Ocean Engineering 237 (2021) 109584

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

Wave conditions encountered by ships—A report from a larger shipping company based on ERA5

Ulrik D. Nielsen [a,b,*], Angelos Ikonomakis [a,c]

[a] DTU Mechanical Engineering, Technical University of Denmark, DK-2800 Kgs., Lyngby, Denmark
[b] Centre for Autonomous Marine Operations and Systems, NTNU AMOS, NO-7491, Trondheim, Norway
[c] Maersk R & D, DK-1098, Copenhagen, Denmark

ARTICLE INFO

Keywords:
Encountered wave heights by ships
ERA5
Global wave statistics
Weather routing
Temporal and spatial variation along ship routes
Gumbel and *t* location-scale distributions

ABSTRACT

This study reports about the wave conditions encountered by the fleet of Maersk Line. Herein, *wave conditions* refer to integral wave parameters in terms of significant wave height, zero-upcrossing period, and wave direction (equivalently relative wave heading), but main focus is on the significant wave height. The study includes data from 189 container ships, where auto-logged data has been collected over a three-years period (2017–2020) during operations in the majority of the world's larger oceans. In total, the data corresponds to 1 million hours of operation. The encountered wave heights are compared with the information from wave scatter diagrams given by the Global Wave Statistics (British Maritime Technology, 1986). The study shows that weather routing and seamanship have an effect, but the effect is not as pronounced as reported in a previous study (Olsen et al., 2006). The study also presents findings in relation to the temporal and spatial *variation* in encountered wave height along ship routes at consecutive wave points spaced by the sailing distance in 30–120 min. It is shown that the variation can be large and, as such, the concept of "stationary conditions" might be compromised sometimes in the analysis of wave-ship interactions.

1. Introduction

The safety and fuel efficiency of a ship sailing in open sea depend largely on the encountered wave conditions equivalently the sea state. Concerning structural safety this is recognised and explicitly accounted for by structural rules of ships, while more general aspects of safety associated with critical situations (deck wetness, slamming, high acceleration levels, etc.) also are highly important. The concern in relation to fuel efficiency associates to the added resistance experienced in waves and lower propeller efficiency. In all cases, low(er) sea states are beneficial to the sailing operations.

Most shipping companies rely on weather routing in combination with good seamanship of the crew. The effect is that ships, statistically, are expected to observe less severe wave conditions than stipulated by, e.g., structural rules (IACS-REC 034, 2001; IACS-REC 106, 2009) and as derived by global wave statistics from wave scatter diagrams (British Maritime Technology, 1986). This has been investigated by, e.g., Olsen et al. (2006).

The measuring of waves encountered by ships is difficult. On the other hand, numerical modelling with spectral wave models in combination with data assimilation has led to comprehensive reanalyses. One of these is the ERA5 database (Copernicus Climate Change Service Information, 2020; Hersbach et al., 2021) that provides hourly updates of wave spectra in grid points spaced by 0.5° in latitudes and longitudes for all parts of the world's oceans.

1.1. Scope and objective

This study considers output from the ERA5 database in combination with in-service data from about 200 ships of Maersk Line. The ship data has been collected over a three-years period (2017–2020) from operations in the majority of the larger oceans, see Fig. 1. The focus is exclusively on encountered wave conditions and, as such, the necessary ship data is limited to GPS and compass heading, but vessel forward speed is also included in the analysis. The study has similarities with Olsen et al. (2006) but is a large extension in the sense that the present study includes more than 1 million hourly observations in total, compared to about 25,000 observations included in Olsen et al. (2006). Moreover, the ship data has been collected automatically from the single vessels using an onboard continuous monitoring system with sampled output every 10 min. The study presents statistics of the encountered wave conditions which are compared with results from existing wave scatter diagrams (British Maritime Technology, 1986).

* Corresponding author at: DTU Mechanical Engineering, Technical University of Denmark, DK-2800 Kgs., Lyngby, Denmark.
E-mail address: udn@mek.dtu.dk (U.D. Nielsen).

https://doi.org/10.1016/j.oceaneng.2021.109584
Received 20 March 2021; Received in revised form 24 July 2021; Accepted 27 July 2021
Available online 3 August 2021

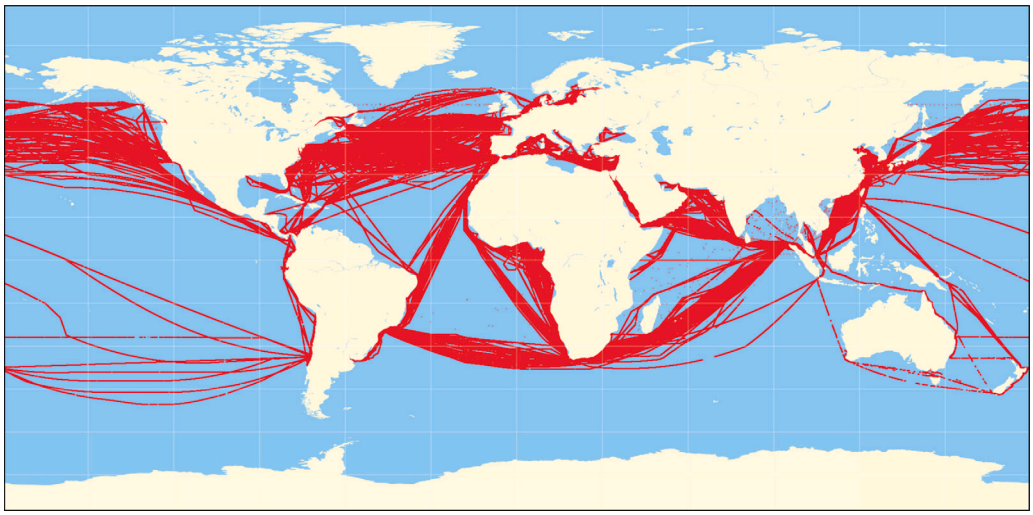


Fig. 1. Path projections of the considered voyages.

Besides, the study investigates and discusses the variation in encountered wave height along the route paths of the ships. In this sense, the paper follows up on the study by Nielsen (2021). Above all, the main objective is to give an account of the wave conditions that ships from the Maersk Line fleet encounter. Due to the large number of ships considered, and the vast amount of data, it is believed that the findings can be extrapolated to container ships in general, although seamanship and the level and sophistication of company-specific weather routing can be influential.

It should be emphasised that the study primarily deals with findings associated to significant wave height. Little attention is given to zero-upcrossing period and (relative) wave direction, but a few results are included to complement the analysis. Besides, it is noteworthy that, in general, a ship's response to waves depends strongly also on the distribution of energy, and not just on the amount of energy, of the encountered wave system. In particular, it would therefore be relevant to study observations of the zero-upcrossing period to the same detail as done herein for the significant wave height, albeit this is left as a future exercise.

1.2. Composition of paper

The next section presents the basic concepts and the methodology of the study. A brief description of the studied ship data is given in Section 3. The report on the encountered wave conditions is included in Section 4 and associated discussions follow in Section 5. Finally, concluding remarks are presented in Section 6.

2. Fundamentals and methodology

2.1. Encountered wave conditions (ERA5)

The ERA5 hourly ocean wave data on single levels (Copernicus Climate Change Service Information, 2020; Hersbach et al., 2021), as used in this study, is available on a regular latitude–longitude grid at $0.5° \times 0.5°$ resolution with updates every 60 min. In principle, the 2D wave spectrum is available but this study considers the associated integral parameters only. The following parameters are considered: significant wave height H_s, zero-upcrossing period T_z, and mean wave direction D_m (*coming from*); all mathematically expressed in accordance with their standard definitions, (e.g., ECMWF, 2017).

2.2. Spatial and temporal interpolation

Realising that a ship rarely is exactly at a position coinciding with one of the points on the ERA5 grid, interpolation becomes necessary. Spatially, the ERA5 data is interpolated to the exact position of a ship using bilinear interpolation (Nielsen, 2021). Interpolation in significant wave height H_s and in zero-upcrossing period T_z can be made immediately. On the other hand, mean wave direction is circular, which leads to the ambiguity that $D_m = 0°$ and $D_m = 360°$ are equivalent. The interpolation in wave direction is therefore based on the Cartesian vector components of the particular directions that enter the interpolation. To account for any variation in significant wave height from grid point to grid point, the interpolation is weighted by H_s. Consequently, interpolation in (mean) wave direction is based on a new set of parameters (A, B) calculated for all the grid points (i, j), with:

$$A(i,j) = H_s(i,j) \cdot \cos(D_m(i,j)) \tag{1}$$

$$B(i,j) = H_s(i,j) \cdot \sin(D_m(i,j)) \tag{2}$$

In this case, interpolation at an arbitrary point off the grid can be made in both A and B. The mean wave direction at the considered geographical point is subsequently calculated by,

$$D_m = \text{atan2}(B, A) \tag{3}$$

Initially, (spatial) interpolation is made so that the wave conditions at the ship positions are obtained exactly at the hourly updates of the ERA5 dataset. Subsequently, to account for the variation in between the hourly updates at the particular hourly positions, temporal linear interpolation is made, so that the wave parameters (H_s, T_z, D_m) are obtained with 10-min spacing between consecutive waypoints. An example of such time series of encountered wave conditions from *one* ship is shown in Fig. 2. On the horizontal axis, sample indices are used instead of absolute time, which means that the spacing between two indices is 10 min. The gaps in the time series correspond to periods where the ship has been out of service, or data might have been removed, for instance due to quay stays or maintenance, cf. Section 3. It is realised that the data from the particular ship covers a period of about three years and two months.

