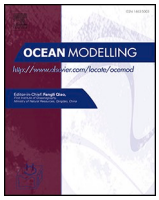
Ocean Modelling 188 (2024) 102341

Contents lists available at ScienceDirect

Ocean Modelling

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



SHyTCWaves: A stop-motion hybrid model to predict tropical cyclone induced waves

Sara O. van Vloten [a], Laura Cagigal [a,*], Beatriz Pérez-Díaz [a], Ron Hoeke [b], Fernando J. Méndez [a]

[a] Departamento de Ciencias y Técnicas del Agua y del Medio Ambiente, Geomatics and Ocean Engineering Group, E.T.S.I.C.C.P, Universidad de Cantabria, Santander, Spain
[b] Climate Science Centre, Commonwealth Science and Industrial Research Organization (CSIRO), Aspendale, VIC, Australia

ARTICLE INFO

Keywords:
Tropical cyclone
Metamodel
Vortex-type winds
Storm parameterization
Hybrid modeling

ABSTRACT

Waves produced by tropical cyclones (TCs) can be estimated using non-stationary wave models forced with time-varying wind fields. However, dynamical simulations are time and computationally demanding at regional-scale domains since high temporal and spatial resolutions are required to correctly simulate TC-induced wave propagation processes. Applications such as early warning systems, coastal risk assessments and future climate projections benefit from fast and accurate estimates of wave fields induced by close-to-real storm tracks geometry. The proposed SHyTCWaves methodology constitutes a novel tool capable of estimating the spatio-temporal variability of directional wave spectra produced by TCs in deep waters, using a hybrid approach and statistical techniques to reduce CPU time effort. This work demonstrates that TC-induced waves can be reconstructed using a stop-motion approach based on the addition of successive 6 h periods of time-varying storm conditions. The developed hybrid model reduces a TC track to a number of segments that are parameterized in terms of 10 representative TC features, and generates a library of cases dynamically pre-computed which allow to ensemble consecutive 6 h analog segments representing the original TC track. The metamodel has been compared and corrected with available satellite data, and its applicability is exemplified for TC Ofa in the South Pacific.

1. Introduction

Tropical Cyclones (TCs) are non-frontal, synoptic scale low-pressure systems driven by heat transfer from the ocean, and triggered by pre-existing weather disturbances above warm tropical or subtropical waters (Emanuel, 2003). These phenomena typically cause intense winds, large waves and extreme water levels which can derive into severe destruction and extensive damage along exposed coastlines and flood-prone areas. Spectral (phase-averaged) numerical wave models forced with wind fields are a standard tool to simulate spatio-temporal waves associated with TC events. For very intense low-pressure systems, the main vortex structure of the winds seems to solely govern the spatial distribution of the waves and their associated directional characteristics (Kudryavtsev et al. 2021). However, the physical processes of wave generation and non-linear interactions with rapidly changing wind conditions are complex, and such models are very demanding in terms of computational time and resources. Thus, efforts have been made to develop parametric models which are simpler and appropriate to predict wave induced conditions. Hasselmann et al. (1976) suggested that the main principle for the construction of a parametric model is that the energy and momentum sources can reproduce empirical fetch laws for idealized cases (i.e., constant wind forcing condition). Holland (1980) proposed an analytic model for the radial profiles of sea level pressures and winds. Young (1988) conducted numerical experiments to determine the influence of the wind field parameters to the resulting spatial distribution of the maximum significant wave height and its associated spectral peak frequency based on the concept of the equivalent fetch. The advent and subsequent regular collection of wind and wave field observations by satellite in recent decades, has confirmed that both wind and wave fields around TCs are asymmetric, with stronger winds/larger waves to the right of the eye of the storm (Tamizi and Young, 2020, Cagigal et al., 2022) in the northern hemisphere (left in the southern hemisphere). Recently, Cagigal et al. (2022) developed an empirical model to estimate the wind wave footprint of TCs taking advantage of the increasing amount of remote sensing observational data. Moreover, the wave field distribution is more asymmetric than the wind field since

* Corresponding author.
E-mail address: laura.cagigal@unican.es (L. Cagigal).

https://doi.org/10.1016/j.ocemod.2024.102341
Received 25 August 2023; Received in revised form 19 December 2023; Accepted 14 February 2024
Available online 15 February 2024

large waves are generated with a necessary substantial fetch over which the intense winds are blowing. This concept of extended fetch was used to map the distribution of wave energy, peak frequency and direction along wave-rays (Kudryavtsev et al. 2021) dependent on sets of TC parameters (wind speed, radius of maximum winds, ratio of wind speed to the TC forward velocity). Recently, Grossmann-Matheson et al. (2023) developed a parametric model to predict TC wave height in terms of a synthetic TC wave field database.

These empirical approaches aim to provide a simpler and cost-effective alternative to full spectral wave models, however they are generally incapable of estimating full time-varying directional wave spectra at a level of fidelity important to many applications, particularly further dynamic coastal downscaling. To overcome these limitations another conceptual approach, a new parametric hybrid modelling approach is introduced in this study, for when the use of full dynamic (numerical) wave models for ensembles of large number of TC-wave simulations for the estimation of coastal hazard becomes impracticable. In that context hybrid models have emerged as powerful computational tools that enable a comprehensive coastal and risk hazard assessment (Irish et al., 2009; Jia et al., 2016). Several studies have shown that these hybrid models provide fast-to-compute data driven approximations quantifying the TC induced storm surge (Kyprioti et al. 2021, Jia et al. 2016, Kim et al., 2015) and hydro-meteorological conditions (Bakker et al., 2022). Results have shown that when properly calibrated, they are capable of replacing with a high level of accuracy, the high-fidelity numerical model used to create the original database, while maintaining the detailed underlying representation of hydrodynamic processes (Resio and Westerink, 2008) at a significantly lower computational cost. One key feature of these hybrid methods is their focus on reducing the number of simulation events that are necessary to represent accurately a wide range of plausible scenarios (Resio et al., 2009, Jia and Taflanidis, 2016, Bakker et al., 2022, van Vloten et al., 2022) by taking advantage of statistical techniques to conduct a smart selection of cases to build a library of precomputed high-fidelity simulations, and the subsequent reconstruction for any given TC. This approach also involves a reduction of the high-dimensionality of the original problem into a simplified parametric space of relevant variables, thus the original storm tracks can be simplified (van Vloten et al., 2022).

