Albuquerque, J., Antolinez, J.A.A., Mendez, F.J., Coco, G., 2022. On the projected changes in New Zealand's wave climate and its main drivers. N. Z. J. Mar. Freshwater Res. .

Bao, Y., Song, Z., Qiao, F., 2020. FIO-ESM Version 2.0: Model Description and Evaluation. J. Geophys. Res. Oceans 125, 1–21.

Casas-Prat, M., Wang, X.L., Mori, N., Feng, Y., Chan, R., Shimura, T., 2022. Effects of internal climate variability on historical ocean wave height trend assessment. Front. Mar. Sci. 9, 847017.

Chavez-Demoulin, V., Davison, A., 2005. Generalized additive modelling of sample extremes. J. Roy. Statist. Soc. Series C: Appl. Stat. 54, 207–222.

Ewans, K., Jonathan, P., 2008. The effect of directionality on northern North Sea extreme wave design criteria. J. Offshore. Arct. Eng. 130, 041604:1–041604:8.

Ewans, K., Jonathan, P., 2020. Extreme conditions, in: Young, I., Babanin, A. (Eds.), Ocean wave dynamics. World Scientific, pp. 271–319.

Ewans, K., Jonathan, P., 2023. Uncertainties in estimating the effect of climate change on 100-year return period significant wave heights. Ocean Eng., under review URL: `https://arxiv.org/abs/2212.11049`.

Jonathan, P., 2021. Simple non-stationary extremes estimation for peaks-over-threshold data, parameters linear in covariate, MCMC inference. https://github.com/ygraigarw/pGpNonStt.

Jonathan, P., Ewans, K., 2013. Statistical modelling of extreme ocean environments with implications for marine design : a review. Ocean Eng. 62, 91–109.

McSweeney, C., Jones, R., 2013. No consensus on consensus: the challenge of finding a universal approach to measuring and mapping ensemble consistency in GCM projections. Clim. Change 19, 617–629.

Meucci, A., Young, I.R., Hemer, M., Kirezci, E., Ranasinghe, R., 2020. Projected 21st century changes in extreme wind-wave events. Sci. Adv. 6, 1–10.

Meucci, A., Young, I.R., Hemer, M., Trenham, C., Watterson, I.G., 2022. 140 Years of Global Ocean Wind-wave Climate Derived from CMIP6 ACCESS-CM2 and EC-Earth3 GCMs. Global Trends, Regional Changes, and Future Projections. J. Clim. doi:10.1175/JCLI-D-21-0929.1.

Morim, J., Trenham, C., Hemer, M., Wang, X.L., Mori, N., Casas-Prat, M., Semedo, A., Shimura, T., Timmermans, B., Camus, P., Bricheno, L., Mentaschi, L., Dobrynin, M., Feng, Y., Erikson, L., 2020. A global ensemble of ocean wave climate projections from cmip5-driven models. Sci. Rep. 7, 105.

Ross, E., Randell, D., Ewans, K., Feld, G., Jonathan, P., 2017. Efficient estimation of return value distributions from non-stationary marginal extreme value models using Bayesian inference. Ocean Eng. 142, 315–328.

Sardana, D., Kumar, P., Bhaskaran, P.K., Nair, T.M.B., 2022. The projected changes in extreme wave height indices over the Indian Ocean using COWCLIP2.0 datasets. Climate Dynamics doi:10.1007/s00382-022-06579-5.

Song, Z., Bao, Y., Zhang, D., Shu, Q., Song, Y., Qiao, F., 2020. Centuries of monthly and 3-hourly global ocean wave data for past, present, and future climate research. Scientific Data 7, 1–11.

IPCC, 2021. Summary for Policymakers. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_SPM.pdf.

Vanem, E., 2015. Non-stationary extreme value models to account for trends and shifts in the extreme wave climate due to climate change. Appl. Ocean Res. 52, 201 – 211.

Young, I.R., Ribal, A., 2019. Multiplatform evaluation of global trends in wind speed and wave height. Science 364, 548–552.