2.3. Information from the Global Wave Statistics

The wave scatter diagrams provided by British Maritime Technology (1986) give the joint distribution of combinations of H_s and T_z for almost any area of the world's oceans. The sub-areas are shown on

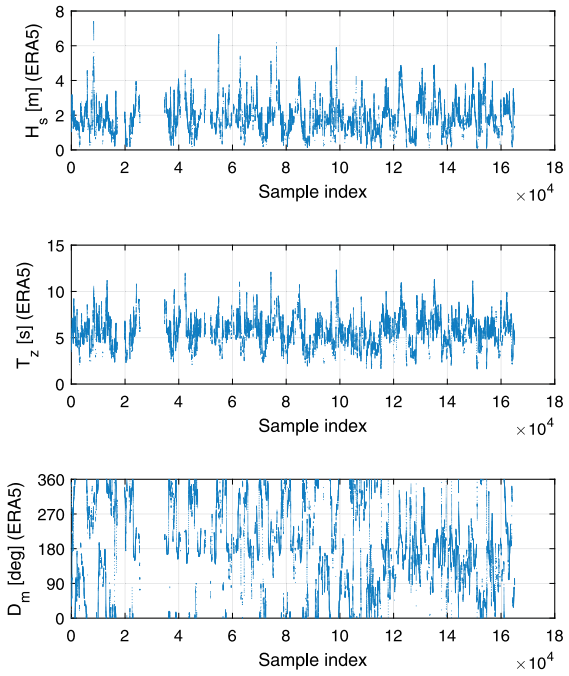


Fig. 2. Time series of significant wave height (top), zero-upcrossing period (middle), and mean wave direction (bottom), as encountered by one ship. The spacing between two consecutive sample indices correspond to 10 min.

the map in Fig. 3, noticing that the individual rectangular regions also are referred to as Marsden areas. The actual information for a specific Marsden area is given in terms of the number of times a particular combination of H_s and T_z is expected to occur out of 1,000 occurrences of given combinations, prescribed by a range of zero-upcrossing periods and a range of significant wave heights. The (marginal) probability density function (PDF) of the significant wave height corresponding to a specific Marsden area can therefore easily be derived.

2.4. Evaluation of the outcome

Essentially, the outcome of the present study is the set of interpolated wave parameters, as obtained for all considered vessels at their exact route paths. As such, computed statistics can be used to assess the final outcome.

The calculation of absolute statistics, such as mean value and standard deviation, is straightforward for H_s and T_z. The circularity of wave direction implies that the mean value $D_{m,\text{mean}}$ of a sequence of mean wave directions must be based on the Cartesian vector components of the single mean wave directions forming the sequence. This is equivalent to how interpolation in wave direction was made, cf. Eqs. (1) to (3), but without the H_s-based weighting. In quantitative terms, the formula reads,

$$D_{m,\text{mean}} = \text{atan2}(\bar{B}, \bar{A}), \tag{4}$$

$$\bar{A} = \frac{1}{K}\sum_{k=1}^{K}\cos(D_{m,k}),\quad \bar{B} = \frac{1}{K}\sum_{k=1}^{K}\sin(D_{m,k}) \tag{5}$$

where $k = 1 : K$ refers to a sequence of wave directions. The corresponding standard deviation σ is calculated in line with the mathematical definition of standard deviation. Specifically, it is obtained by

$$\sigma^2 = \frac{1}{K}\sum_{k=1}^{K}(\min_{abs}\{D_{m,k} - D_{m,\text{mean}}\})^2. \tag{6}$$

where the minimisation function $\min_{abs}\{\cdot\}$ is applied to ensure that it is always the minimum difference, in absolute terms, which is computed, i.e. $(D_{m,k} - D_{m,\text{mean}}) < 180°$ always, thus addressing that wave direction is circular. In addition to the mean value and standard deviation, metrics representing the maximum value and the 95-percentile will also be calculated for H_s and T_z, noting that the number 'p95' means that 95% of the observations have a value less than the number 'p95'.

The Gumbel distribution, also known as a Generalised Extreme Value distribution Type-I, is sometimes used to describe the probability distribution of significant wave height when addressing, e.g., extreme value predictions (Jensen, 2001). The Gumbel distribution is characterised by the location and scale parameters α and β, respectively, through the cumulative distribution function,

$$F(x|\alpha,\beta) = e^{-e^{-(x-\alpha)/\beta}},\quad \beta > 0 \tag{7}$$

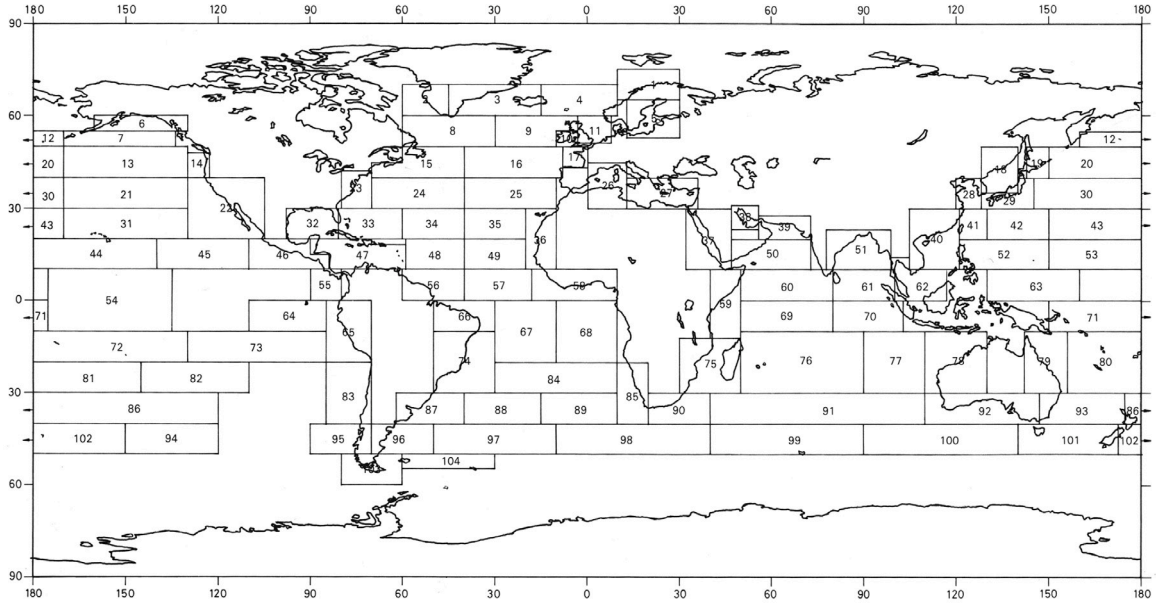
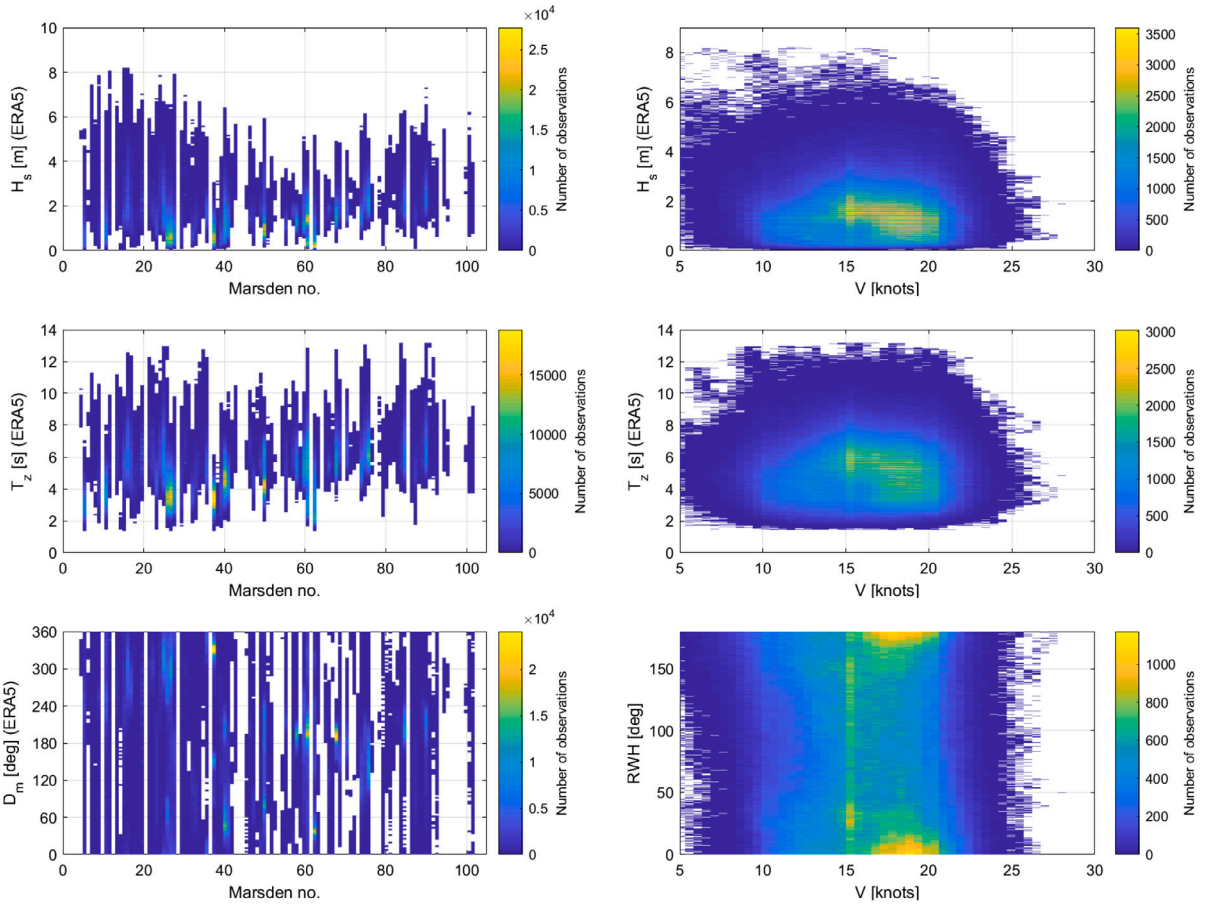