Here, we use a hybrid model to generate estimations of directional wave spectra induced by close-to-real storm tracks geometry, by using high-fidelity simulations of a statistically representative set of parameterized storm segments, and a simplified version of the dynamic downscaling based on a stop-motion approach which preserves the physics of wave propagation processes. This approach allows to obtain the directional wave spectra produced by a TC at regional scales, which is relevant as it can serve as much more accurate hydraulic boundary conditions for downscaling to nearshore coastal areas (e.g. for coastal hazard warnings or analyzes for engineering applications) compared to earlier parametric and wind-fetch based approaches. Particularly for island territories impacted by TCs and surrounded by deep ocean waters, the characterization of the directional wave spectra is essential to take into account multimodal energy systems. Thus, the proposed methodology is suitable to be implemented at such target locations, and will be exemplified in this study for the archipelago of Samoa.

This paper is structured as follows. Section 2 describes the sources of data employed at different stages of the study and the numerical wave model. Section 3 presents the method of the proposed stop-motion approach, while Section 4 describes the application of the SHyTC-Waves methodology. Section 5 focuses on the bias correction of the hybrid model using wave altimeter data. Section 6 presents a summary and conclusions.

2. Data and methods

2.1. Tropical cyclones

For the purpose of generating a worldwide database of storm tracks, the International Best Track Archive for Climate Stewardship (IBTrACS, Knapp et al. 2010, 2018) global dataset was developed by the NOAA National Climatic Data Centre, and it has been used here to extract historical storm track data across all ocean basins. In order to obtain the largest possible population of historical storm tracks, both the official records reported by the responsible agency in each location provided by the World Meteorological Organization (WMO), as well as records from all Regional Specialized Meteorological Centres (RSMCs) have been used. The IBTrACS version 04r00 database (https://www.ncei.noaa.gov/products/international-best-track-archive) includes the longitude, latitude, central pressure, maximum wind speed and, in the last two decades, the radii of maximum winds. Given that RSMCs employ different procedures to compute the maximum sustained winds, a conversion factor is applied to obtain the estimated 1 min average maximum sustained winds. Harper et al. (2010) recommended 0.93 to convert from 10 min to 1 min at sea conditions instead of the traditional 0.88 factor.

2.2. Satellite wave data

In this study, satellite wave data was extracted from the Integrated Marine Observing System (IMOS) database (Ribal and Young, 2019) which is intended to be updated every six months, as an ongoing surface altimeter database for wave height and wind speed available through the Australian Ocean Data Network portal (AODN: http://portal.aodn.org.au/). Since 1985, data has been obtained from 13 altimeters, and it is provided globally with 1° grid resolution; this altimeter database has been calibrated and validated against buoy data from the National Oceanographic Data Center (NODC), as well as cross-validated between altimeters in order to guarantee consistency. The whole dataset has been used from 1985 to 2020, except for flagged outliers that were discarded for the purpose of the present study.

2.3. Numerical modelling

The spectral evolution and propagation of wave energy in time and space due to the passage of a storm is predicted using the Simulating Waves Nearshore (SWAN) third-generation model (version 41.31) with non-stationary runs forced with time-varying parametric vortex-type storm wind fields over the numerical domain. SWAN simulations were run on a 15km resolution mesh, storing wave results at spatially gridded nodes. The model configuration includes a variable time step (10 or 20 min) depending on the storm's mean forward speed (higher or lower than 20km/h respectively); the direction space is discretized into 72 sectors (5° per sector); the frequency space ranges from 0.03 to 1Hz discretized into 34 bins on a logarithmic scale. White-capping, quadruplets and triads are activated. The wind source term implementation is set to ST6 physics package (Rogers et al. 2012) and the wind drag coefficient is capped at 50m/s according to Hwang et al. (2011) to prevent overestimation due to very intense wind conditions during TCs.

The bathymetric data used in the numerical simulations performed for the control case defined in the proof of concept (section 3) is the GEBCO (General Bathymetric Charts of the Ocean, http://www.gebco.net) global bathymetric dataset. The GEBCO's gridded bathymetric dataset (GEBCO_2019 grid) is a global terrain model for ocean and land providing elevation data on a 15 arc-second interval grid, equivalent to 450m spatial resolution.

The storm wind fields associated to any given set of storm conditions in the proposed methodology are defined by a wind parametric model which requires a small number of input variables. Here, we used the Dynamic Holland Model which is based on the vortex-type model

developed by Holland (1980) to estimate maximum wind speed using an analytic model of the sea level pressure and wind profiles. Moreover, we included the modification implemented by Fleming et al. (2008) which accounts for estimated winds at surface level due to dynamically moving changes of the TC parameters along the track, rather than the original model which was parameterized to fit the TC instantaneous gradient wind level. The spatial distribution of sea surface wind fields is thus estimated from the forward speed velocity, the maximum wind speed at 10m with a 1-minute sampling interval, the central pressure, the latitude and the radii of maximum winds (RMW).

3. Stop-motion method

The Stop-motion Hybrid TC-induced Waves (SHyTCWaves) metamodel aims to produce accurate estimates of the directional wave spectra produced by a TC at regional scales. Here the Samoa Islands, located in the South Pacific were chosen as the study area, defined by [180°,193°]E, [-20°,-7.5°]N, to perform several TCs control cases to validate the proposed stop-motion method by comparing the full dynamical modeling results with the stop-motion simulation and reconstruction. For this purpose, several control cases (historical storms hitting near Samoa) were tested, whereas TC Ofa (January-February 1990) was selected to illustrate the results hereafter. This storm

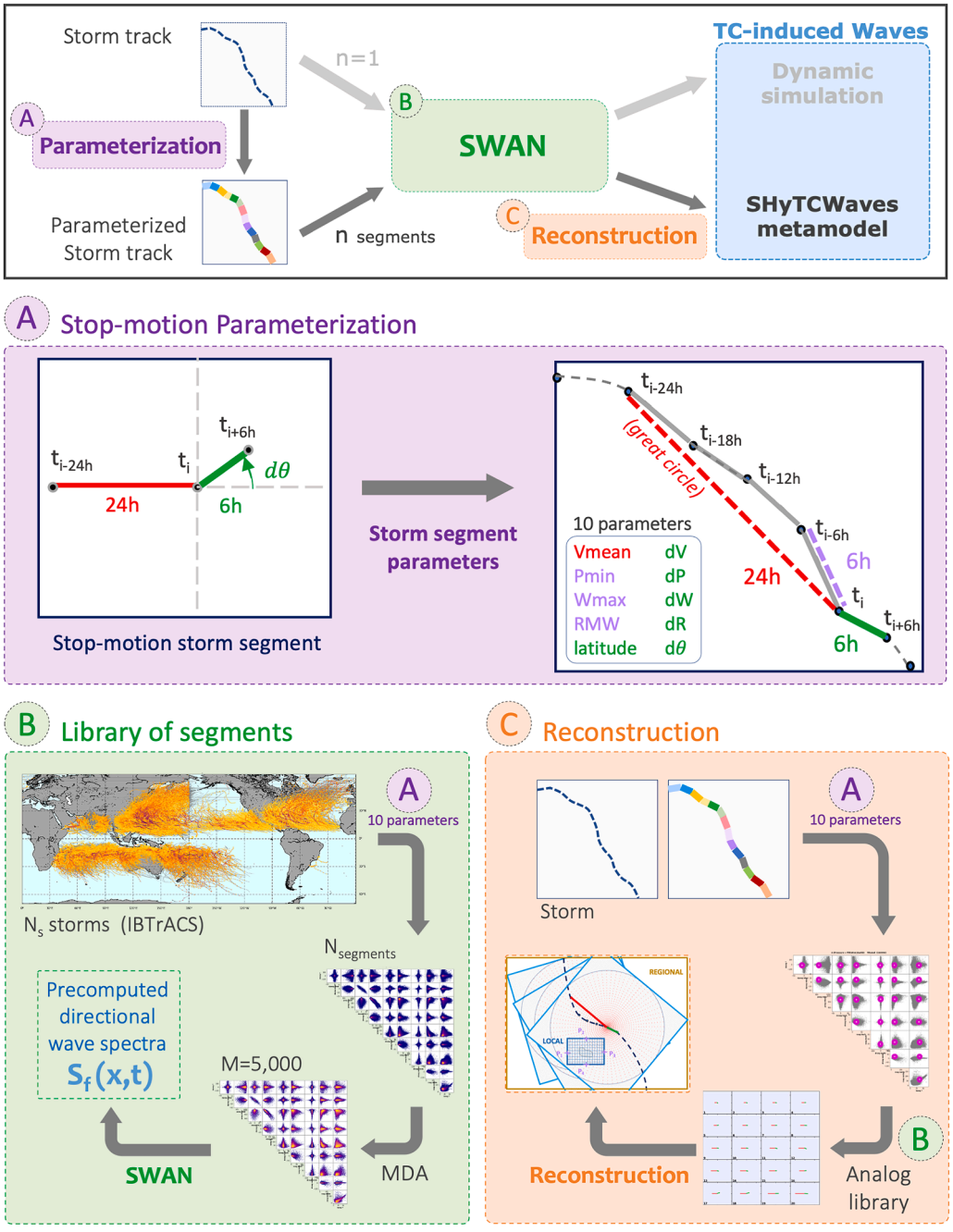


Fig. 1. Methodology flow charts: (upper panel) workflow diagrams of Dynamic (grey arrows) and Stop-motion (black arrows) modeling; (A) definition of SHyTCWaves storm segment (24 h warm-up segment in red plus 6 h target segment in green) and its parameterization into 10 parameters; (B) generation of a smart library of precomputed SHyTCWaves segments; and (C) reconstruction of the directional wave spectra using parameterized SHyTCWaves segments associated to any original TC.

intensified as it approached Samoa from the west heading poleward and reaching category 4 (according to Saffir-Simpson central pressure scale) with maximum sustained winds of 40m/s causing severe impacts due to a combination of storm surge and high seas (Hoeke et al., 2015) estimated at 130 million USD and 7 fatalities were reported (Ready et al. 1992). The control storm TC Ofa was employed to perform a series of tests to establish the optimum configuration of the SHyTCWaves methodology, based on the parameterization of storm segments, obtaining minimum discrepancies when comparing the control case dynamical and stop-motion numerical modeling outputs as depicted in the workflow diagram (Fig. 1 upper panel). Fig. 1 outlines the three main steps of the methodology; the parameterization of TC into storm segments (Fig. 1A); the generation of a library of storm segments (Fig. 1B); and the reconstruction of the original TC (Fig. 1C).

3.1. Full dynamical base case

The full dynamical non-stationary run of TC Ofa was conducted forcing SWAN with parameterized vortex-type wind fields (Holland 1980, Fleming et al. 2008) during a 7-day simulation period. The model computation resolution is 0.136° (approximately 15km) and the time-step is 20 min for TC Ofa that moves with translation velocities smaller than 20km/h. The model wind drag coefficient is capped to prevent overestimation of the estimated wave growth due to very high wind speeds, as indicated in Section 2.3. The resulting time-varying significant wave height (SWH) is used to evaluate the degree of agreement of the proposed methodology. Moreover, the maximum significant wave height (MSWH) field over the simulation period (also called swath map) is used to evaluate the extreme waves differences as a relevant variable which informs about the TC potential impact. Fig. 2a illustrates the storm track position and category according to the Saffir-Simpson central pressure scale; Fig. 2b the parameterized maximum wind speed (MWS) fields and Fig. 2c the MSWH output. It can be observed that TC Ofa intensified while moving southwards with MWS over 40m/s and MSWH around 8m in the surroundings of Samoa.

3.2. Stop-motion parameterization (step A)

The stop-motion approach consists of splitting the TC track into discretized 6 h storm segments, a requisite of the described hybrid approach; each segments may be considered independent storm units capable of being pre-computed. This means that each storm segment is dynamically simulated independently and eventually all segments are ensembled to reconstruct the original storm track induced waves. Since IBTrACS's storm track positions are usually provided every 3 or 6 h depending on the responsible RSMC, a minimum of 6 h period segments is both practical and short enough for the stop-motion procedure. Each track coordinate of historical observations is characterized by the variables: minimum central pressure, mean forward velocity, maximum wind speed, radii of maximum winds, latitude and the angle of forward direction. Fig. 2d illustrates the segmented TC Ofa track, with a total of 21 segments of 6 h within the numerical domain used for the full dynamic control case (Section 3.1). Once the 6 h storm segments are defined, each segment needs to be simulated independently using non-stationary runs of SWAN forced with the corresponding vortex-type wind fields using the Dynamic Holland Model as a function of the storm variables above mentioned. Fig. 2e depicts the SWAN wind forcing timeline for the stop-motion events, as opposed to the full dynamic simulation. Thus, each segment's simulation is divided into three categories: (a) the 24 h warm-up (or model spin-up) segment in red; (b) the 6 h target segment in blue; and (c) the 42 h propagation period in orange. The proposed configuration incorporates the warm-up period leading to the target segment since the wave model is not provided with initial conditions. Instead, the warm-up period accounts for the immediately preceding storm history (i.e., the target segment may be preceded by weaker or stronger intensities, with a change in direction, etc.) thus