Fig. 3. Marsden areas as defined by British Maritime Technology (1986).



(a) Variation with Marsden area. (b) Variation with speed.

Fig. 4. Distribution of wave parameters as encountered in all voyages. Top: Significant wave height (H_s); middle: zero-upcrossing period (T_z); bottom: mean wave direction (D_m) *or* relative wave heading (RWH). Note the difference in the colour scale, counting the number of observations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.5. ERA5 as ground truth

This study relies fundamentally on the assumption that the ERA5 dataset represents the ground truth of the encountered wave conditions, noticing that the dataset is the output of the wave model ECWAM (ECMWF, 2017) assimilated with measurements of wind and waves. It is beyond the scope herein to investigate the assumption, but the literature contains many interesting studies in this direction, for instance, Hersbach et al. (2020), Stefanakos (2019), Hauteclocque et al. (2020), Takbash and Young (2020) and Timmermans et al. (2020) to mention just a few more recent studies justifying the use of ERA5. Obviously, in reality, the dataset is nothing more than an estimate, and some reports have informed about the underestimation of the significant wave height by ERA5, e.g., Naseef and Kumar (2020) and Belmadani et al. (2021) stating that ERA5 H_s data underestimates buoy observations at various locations including the northwestern Atlantic. Another noteworthy comment is that, in the present study, no attempt is made to compare with the outcome from other wave models such as WAVEWATCH III (Tolman et al., 2002; The WAVEWATCH III Development Group, 2019), saying this just to add that, although small differences in the results could appear, it is believed that the overall findings would remain the same.

3. Data

The fundamental and necessary ship data consists of measurements of the geographical position given by GPS, and the compass heading that can be used together with the mean wave direction D_m for computing the relative wave heading (RWH), equivalent to the wave encounter angle, with RWH = 0° and RWH = 180° being following sea and head sea, respectively. As will be seen later, vessel forward speed is also included in the analysis to study its influence on the temporal variation in wave conditions as encountered by a ship. An overview of the input to the analysis is given by Table 1; noticing that the limiting values on the ranges are just indicative.

In total, data from 189 ships has been included in the analysis, and the time spans from January 2017 until February 2020. However, there is a large variation in the number of 'measurement points' from the different ships. Some ships contribute with less than a month of data while others contribute with many hundreds of days, see e.g. Fig. 2. On average, each ship brings data from about 380 days of operation into the analysis, where *operation* refers to all parts of shipping operations; that is, ocean transiting, quay stays, anchoring, manoeuvring, maintenance (dock stays), etc. In order to focus on sailing in seaways exclusively, data is simply included *only* if the forward speed is above 5 knots. This removes most of the non-seaway related data, including manoeuvring data in and around harbours, and, after additional removal of corrupted GPS data, it leaves the analysis with more than six million 10-min spaced observations; defining an *observation* to be that of an interpolated integral wave parameter, i.e. (H_s, T_z, D_m), from the ERA5 dataset. The route paths corresponding to the observations are seen in Fig. 1. Obviously, as realised by looking at the route paths and as described in the preceding, the six-plus million observations, corresponding to 1 million hours of observation, are not unique in time but they are in space–time.

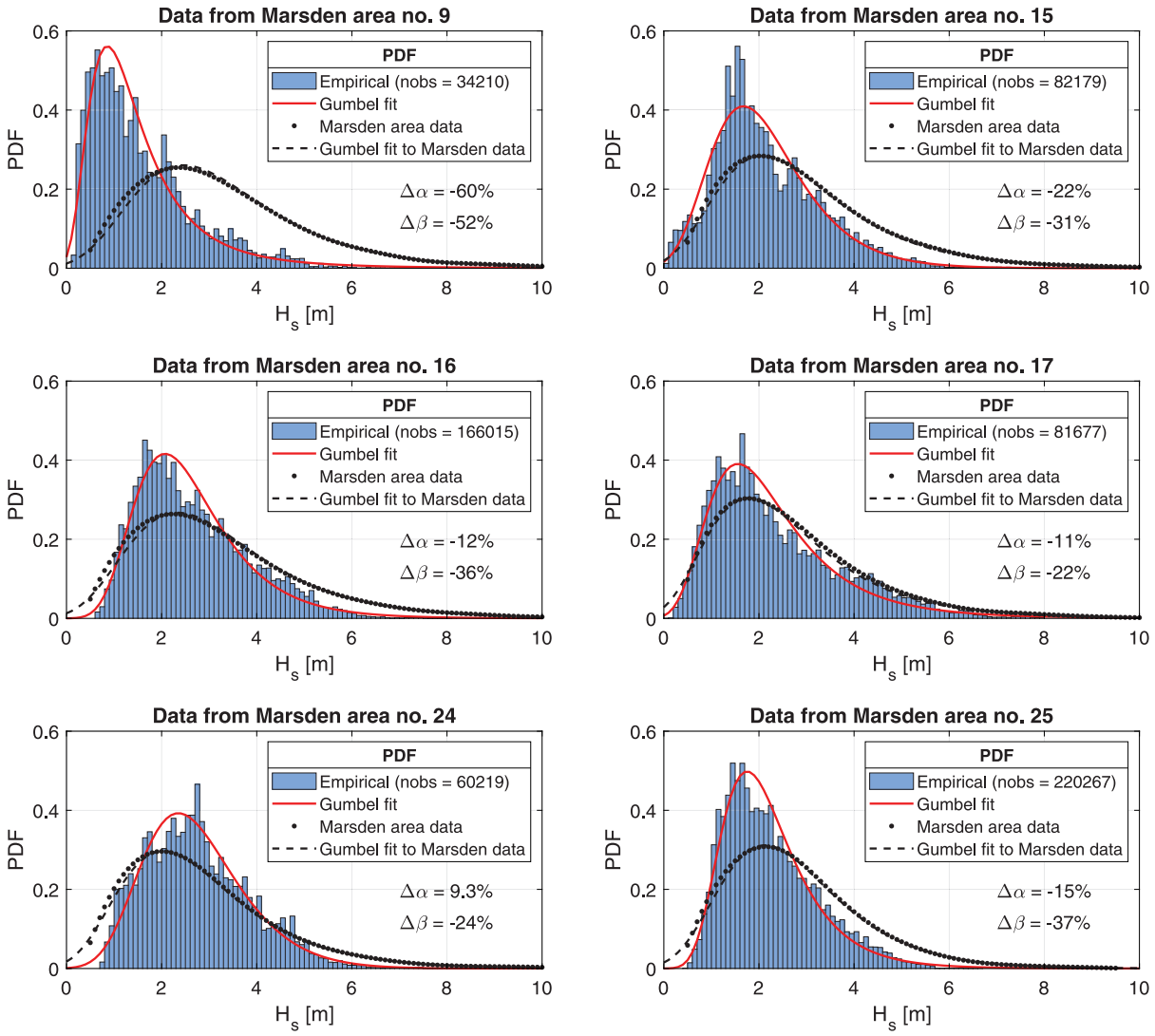


Fig. 5. Probability density functions of significant wave height for Marsden areas within the North Atlantic region. The relative deviations, $\Delta\alpha$ and $\Delta\beta$, in the Gumbel-parameters are printed in the plots, cf. Eq. (8).

Table 1
Input to the analysis: Type, description, sampling time, value range and units for dataset. Note that the ERA5 data is interpolated, spatially and temporally, to actual ship positions spaced by 10 min, as described in Section 2.2.

Type	Description	Sample time	Range	Unit
Time	UTC date-time	10 min.	01/01/2017–22/02/2020	[–]
Ship data	Latitude	10 min.	−90–90	[°]
	Longitude	10 min.	−180–180	[°]
	Compass heading	10 min.	0–360	[°]
	Speed	10 min.	0–29	[kn]
ERA5	Significant wave height	60 min. (10 min.)	1–9	[m]
	Mean wave direction	60 min. (10 min.)	0–360	[°]
	Zero-upcrossing period	60 min. (10 min.)	4–14	[s]

3.1. ERA5 data as 'observed data'

It was previously explained that the ERA5 reanalysis data is considered as the ground truth, see Section 2.5. As a consequence of this assumption, the encountered wave condition, as obtained from the ERA5 dataset, is referred to as 'observed data' and/or 'actual observation', although in the strict sense it should be termed 'reanalysis data'. The use of the term *observed*, or *observation*, is a choice made to reflect that it is data which is applicable to the exact time and position of any ship; like the data was measured or observed by a ship-board sensor (Thornhill and Stredulinsky, 2010; Nielsen, 2017) with a 10-min spacing between the measurements.

4. The encountered wave conditions

4.1. Overview and descriptive statistics

Fig. 4 presents how the encountered wave conditions are distributed, when data from all ships is considered. Fig. 4a shows the variation with Marsden areas, cf. Fig. 3, while Fig. 4b shows the

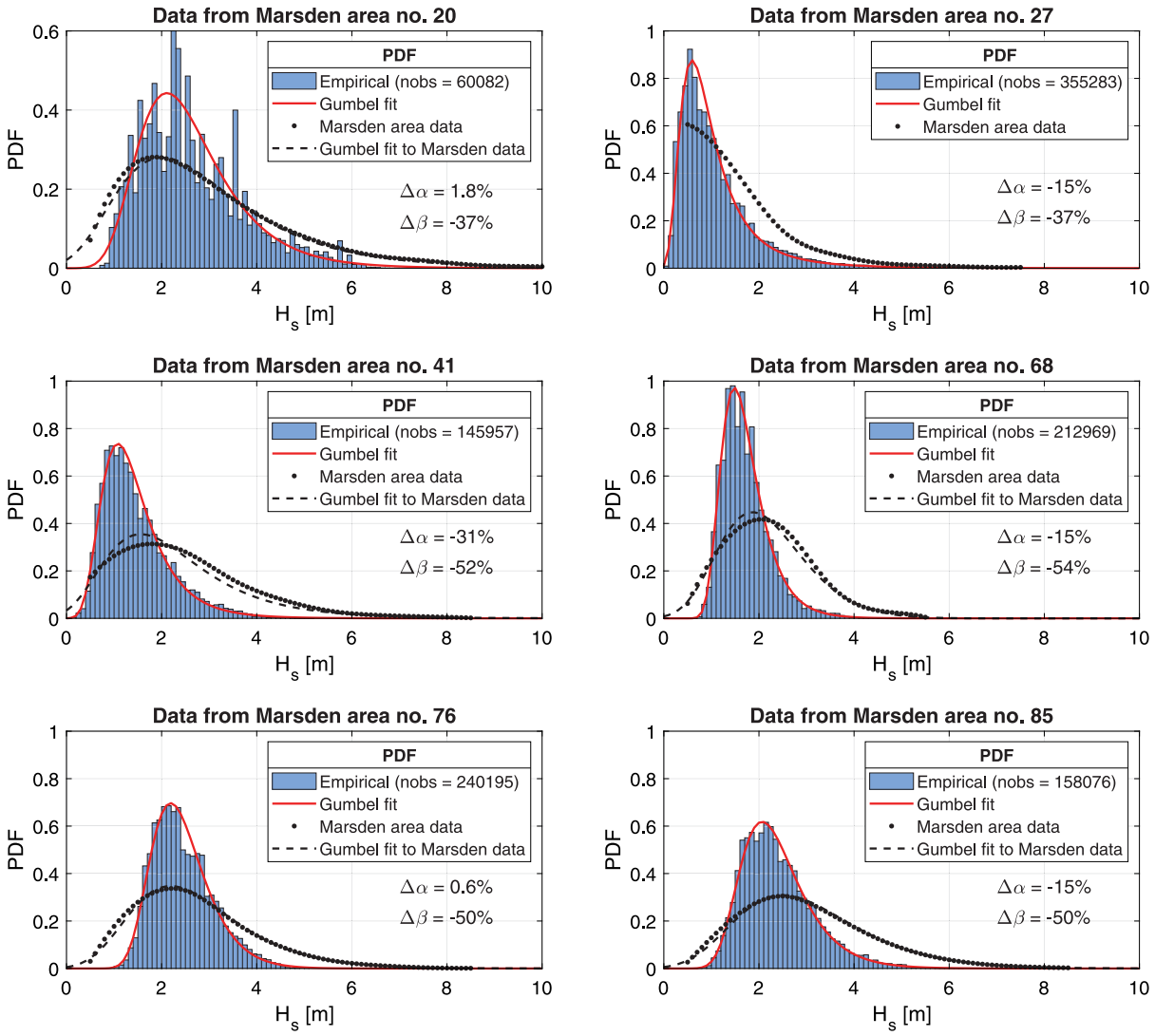
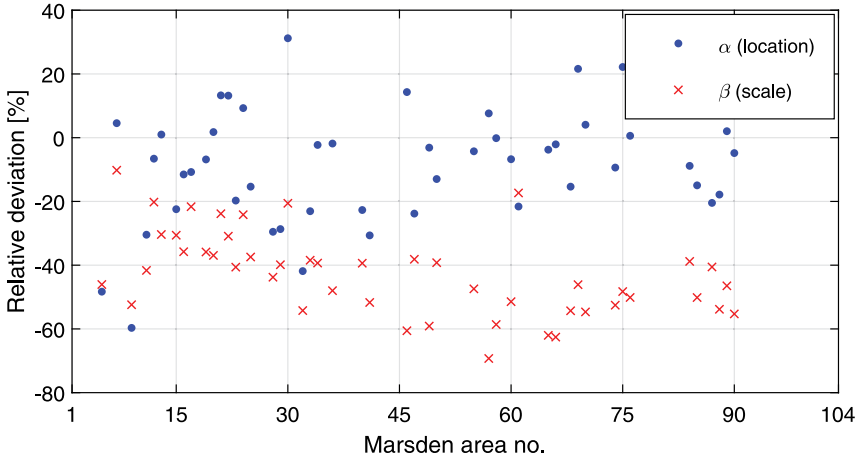


Fig. 6. Probability density functions of significant wave height for different regions of the world's oceans, cf. Fig. 3. Note the difference on the vertical scale. The relative deviations, $\Delta\alpha$ and $\Delta\beta$, in the Gumbel-parameters are printed in the plots, cf. Eq. (8).

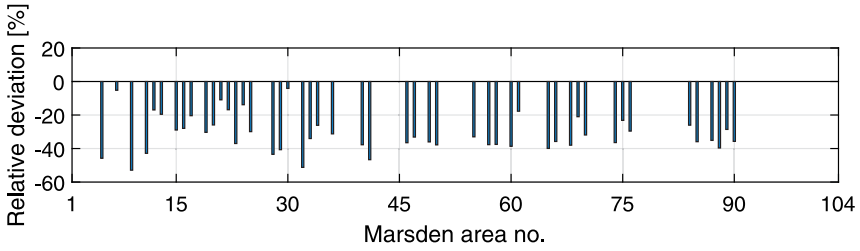
Table 2
Statistics of the wave parameters characterising the encountered sea states. Statistics of the actual scatter diagrams (British Maritime Technology, 1986) are included in parentheses.

Ocean (Mars. no.)	N. Atl. (9,15,16,24,25)	S. Atl. (67,68,84)	N. Pac. (13,20,21,30)	Ind. O. (60,61,69,70)	World (1–104)
N. obs.	562,890	240,473	139,788	752,830	6,023,263
Significant wave height [m]:					
Mean	2.36 (3.01)	1.79 (2.33)	2.79 (2.91)	1.57 (1.96)	1.64
StD	1.11 (1.72)	0.57 (1.03)	1.09 (1.65)	0.73 (1.06)	0.97
Max	8.20 (11.5)	6.12 (6.50)	7.00 (10.5)	5.08 (7.50)	8.20
p95	4.50 (6.50)	2.86 (4.50)	4.85 (6.50)	2.80 (3.50)	3.49
Zero-upcrossing period [s]:					
Mean	6.18 (8.21)	6.16 (7.58)	6.51 (8.13)	5.73 (6.90)	5.16
StD	1.59 (1.46)	1.15 (1.54)	1.31 (1.46)	1.56 (1.52)	1.65
Max	13.0 (13.5)	13.1 (13.5)	12.2 (13.5)	12.9 (13.5)	13.2
p95	9.04 (10.5)	8.46 (10.5)	8.90 (10.5)	8.29 (9.50)	8.03
Mean wave direction [°]:					
Mean	303	189	270	189	N/A
StD	61	26	71	67	N/A
Max	N/A	N/A	N/A	N/A	N/A
p95	N/A	N/A	N/A	N/A	N/A

variation with speed. The single plots are 2D histograms, which means that it is possible to infer how many observations of a given wave parameter, H_s (top), T_z (middle) and D_m *or* RWH (bottom), that have a specific value. It can be appreciated that ships sail (very) fast only in



(a) Deviation in location and scale parameters of the Gumbel distribution.



(b) Deviation in 95-percentile of significant wave height.