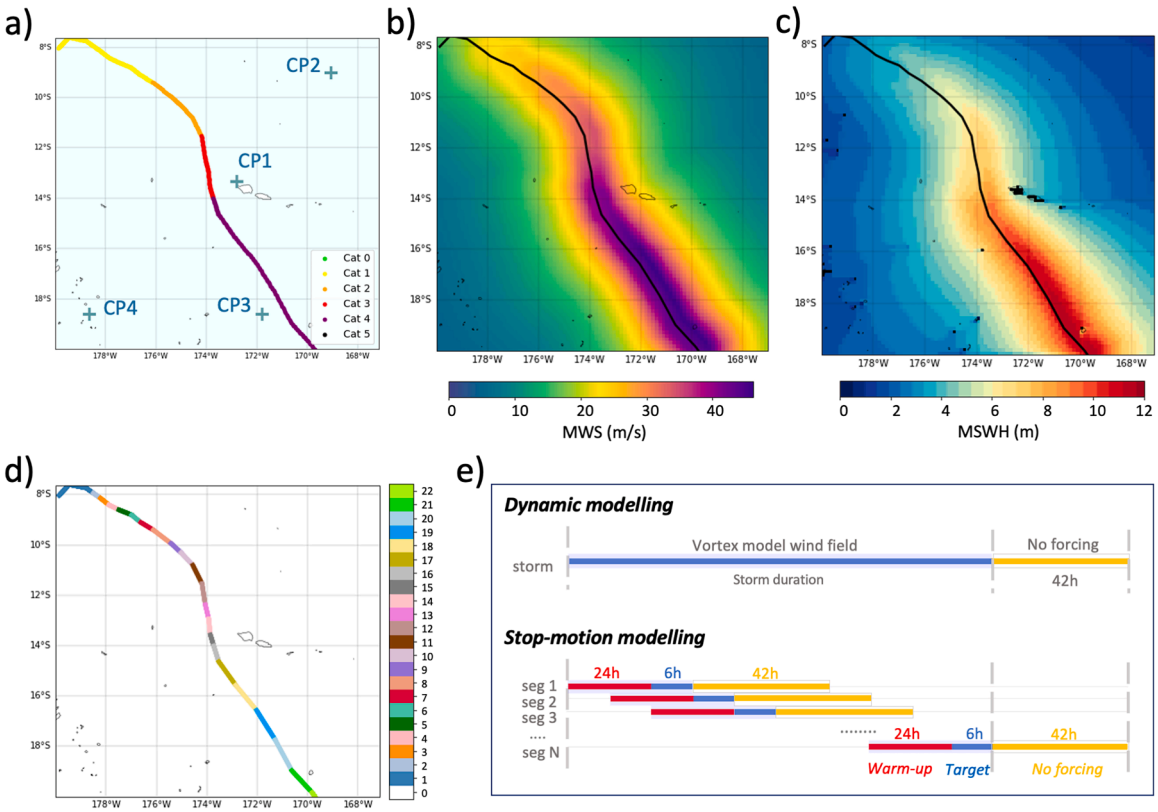


Fig. 2. (Upper row) TC Ofa control case dynamic modeling (a) storm track, (b) maximum wind speed, (c) maximum significant wave height. (Lower panel) (d) 6 h storm track segmentation, and (e) SWAN wind forcing configuration for dynamic and stop-motion modeling.

generating approximate initial conditions at the beginning of the target segment. However, the 24 h initial results of the warm-up segment will be discarded. During the first two periods (a,b) the model is forced with the corresponding time-varying winds, whereas in the last period (c) there is no wind forcing, as it is only meant to allow the wave model to propagate the wave energy generated in the previous 30 h across the whole domain. The effect of this propagation period is significant in the context of the developed framework when reconstructing the original storm track.

Additionally, each stop-motion unit is defined by the warm-up segment plus the target segment. Moreover, each unit is parameterized in terms of a number of parameters which characterize relevant features of the TC. In this study, 10 parameters have been defined as the minimum necessary number to preserve the most information of the track geometry, intensity, size, location and preceding conditions. The parameterization allows a trade-off between reducing the complexity and dimensionality of dealing with all plausible storm segments, and enabling a metamodel to be built based on the 10-dimension parametric space, as explained in Section 3.3.2. These 10 parameters are: (1) the central pressure, *Pmin*; (2) the maximum wind speed, *Wmax*; (3) the radii of maximum winds, *RMW*; (4), the mean forward velocity, *Vmean*; the absolute change of those four variables between the warm-up and the target segment, (5) *dP*; (6) *dW*; (7) *dR*; (8) *dV*; (9) the difference of the heading direction, $d\theta$; and (10) the latitude (*lat*). Fig. 1A illustrates a sketch of the SHyTCWaves parameterization for each unit of 5 consecutive 6 h segments. The first 4 segments constitute the 24 h warm-up segment, which is assimilated as one single segment over a "great circle" (shortest) path between the start/end coordinate of the first/last segment (dashed red line). This geodesic segment is parameterized by: (a) the intensity and size parameters of the last 6 h warm-up segment which immediately precedes the target segment (*Pmin, Wmax, RMW*); (b) the forward velocity is obtained from the geodesic distance and the time between the segment endpoints (*Vmean*); and (c) the forward heading direction is calculated from the geodesic segment (θ). This constitutes the final SHyTCWaves parameterization which proved to return optimum results in the various tests performed over the control case to evaluate the sensitivity analysis on the effect of the warm-up segment period (e.g. 6, 12, 18 and 24 h) and its parameterization (e.g. accounting for different criteria).

Once the 21 segment runs were executed for TC Ofa, the reconstruction consists of selecting for each computational node the SWH envelope (maximum value) among the active segment simulations for each given timestep. Fig. 3 illustrates the comparison of the SWH output for the dynamical control case (black line) and the subsequent segment simulations (colored lines). The grey dashed lines correspond to the warm-up period during which the output is discarded for the envelope calculation. The four control points are located at the northwest of Samoa (CP1), at the northeast corner of the study area (CP2), at the south boundary (CP3) and at the southwest corner (CP4) as shown in Fig. 2a. These four points showcase the results both at locations close to the storm track (CP1, CP3) and thus subjected to local intense wind speeds when OFA reaches close to the island, as well as at a location far more distant from the storm path (CP2, CP4). It can be seen that similar results have been obtained for all locations, whether waves are generated by local wind seas or waves are produced by far distant generated swells which arrive over time, demonstrating that the 24 h warm-up period is capable of reproducing the fully developed wave growth especially at the early stages when the wind intensity starts to pick up.