Fig. 7. Relative deviation between observations and scatter diagrams, cf. Eq. (8), when the PDFs are fitted with the Gumbel distribution.

good weather. It is also realised that the most significant speed range is 15–20 knots. Moreover, it is evident that ships tend to either sail directly against or with the waves, reflected by head sea (RWH = 180°) and following sea (RWH = 0°), respectively.

Descriptive statistics of the encountered wave conditions are presented in Table 2. The table includes results for different ocean regions: the North Atlantic ('N. Atl.'), the South Atlantic ('S. Atl.'), the North Pacific ('N. Pac.'), the Indian Ocean ('Ind. O'), and all data ('World'). The particular regions are associated with sets of Marsden areas, as selected/defined by the authors with reference to Fig. 3. The total number of observations ('N. obs.') within each region is specified, and then follows statistics for significant wave height, zero-upcrossing period, and mean wave direction; noting that the statistical metrics were defined in Section 2.4. It is clear from the table that the observed sea states is significantly less severe than prescribed by the standard wave scatter diagrams suggested by British Maritime Technology (1986).

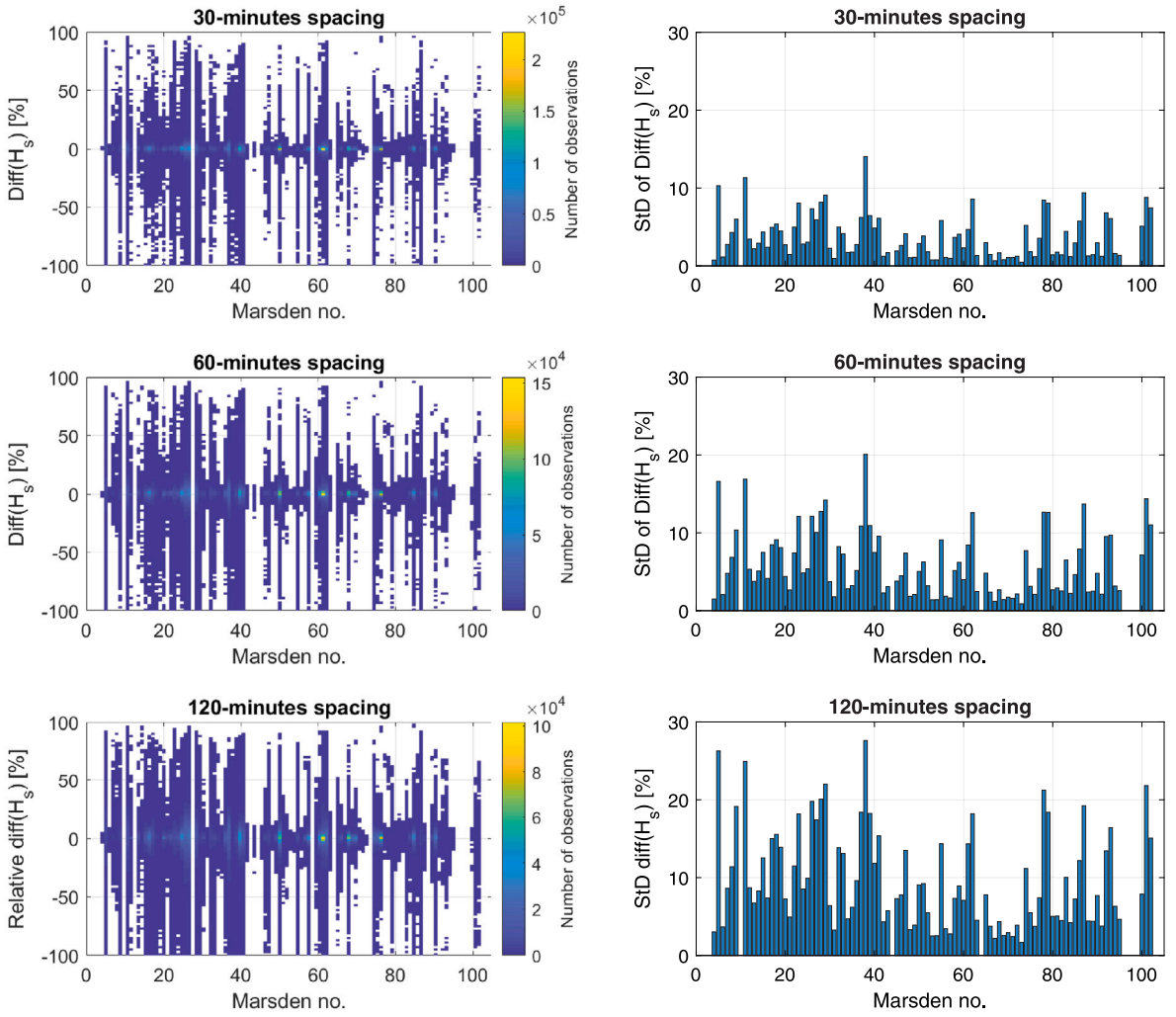
4.2. Encountered wave height distributions

Figs. 5 and 6 show samples of the probability density function (PDF) for the significant wave height as encountered in some selected Marsden areas. The individual plot includes the empirical distribution based on the actual observations within the given Marsden area. The empirical distribution is fitted with a Gumbel distribution. It can be seen that the Gumbel distribution generally fits the empirical distribution nicely. The plots in Figs. 5 and 6 present also the "empirical" PDFs together with associated Gumbel-fits (if obtainable) of the scatter diagrams (British Maritime Technology, 1986) for the particular Marsden areas.

Visual assessment of PDFs obtained in all Marsden areas from where observations exist is not practical. On the other hand, an evaluation can be made by comparing the location and scale parameters of the Gumbel distribution, cf. Eq. (7). The relative deviations in the location and scale parameters have therefore been computed:

$$\Delta\rho = \frac{\rho_{\text{obs}} - \rho_{\text{mrs}}}{\rho_{\text{mrs}}} \cdot 100\% \tag{8}$$

where ρ represents the location or scale parameter, α and β, respectively, coming from observations ('obs') or the scatter diagram ('mrs'). The numbers are printed in the plots in Figs. 5 and 6, and the meaning of the parameter names can easily be inferred; α refers to the *location* of the peak on the horizontal axis (H_s-value), while β is a measure of the peak level or *scale* in the sense that the higher β, the more narrow a distribution. An overall assessment of the relative deviations for all Marsden areas is shown in Fig. 7a; noting however that results are only computed for the areas where more than 10,000 observations are available within the particular Marsden area, to ensure statistically reliable results. Like indicated from the visualisations of the PDFs, Fig. 7a shows that there is a reasonable agreement in α with both positive and negative deviations; that is, the peak of the fitted Gumbel PDF of the observations is located at both smaller and larger H_s than prescribed by the wave scatter diagrams. On the other hand, the scale parameter β is consistently smaller for the fitted Gumbel PDF of the observations, which means that the observations generally represent a range of significant wave heights distributed in a more narrow interval. The further consequence of this is that the extreme values of significant wave height, as derived from the fitted PDFs for the observations, is smaller than what the scatter diagrams predict. This is easily confirmed by inspection of Fig. 7b that presents the deviation between the 95-percentile of H_s computed from the fitted Gumbel PDF of the observations relative to the fitted PDF of the scatter diagrams. The relative deviation is calculated in accordance with Eq. (8). It is seen that the scatter diagrams generally lead to significantly larger 95-percentiles.



(a) 2D histograms showing variation with Marsden areas.

(b) Standard deviation of the bins in the 2D histograms.

Fig. 8. Relative difference in significant wave height, as computed by Eq. (9), when the sailing distance between consecutive waypoints is spaced by $\Delta t = 30$ min (top), $\Delta t = 60$ min (middle), and $\Delta t = 120$ min (bottom). Note the difference in the colour scale, counting the number of observations, in the 2D histograms. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.3. Variation in significant wave height at consecutive waypoints

This subsection investigates what level of variation in significant wave height that occurs as a result of evolvement in both space and time, considering a given ship advancing on its route. Specifically, the wave height observed at a waypoint at time t_k is compared with the wave height observed at a waypoint at a later time $t_k + \Delta t$, emphasising that any variation results partly because of the evolvement Δt in time, and partly because the ship moves a distance $\Delta s = U \cdot \Delta t$ dependent on the ship's speed U. The formula of the relative difference[1] therefore reads, intentionally using Diff() to use another symbol than used with Eq. (8),

$$\mathrm{Diff}(H_{s,k}) = \frac{H_{s,k+m} - H_{s,k}}{H_{s,k}} \cdot 100\% \tag{9}$$

where index $k + m$ represents a waypoint at a time later than the waypoint represented by index k; the value of m is set in accordance with Δt. In the following, three situations are considered, thus the consecutive waypoints are spaced by the sailing distance during $\Delta t = 30$ min, $\Delta t = 60$ min, and $\Delta t = 120$ min. It should be stressed that the deviation between waypoints, as defined by Eq. (9), obviously is computed on a ship-specific basis.

Fig. 8a shows how the relative difference, all observations considered, varies with the Marsden areas, noting that the single plots correspond to different Δt and that the plots are 2D histograms. A more specific evaluation of the relative difference is shown in Fig. 8b which is presenting the standard deviation, as computed for the observed differences in the given Marsden areas. That is, the left-side and right-side plots in Fig. 8 correspond as pairs. In general, it can be appreciated from Fig. 8 that the relative difference in encountered significant wave heights can be substantial between consecutive waypoints, even in just 30 min.