Additionally, Fig. 4 exhibits a series of swath maps, or the maximum values during the simulation period at each grid node, which depicts the "footprint" of extreme waves. Fig. 4a and b show the consecutive segments' wind forcing and MSWH outputs respectively, in a grid of 20 independent stop-motion units. Fig. 4c corresponds to the reconstructed MSWH, and Fig. 4d the difference between the reconstructed MSWH (Fig. 4c) and the MSWH from the dynamic control case (Fig. 2c) (blue values depict underestimation of the dynamical control case). The sensitivity tests with increasing warm-up periods (not shown) demonstrated that when there is little to no warm-up period, the envelope lacks a great deal of wave growth with maximum differences of about -4m and localized jumps along the storm track. Evidently with successive longer warm-up durations the swath difference decreases considerably. It should be noted that the largest discrepancies are spotted around the exit of the storm, probably due to the effect of the bottom boundary of the numerical domain, and the fact that the last full target segment ends some time earlier than the dynamic control case. Overall, it was concluded that a minimum of 24 h warm-up period forcing conditions is required prior to the target segment to minimize swath differences.

Regarding the parameterization of the warm-up segment, a

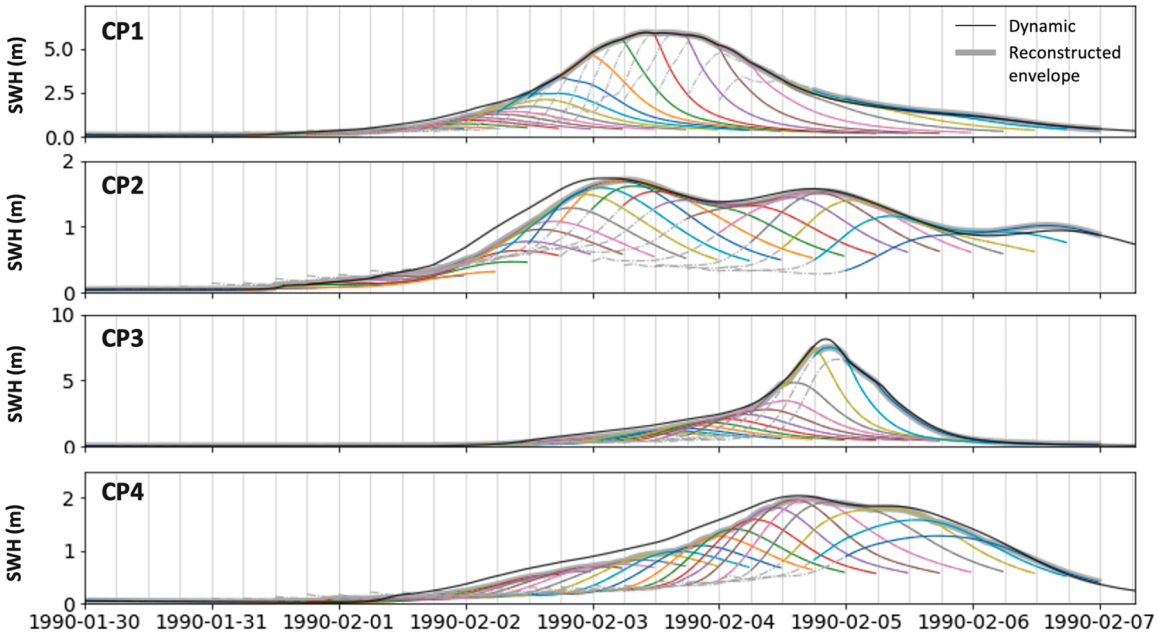


Fig. 3. Time series of SWH for control points CP1, CP2, CP3, CP4: dynamical modeling (black line) and stop-motion segments modeling (colored lines) with 24 h warm-up period. Grey dashed lines for warm-up periods are discarded; grey line for the reconstructed envelope from stop-motion colored lines.