It could be anticipated that vessel advance speed will be quite influential on the variation in significant wave height between consecutive waypoints, but Fig. 9 shows that the relative difference is not highly dependent on advance speed for the given data. However, an important point in this connection is that ships generally sail relatively faster in good weather and low sea states. Furthermore, it is presumed that such conditions develop more gradual than those represented by "medium" weather and higher sea states. Consequently, the most logical reasoning, saying that high(er) speed naturally leads to large(r) relative differences in observed wave heights, is not true, since ships mostly

[1] Note that *difference* is used, as this term is considered more appropriate than *deviation*.

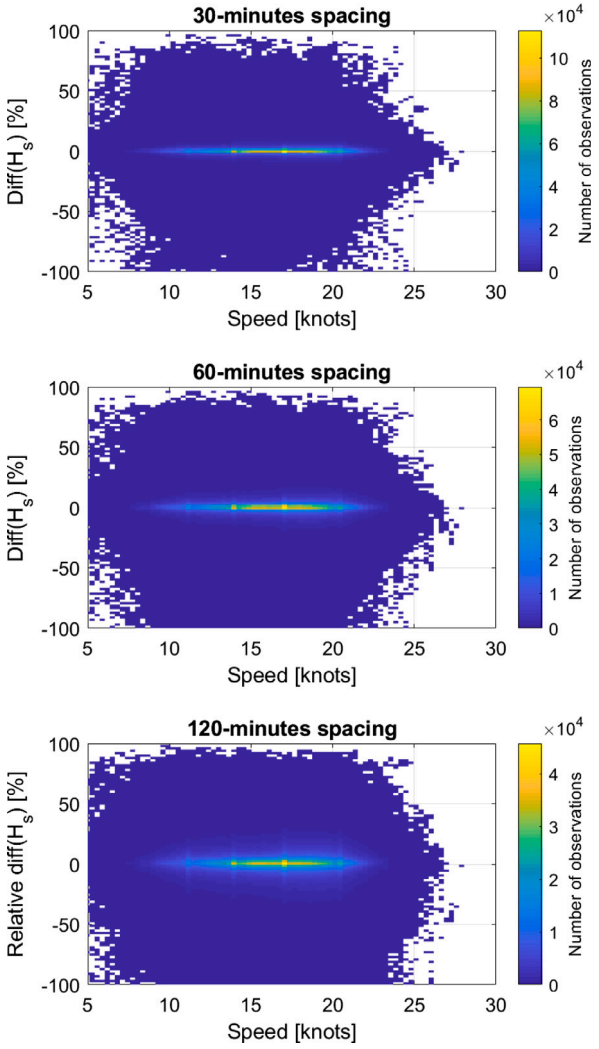


Fig. 9. Relative difference in significant wave height (Eq. (9)) and the dependency with advance speed. Note the difference in the colour scale, counting the number of observations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

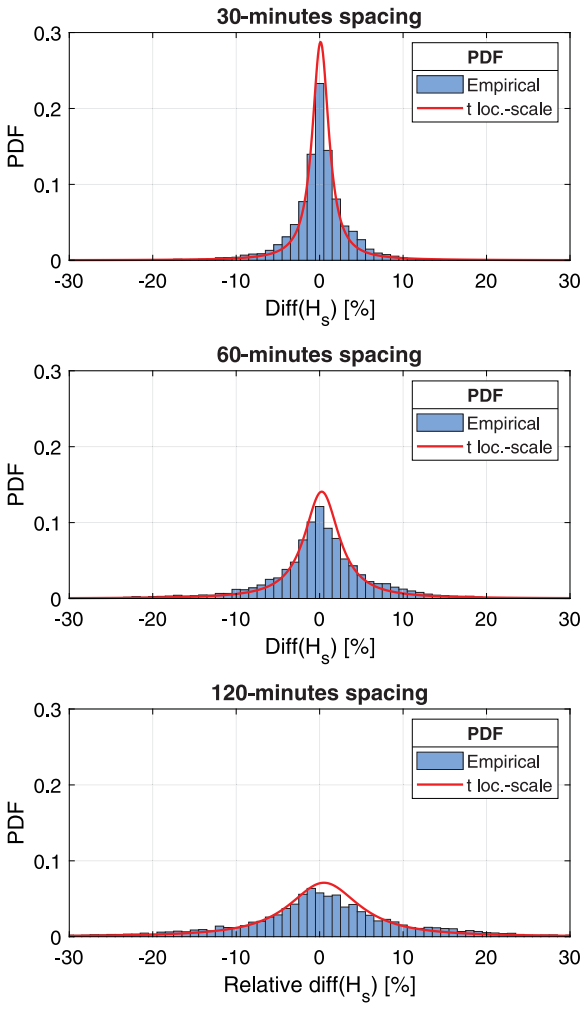


Fig. 10. Probability density functions of the relative difference in significant wave height (Eq. (9)) when the distance between consecutive is spaced by Δt = 30 min (top), Δt = 60 min (middle), and Δt = 120 min (bottom).

only sail fast in a calm sea. Obviously, this point cannot be generalised but will be specific to the particular shipping company, depending on for instance, company operational strategy, type of weather routing, seamanship of the ship master, etc.; all points that could be further investigated in a future study.

Despite the dependency with Marsden areas and speed, as illustrated in Figs. 8 and 9, an overall evaluation of the observed relative differences in encountered wave heights is presented in Fig. 10, with results calculated for the different spacings between waypoints as controlled by Δt. The individual plots show the empirical PDF and a fitted PDF based on the t location-scale distribution that has been found previously to give good results for modelling spatio-temporal variation in sea states (Nielsen, 2021). The t location-scale distribution is useful for modelling data distributions with heavier tails than the normal distribution, and the distribution is controlled by the probability density function (e.g. MATLAB R2020b, 2020):

$$f_t(x|\mu,\hat{\sigma},\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sigma\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left[\frac{\nu+\left(\frac{x-\mu}{\sigma}\right)^2}{\nu}\right]^{-\left(\frac{\nu+1}{2}\right)} \tag{10}$$

where $\Gamma(\cdot)$ is the gamma function, μ is the location parameter, σ is the scale parameter, and ν is the shape parameter.

It should be repeated that the results in this subsection and thus also the PDFs in Fig. 10 have been calculated by, first, evaluating Eq. (9) for each ship individually, and then putting together all results as a whole. As an interesting exercise left for future work it should be relevant to study detailed PDFs applicable to different ocean regions.

The statistics corresponding to Fig. 10 are given in Table 3, where RMSE is the root mean square error and p95 is the 95-percentile; noting that both numbers are relative, see Eq. (9), but the dimensional value is included in parentheses. Intentionally, the maximum number, i.e. `max()` and derived as a single number from the entire population, is *not* included, since this number will not necessarily reflect the real physics of the variation in H_s, observed while sailing in ocean areas. Thus, the number could be affected by arrivals/departures from sheltered areas. Notwithstanding, it can be mentioned that the maximum difference, in dimensional form, takes the values 2.1 m (30 min), 2.2 m (60 min), and 2.5 m (120 min), for all considered data but without knowing if the numbers are from the same event.

It is interesting to note that, although the present analysis does not differentiate between the various ocean regions, the outcome is fairly consistent with the findings given by Nielsen (2021), albeit the referred study was based on virtual shipping routes for specific ocean regions but with no account for weather routing. This observation suggests that, while weather routing clearly lead to the encountering of less severe wave conditions compared to existing scatter diagrams, cf. Section 4.2,