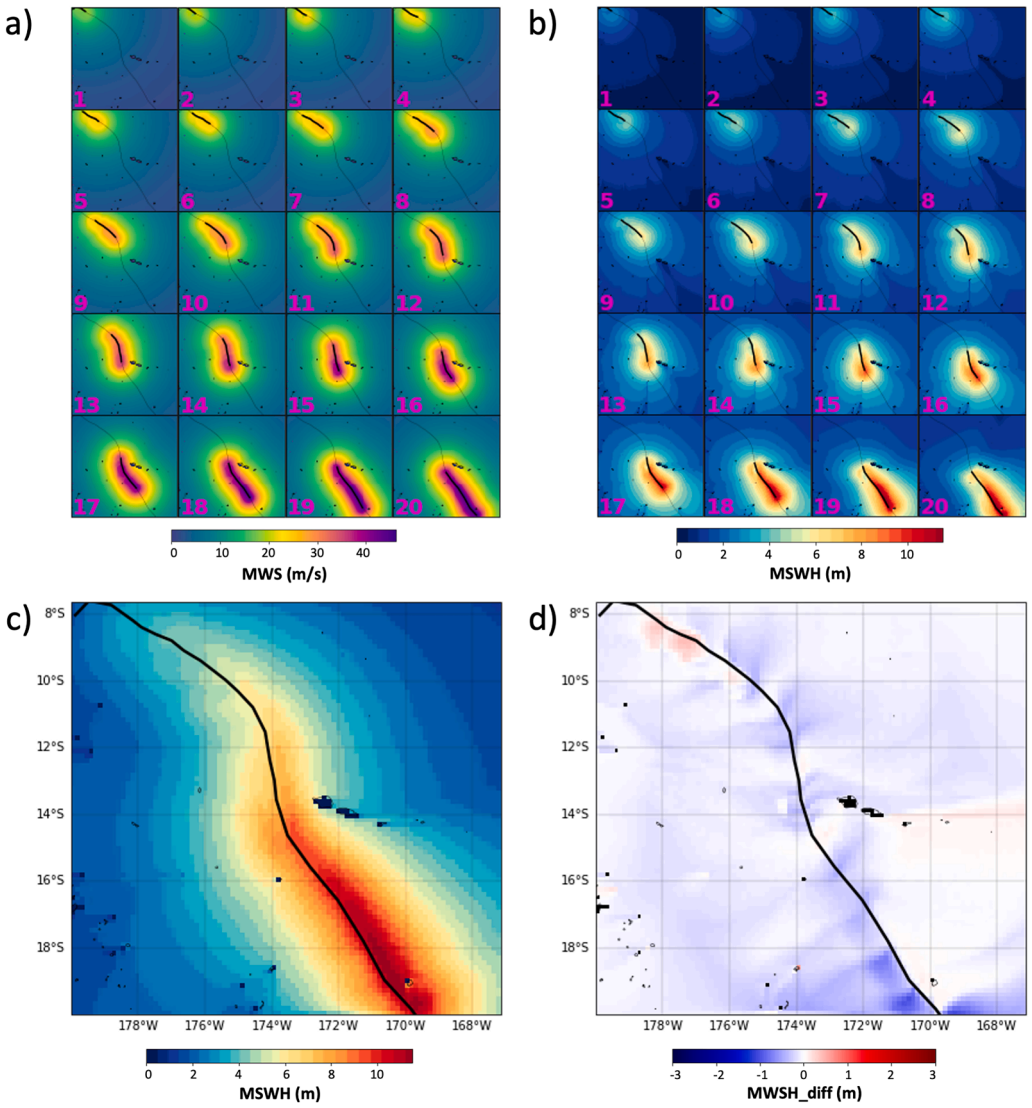


Fig. 4. Maximum fields for 6 h segments of (a) wind speeds (m/s), (b) significant wave height (m); (c) stop-motion reconstructed MSWH, and (d) difference between the stop-motion and dynamic modeling of TC Ofa.

sensitivity analysis was performed to evaluate the effect of different criteria to account for the storm intensity prior the target segment: (a) a constant value as the mean average of the four preceding 6 h segments; (b) variable values as the linear increasing weighted average of all four segments until the target segment, and (c) a constant value as the mean average of the immediately preceding 6 h segment. Eventually it was concluded that option (c) is the optimum configuration whereas options (a) and (b) tend to underestimate the reconstructed SWH (not shown).

Fig. 5a shows the swath map of differences of MSWH between the stop-motion and the dynamical control case. Discrepancies are bound to +0.2m and -0.3m when neglecting the outliers near the simulation southern boundary (as mentioned above). Therefore, the degree of uncertainty that is assumed with constant 6 h segments is admissible. Moreover, three error metrics were evaluated to assess differences along the simulation period: RMSE, correlation coefficient and bias (Fig. 5b–d). In all cases, it can be observed that the storm entrance and/or exit of the domain should be neglected due to the boundary effect of the forcing. The correlation coefficient displays values over 0.98. The RMSE ranges from 5 to 30cm, even though some high values remain near the storm track (relative to SMH of the order of 10m). The bias informs of a tendency to overestimate SWH while the storm intensity is weaker (up to category 3, upper part of the domain) and once it reaches category 4 or higher, the stop-motion ensemble tends to underestimate up to 8cm.

Lastly, Fig. 5e and f illustrate the swath plots of the directional wave spectra corresponding to four control points across the domain. On one hand, points CP1 and CP3 are located closer to the storm track, so that over time they are subjected to high wind speeds and therefore exhibit a much higher density energy, as well as a wide range of incoming directions as the storm moves forward. On the other hand, points CP2 and CP4 are located far distant at each side of the track, where the swath directional wave spectra shows that the wave energy is concentrated at lower frequencies (swells) and with a reduced directional spreading. The swath difference plots (Fig. 5f) account for the frequential and directional discrepancies of the stop-motion envelope, which are more noticeable in points near the track (CP1, CP3). Nevertheless, differences are relatively marginal since the stop-motion approach is able to preserve the continuum transfer of energy from high to low frequencies over time and during the simulation period. Also, the effect of islands obstacles can be distinguished in CP2 which stop-motion approach is not able to reproduce.

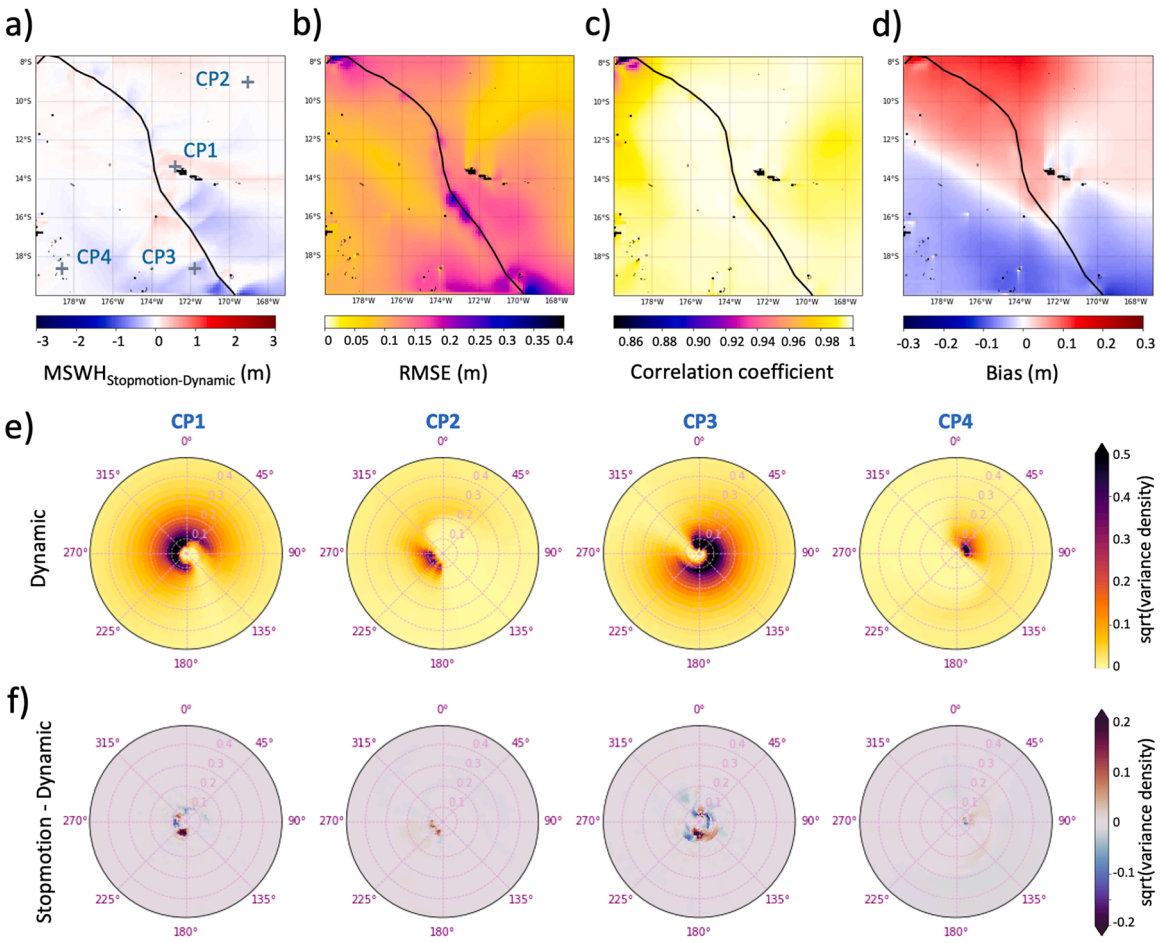


Fig. 5. Error metrics of the stop-motion parameterization compared to the dynamical control case: (a) absolute difference of MSWH; (b) root mean squared error; (c) correlation coefficient; (d) bias. Directional wave spectra at control points (e) dynamical control case, and (f) difference between stop-motion and dynamical.