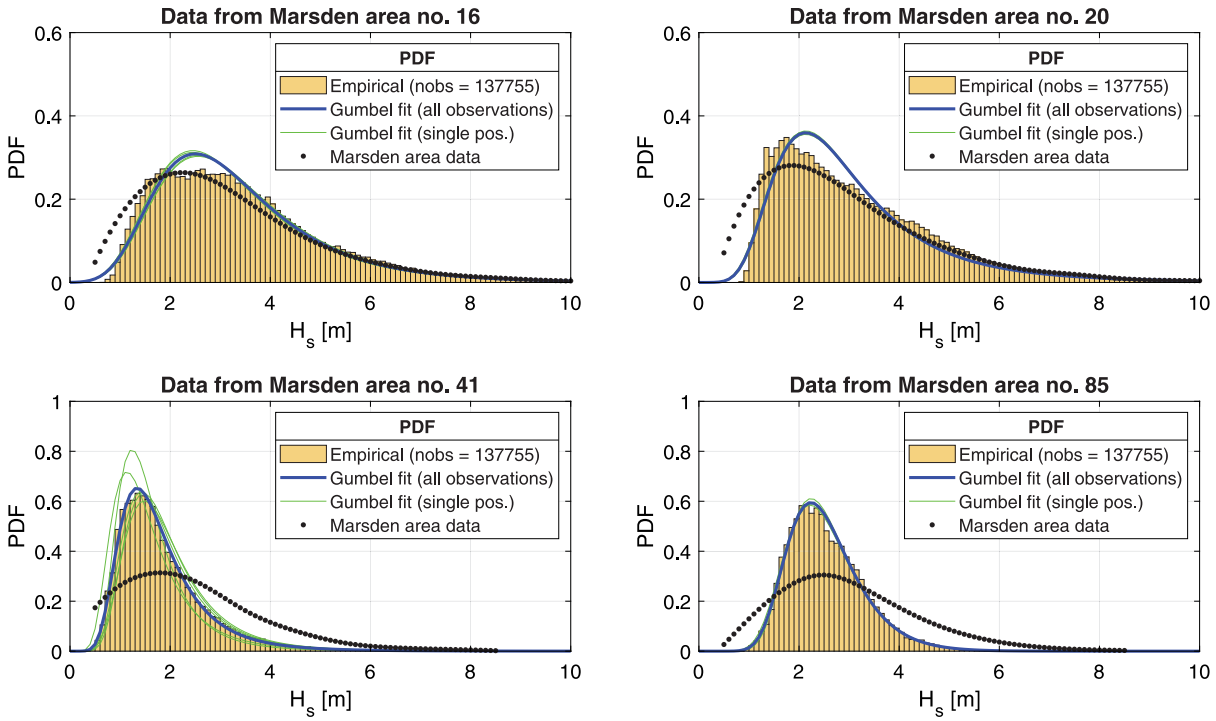


Fig. 11. Probability density functions of H_s for fixed positions within given Marsden areas. Note the difference in scales on the vertical axes. Intentionally, the colours are different compared to Figs. 5 and 6.

Table 3
Statistics of the relative difference [%] in encountered significant wave height (Eq. (9)).

Δt	30 min	60 min	120 min
t **location-scale parameters**:			
μ	0.11	0.25	0.52
σ	1.15	2.38	4.76
ν	1.25	1.33	1.43
Statistics of Diff(H_s):			
RMSE	5.42 (0.07 m)	8.67 (0.12 m)	13.9 (0.19 m)
p95	8.28 (0.11 m)	16.1 (0.21 m)	29.7 (0.39 m)

Table 4
Latitudes and longitudes for some fixed positions within different Marsden areas.

Mars. no	Region	Positions (Lon,Lat)				
		1	2	3	4	5
16	North Atl.	(−38°,44°)	(−38°,46°)	(−36°,44°)	(−36°,46°)	(−37°,45°)
20	North Pac.	(162°,44°)	(162°,46°)	(164°,44°)	(164°,46°)	(163°,45°)
41	Indian O.	(122°,22°)	(122°,24°)	(124°,22°)	(124°,24°)	(123°,23°)
85	South Atl.	(12°,−30°)	(12°,−28°)	(14°,−30°)	(14°,−28°)	(13°,−29°)

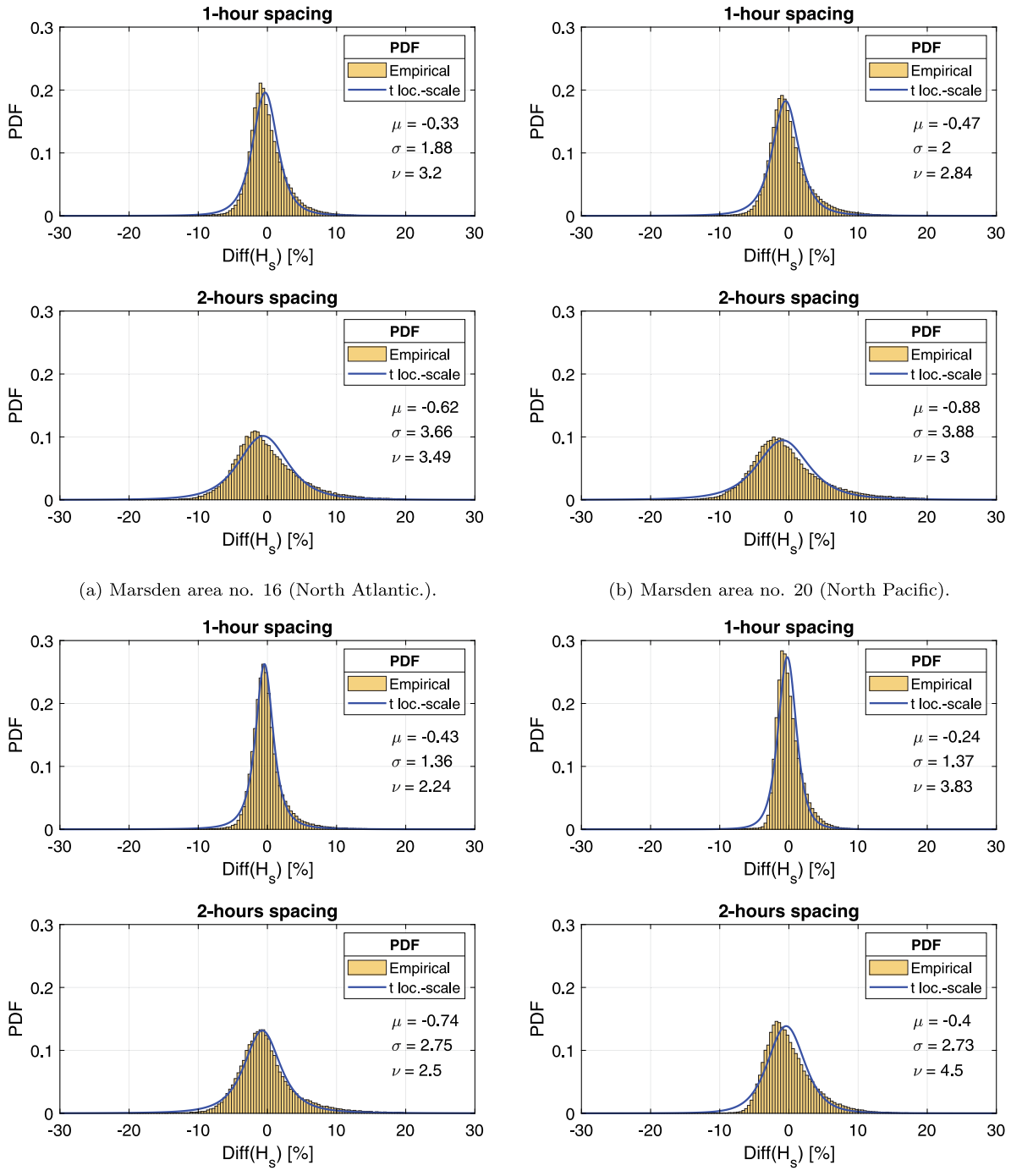
the *variation* in encountered sea state is not affected to the same degree by the use of weather routing systems.

5. Discussions

5.1. Encountered wave conditions

It seems relevant to ask the question whether the discrepancy between the actual encountered wave conditions and the conditions prescribed by the wave scatter diagrams (British Maritime Technology, 1986) is solely a consequence of weather routing and good seamanship. In the strict sense, this is an almost impossible question to answer, as it probably will require much more data than studied herein. Nonetheless, indications can be found by replicating the analysis made in Section 4 but for fixed geographical positions. In this case, Fig. 11 presents the findings. The plots in the figure reproduce the PDFs of the "encountered" wave height distribution, as obtained for some selected Marsden areas; similar to the plots in Figs. 5 and 6. However, it is emphasised that, in this case, fixed positions are considered. Specifically, the single plots are the result of an analysis where five positions are considered, within the given Marsden area. The five positions are taken to form a square, with the four corners spaced by 2 degrees (lon, lat), and one centre-position at the intersection of the diagonals. Relevant information is outlined in Table 4 and it is noted that the considered period used for this particular sub-study is identical with the period from which the ship data has been collected, that is, 01 Jan. 2017 to 22 Feb. 2020, cf. Section 3. Contrary to the former analysis with ship data, the plots in Fig. 11 are made from hourly observations corresponding to the exact hourly updates in the ERA5 dataset. For the considered period, this gives 27,551 observations for each position, resulting in totally 137,755 observations for all five positions in a given Marsden area.

The plots in Fig. 11 leave no definitive answer to the question raised above. It can be seen that there is a fairly good match between the wave height distribution from the observations (ERA5) and the scatter diagram in Marsden areas no. 16 and 20 (North Atl. and North Pac.), although it could be argued that the observations, in fact, indicate an increase in any prediction of extreme values. In a context of ship operations, the relatively good match between the observations and the scatter diagram suggests that the discrepancy observed previously (Fig. 5) is a result of weather routing and good seamanship. However, the findings from the other Marsden areas (no. 41 and 85) in Fig. 11 are more blurred: While the match between the empirical (and fitted) distribution of observed wave heights (ERA5) and the distribution based on the wave scatter diagram is marginally better than previously seen, the conclusion is that the scatter diagrams for the particular Marsden areas lead to prediction of significantly larger extreme values. In these cases, it therefore cannot be concluded that weather routing and good seamanship alone is responsible for the encountering of lower sea states than prescribed by the wave scatter diagrams.



(a) Marsden area no. 16 (North Atlantic.).

(b) Marsden area no. 20 (North Pacific).

(c) Marsden area no. 41 (Indian Ocean).

(d) Marsden area no. 85 (South Atlantic).