4. SHyTCWaves method

In this section, we utilize the stop-motion approach to develop a hybrid metamodel (SHyTCWaves) based on a library of pre-run storm segments that can be used to reconstruct any given storm track using analog, and thus obtaining an approximation of the TC-induced waves. The following subsections describe the generation of the SHyTCWaves library and its application.

4.1. Library of segments (step B)

Former studies have demonstrated that a hybrid framework is very efficient since it does not demand large computational resources, and in return, is capable of providing relatively fast estimates with reasonably good results (Camus et al. 2011, van Vloten et al. 2022, Ricondo et al, 2023). Fig. 1B and C schematize the workflow, composed of the generation of a database of TC segments, its selection method, and the execution and storage of those numerical simulations. To build a library of dynamically simulated segment events, it is needed to characterize the empirical joint distribution functions of the 10-dimensional parametric space. Once a segment database is generated from historical tracks ($N_{segments}$ = 320,000), a selection technique will be applied to extract a representative sample of historic segments. Those events will be numerically simulated and will constitute the library of analog events used for the application of the ShyTCWaves methodology to produce estimates of TC-induced waves.

4.2. TC segments database

The historical records compiled in IBTrACS were used to generate the largest possible database of TC segments. However, some historical records exhibit gaps of information lacking in one or two variables of interest (Wmax and/or RMW). For this reason, whenever the storm position and the central pressure are provided, any missing variables were estimated with existing empirical relationships to fill those gaps. Wmax was calculated using a 3-order polynomial regression model as a function of the central pressure for each RSMCs and ocean basin; this empirical relationship has significant statistical correlation. RMW was estimated as a function of storm intensity and latitude as obtained by Knaff et al. (2015) from observed flight-level TC winds. Despite the fact that historical data of RMW is only available from season 2001 onwards, the RMW is still included for two reasons. Firstly, it provides a measure of the storm size that must be provided for the vortex-type wind model calculation. Secondly, one of the applications of the SHyTCWaves metamodel may be to produce TC-induced wave estimates for early warning systems based on forecasted TC tracks from official agencies which usually inform about the RMW.

Once the historical storm track database has been processed, the resulting TC parameterized segment database is populated by 320,000 unique elements which are plotted in Fig. 6 (grey dots) showing the joint statistical relationships among the sample of all 10 parameters. It can be noted the high variability and dispersion of most parameters: all possible changes of forward direction [-180°, 180°], translational velocities up to 120km/h, maximum wind speed up to 180kt, maximum radii of 180nmile, and latitudes from near the equator up to 70°, as latitude has

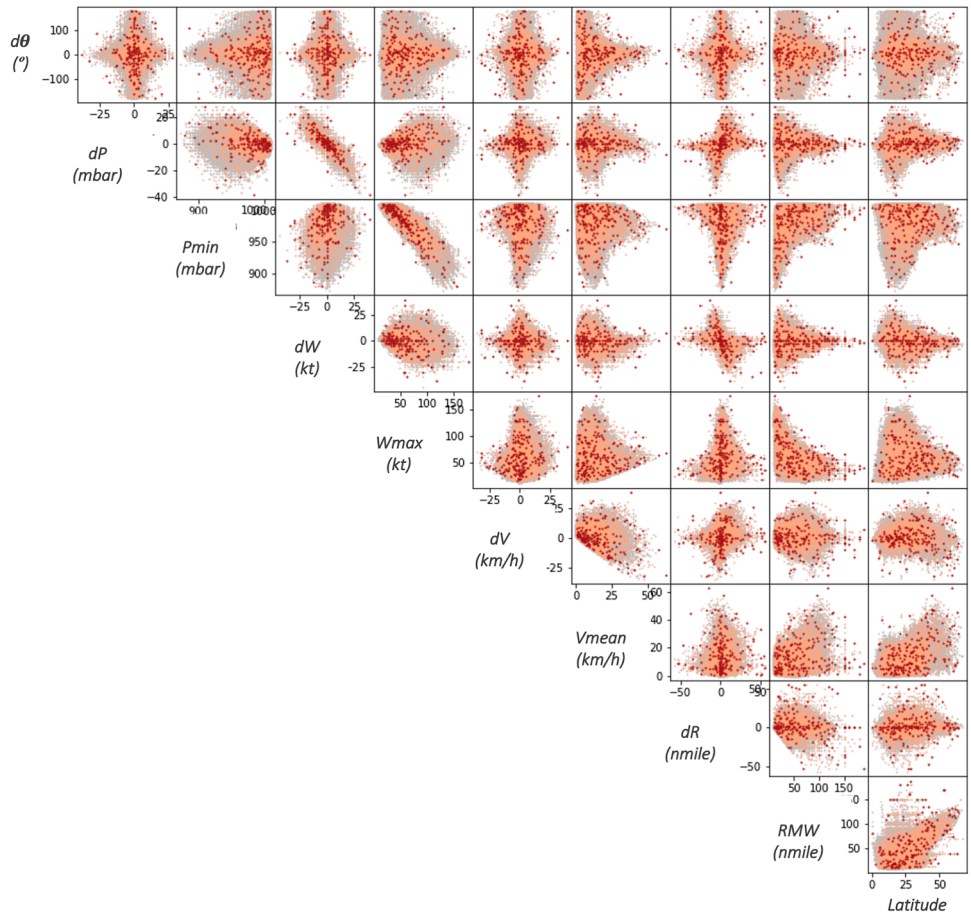


Fig. 6. Density scatter plots of the 10 SHyTCWaves parameters for historical TC database (grey dots) and MDA subset (red and pink dots, for 1–200 and 201–4,800 cases respectively) of parameterized segment events.

been stored as the absolute value to aggregate events from both hemispheres. Also, the effect of the empirical estimate of RMW can be clearly distinguished in the RMW-latitude subplot. Similarly, the Pmin-Wmax relationship is more robust and less scattered. The sign convention for the change of forward direction (dθ) in the north hemisphere is positive/negative when steering towards the right/left hand side. Since storm tracks move poleward starting from low latitudes, there is an axial symmetry around the equator regarding the track forward direction. For this reason, the dθ sign convention is the opposite in the southern hemisphere. The wind rotation is counterclockwise in the north hemisphere and clockwise in the south hemisphere.