Fig. 12. Probability density functions of the relative difference in significant wave height (Eq. (9)) resulting between 1-h and 2-h updates, as observed for five fixed positions in the given Marsden areas.

The above discussion calls for a much more comprehensive analysis, and also calls for an even larger dataset, before solid conclusions can be drawn. This is beyond the scope of the present study; which, in the most basic sense, merely aims to give an account of the wave conditions *actually* encountered by ships sailing in the larger oceans. In the event of such a future work, it would also be important to recognise that, although the Global Wave Statistics (GWS) and the associated scatter diagrams (British Maritime Technology, 1986) are considered reliable for the northern oceans, even with good fits in the tail, the GWS have not been well validated along shipping routes in the southern oceans and at low latitude, as reported by Leenaars et al. (2000).

It seems relevant to mention that, albeit the present study shows that weather routing and seamanship lead to the encountering of less severe sea states for ships in the Maersk Line fleet, the effect is smaller than what was reported by Olsen et al. (2006), where focus was on both tankers, bulk carriers, and container ships but limited to "only" 25,000 observations. Several reasons could potentially explain the discrepancy; to mention some: (a) In the present study, much more data is included, herein 1 million hourly observations; (b) development in wave spectral models and the amount of assimilated measurements, leading to improved reanalyses, herein ERA5; (c) the use of auto-logged

GPS data, as in the present study, giving more reliable information about the position of the ships at reported times.

5.2. Temporal variation in wave conditions

The results in Section 4.3 clearly indicate that the variation in sea state, as encountered along consecutive waypoints of sailing ships, can be relatively large. The observed variation accounts for developments resulting because of the change in *both* time and position. If the position instead is fixed, the corresponding results are shown in Fig. 12. The plots present the relative difference, cf. Eq. (9), in significant wave height for the fixed locations specified in Table 4. Each plot contains the data from all five positions corresponding to a given Marsden area, and the title of the plot informs whether the temporal spacing is 1 h or 2 h.

It is noticed that the t location-scale distribution again fits reasonably well in all cases. The fitting parameters are printed in the single plots of Fig. 12. A comparison to the PDFs of the sailing ships, cf. Fig. 10 and Table 3, reveals that the variation with little surprise is larger when it results because of a change in both position and time, nonetheless the variation can be substantial also for fixed positions. As a final remark, it is interesting to note the positive skewness of the empirical PDFs shown in Fig. 12. The skewness implies a slight deviation relative to the t location-scale distribution in all cases but is the most pronounced for the 2-h spacings. This skewness is a result of the fact that waves, and thus (significant) wave height, grow faster than they decay during the coming and going of storms. It should be realised that skewness is not visible in the PDFs of the difference in significant wave height when sailing ships are considered, cf. Fig. 10. This is likely explained as a consequence of weather routing and good seamanship, meaning that ships tend to avoid being in the actual areas where storms are developing.

6. Final words and conclusions

This paper has given a report about the wave conditions encountered by ships in the fleet of Maersk Line, studying routes across the world. The data of the study originated from 189 container vessels, each one equipped with a continuous monitoring system. Based on measurements of GPS (longitude, latitude), heading, and forward speed, outcome from the ERA5 dataset (Hersbach et al., 2021) was merged with the ship data. As a result, more than six million 10-min spaced observations (= reanalysis data) of significant wave height, zero-upcrossing period, and relative wave heading have been considered. The study showed that weather routing and seamanship have the effect that the ships encounter less severe sea states than would be inferred from the Global Wave Statistics (British Maritime Technology, 1986), although the effect has not been seen to be as significant as reported in a previous, similar, but much smaller study (Olsen et al., 2006) with a focus on conditions specifically in the North Atlantic. To the credibility of the present study, compared to former ones, speaks not only the use of much more data, but also the application of continuous monitoring ship data (GPS and speed) together with reanalysis wave data (ERA5).

The paper also studied the temporal and spatial variation in encountered wave heights; here referring to *variation* as the difference in wave height observed between consecutive waypoints spaced from 30 min to 120 min along the sailing route of a ship. Although it is difficult to exactly define when conditions in practical applications are no longer stationary, the encountered variation in sea state is assessed to be at a level that sometimes might compromise the often imposed (theoretical) requirement that a *stationary process* is assumed, when wave-ship interactions are analysed. Above all, this study has reported about what wave conditions, container ships are expected to encounter during their daily operations in the seas around the world.

CRediT authorship contribution statement

Ulrik D. Nielsen: Conceptualization, Methodology, Software, Writing - original draft, Writing - review & editing, Investigation, Visualization. **Angelos Ikonomakis:** Data curation and collection, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank Dr. Jesper Dietz (Maersk Line) for assistance in connection with accessing data and for useful discussions. The work by the first author has been supported by the Research Council of Norway through the Centres of Excellence funding scheme, project number 223254 AMOS. The second author acknowledges the support by Den Danske Maritime Fond, Denmark (case number 2018-060) and Innovationsfonden, Denmark (case number 8053-00231B).

References

Belmadani, A., Dalphinet, A., Chauvin, F., Pilon, R., Palany, P., 2021. Global wave height trends and variability from new multimission satellite altimeter products, reanalyses, and wave buoys. Clim. Dynam. 56, 3687–3708.

British Maritime Technology, 1986. Global Wave Statistics (Primary Contributors N. Hogben and N.M.C. Dacunha and G.F. Oliver). Unwin Brothers Limited, Middlesex, UK.

Copernicus Climate Change Service Information, 2020. ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. https://cds.climate.copernicus.eu/cdsapp#!/home (accessed 08-01-2021).

ECMWF, 2017. Part VII: ECMWF Wave Model. Technical Report IFS Documentation – Cy43r3, European Center For Medium-Range Weather Forecasts, Shinfield Park, Reading, RG2 9AX, England.

Hauteclocque, G., Johnson, M., Zhu, T., Austefjord, H., Bitner-Gregersen, E., 2020. Assessment of global wave datasets for long term response of ships. In: Proc. of 39th OMAE, Fort Lauderdale, FL, USA.

Hersbach, et al., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. http://dx.doi.org/10.1002/qj.3803.

Hersbach, et al., 2021. ERA5 Hourly Data on Single Levels from 1979 to Present. Copernicus Climate Change Service: Climate Data Store (CDS), http://dx.doi.org/10.24381/cds.adbb2d47, https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview (accessed: 08-01-2021).

IACS-REC 034, 2001. Standard Wave Data (no. 034). International Association of Classification Societies, http://www.iacs.org.uk/download/1978 (accessed 09-03-2021).

IACS-REC 106, 2009. IACS Guideline for Rule Development - Ship Structure (no. 106). International Association of Classification Societies, http://www.iacs.org.uk/download/1903 (accessed 09-03-2021).

Jensen, J.J., 2001. Load and Global Response of Ships. In: Elsevier Ocean Engineering Book Series, vol. 4, Elsevier, Oxford, United Kingdom.

Leenaars, C., Louazel, S., Brugghe, P., 2000. Comparison of Wave Databases and Design Methods for Major Shipping Routes. Technical Report WP5100, EU project COMKISS, Meer, France.

MATLAB R2020b, 2020. t Location-scale distribution. https://se.mathworks.com/help/stats/t-location-scale-distribution.html (accessed: 28-12-2020).

Naseef, T.M., Kumar, V.S., 2020. Climatology and trends of the Indian ocean surface waves based on 39-year long ERA5 reanalysis data. Int. J. Climatol. 40, 979–1006.

Nielsen, U.D., 2017. A concise account of techniques available for shipboard sea state estimation. Ocean Eng. 129, 352–362.

Nielsen, U.D., 2021. Spatio-temporal variation in sea state parameters along virtual ship route paths. J. Oper. Oceanogr. http://dx.doi.org/10.1080/1755876X.2021.1872894.

Olsen, A., Schrøter, C., Jensen, J., 2006. Wave height distribution observed by ships in the North Atlantic. Ships Offshore Struct. 1, 1–12.

Stefanakos, C., 2019. Intercomparison of Wave Reanalysis based on ERA5 and WW3 Databases. In: Proc. of 29th ISOPE, Honolulu, HI, USA.

Takbash, A., Young, I., 2020. Long-term and seasonal trends in global wave height extremes derived from ERA-5 reanalysis data. J. Mar. Sci. Eng. 8, http://dx.doi.org/10.3390/jmse8121015.

The WAVEWATCH III Development Group, 2019. User manual and system documentation of WAVEWATCH III r version 6.07. https://github.com/NOAA-EMC/WW3/wiki/Manual.

Thornhill, E.M., Stredulinsky, D.C., 2010. Real time local sea state measurement using wave radar and ship motions. In: SNAME Annual Meeting, Seattle, WA, USA.

Timmermans, B., Gommenginger, C., Dodet, G., Bidlot, J., 2020. Global wave height trends and variability from new multimission satellite altimeter products, reanalyses, and wave buoys. Geophys. Res. Lett. 47, http://dx.doi.org/10.1029/2019GL086880.

Tolman, H., Balasubramaniyan, B., Burroughs, L., Chalikov, D., Chao, Y., Chen, H., Gerald, V., 2002. Development and implementation of wind-generated ocean surface wave models at NCEP. Weather Forecast. 17, 311–333.