4.2.1. Selection of cases

The TC segment database contains a large collection or plausible segment units with a wide range of possible combinations. Here, some outliers were identified and removed from the segments database with the following criteria: (1) storms move at absolute speeds lower than the maximum winds (Vmean < Wmax needs) which condition is required by the vortex wind calculation; (2) the change of forward velocity is capped with a threshold of 70km/h; (3) some RMW values over 200nmiles are identified as outliers and discarded. The resulting filtered database has 300.000 elements.

For constructing the hybrid model, a reduced number of cases will be numerically simulated. For this purpose, the Maximum Dissimilarity Algorithm (MDA, Camus et al. 2011) is used to obtain a selection of cases well distributed and covering the full parametric data space. This technique chooses the most dissimilar n-dimension data from each other, so that the selected multi-dimensional data is fairly distributed, as well as exploring the bound limits of the parametric data space (Camus et al., 2011). Fig. 6 shows the density scatter plot resulting from a selection of 5,000 cases. This number was selected as a compromise between the computational resources available, and obtaining a subset large enough for the hybrid model. Red dots depict the first 200 most dissimilar cases, whereas the pink dots illustrate the remaining MDA subset, compared to the historical dataset (grey dots).

4.2.2. Numerical modeling

The next step is to perform wave model simulations over the 5,000 selected cases and store the library of time series outputs of directional wave spectra. To facilitate the stop-motion methodology, a few changes are made to the SWAN configuration described in Section 2.3. A cartesian coordinate system (instead of spherical coordinates) is used to allow arbitrary re-projection applicable to any geographical location. A constant latitude is assumed for each segment event, which generally is small in size so that discrepancies of the associated wind fields are marginal. Since the stop-motion events are meant to be employed to reconstruct TCs at any ocean basin, the computational domain is defined as an idealized deep water extent of 1500km on each side with a 15km resolution. The default position of the pair of segments which constitute an event, places the warm-up segment horizontally over the negative x-axis thus heading eastward; the starting point of the target segment falls at the origin of the cartesian coordinates, and it heads forward with an angle corresponding to dθ. The size of the domain ensures that all target segments fit generously with ample space in front so that the stored output can account for wave propagation both at local and far distant locations from the storm track. Some warm-up segments in the subset selection are fast-moving segments, and for those cases the computational domain is extended further to the west.

Regarding the model bathymetry is defined as a uniform constant depth of 2000m, emulating an infinite pool of deep waters without obstacles. It is assumed that the applicability of the proposed methodology is limited to regional scales, and it does not aim to resolve nearshore processes. The wave model is forced with parameterized vortex-type wind fields during the warm-up and target segments. Finally, the hourly directional wave spectra is stored at custom selected points defined by a polar coordinate system of 2200 nodes distributed in a ~7.5° radii circumscribed circle with 28 rings, and 5° directional sectors. This allows to store results at a higher spatial resolution near the storm track with local high winds forcing, whereas far-off distances from the center have a more relaxed resolution as the corresponding wave spectra is not expected to experience significant variations. Once all simulations are performed, and the library of storm segments is generated, the last final step is the application which is presented hereunder.

4.3. Reconstruction (step C)

SHyTCWaves can be applied to any given storm track with known coordinates and central pressure as these are the minimum data required, with remaining key variables either implicit in the storm geometry (latitude; mean forward velocity; angle of forward direction) or can be estimated with empirical relationships whenever data is not available (Wmax; RMW). The flow chart for the SHyTCWaves application is depicted in Fig. 1C consisting of three steps. Firstly, the storm track is broken down into 6 h segments, and parameterized into stop-motion units as described in Section 3.2 and Fig. 1A (24 h segment followed by 6 h segment) in terms of 10 parameters. Secondly, for each parameterized stop-motion unit the analog case from the library of pre-run cases is selected with the criteria of the minimum Euclidean distance in the dimensionless parametric data space. The last step constitutes the postprocessing of the series of analog cases corresponding to all consecutive segments to produce the ensemble of the reconstructed storm track induced waves.

The ensemble postprocessing takes each analog event from the pre-run library (Fig. 7b) and projects the outputs polar grid georeferenced over the real study area, taking into account the sign convention for either hemisphere. The geographical projection places the analog target segment with the orientation that matches the real storm track target segment. This is done consecutively for the series of analog cases, and Fig. 7c exemplifies such a process for TC Ofa. Finally, the ensemble of the SWH envelope is calculated at any control point by spatially interpolating the outputs from all the "active" analog segments and extracting at each time coordinate the analog case with the maximum value of SWH. In the same way, the peak period and mean direction can be ensembled. Therefore, the extent of estimates provided by the SHyTCWaves metamodel is limited to a 7.5° circle moving along the TC track. In terms of computation, the ensemble postprocessing takes only a few min for a storm track, and it varies depending on its duration and the total number of segments to be computed.

Fig. 8 shows the results obtained for TC Ofa. On the left, it is shown the comparison of swath maps along the period of 9 days, for the dynamical control case (Fig. 8a) and SHyTCWaves (Fig. 8b). The footprint of the MSWH along the TC is able to capture the spatial variability and the extreme values at the lefthand side of the track (as winds rotate clockwise in the southern hemisphere). However, the methodology is not prepared to account for the wave energy that is masked due to islands (Samoa in the east, Tonga in the south) and the refraction around obstacles. Nonetheless, the output wave spectra are useful as hydraulic boundary conditions for regional models downscaling to nearshore coastlines which would account for such refraction and dissipation internally. The panels on Fig. 8c show the time series of the SWH and peak period of the dynamical control case (black line), and the SHyTCWaves results with a color code corresponding to each subsequent segment identifier. Taking into account that the use of analog segments introduces an additional uncertainty inherent in the hybrid model framework, it can be observed that the time series of both SWH and peak period are good approximates of the full dynamical solution. The maximum values of SWH are well captured at both control points located near and far from the TC. In Fig. 8d, the directional wave spectra at the four control points are analogous to the ones in Fig. 5e corresponding to the dynamical control case. These demonstrate that both the directional and frequency distribution of the arriving wave energy is very close to the dynamical control case, except for the inability to account for obstacles which result in overestimating the energy at such points as CP2. The same validation of SHyTCWaves methodology against dynamical simulations has been performed for other TCs (e.g. TC Isaac (1982), TC Evan (2013), TC Gita (2018)) with similar results (not

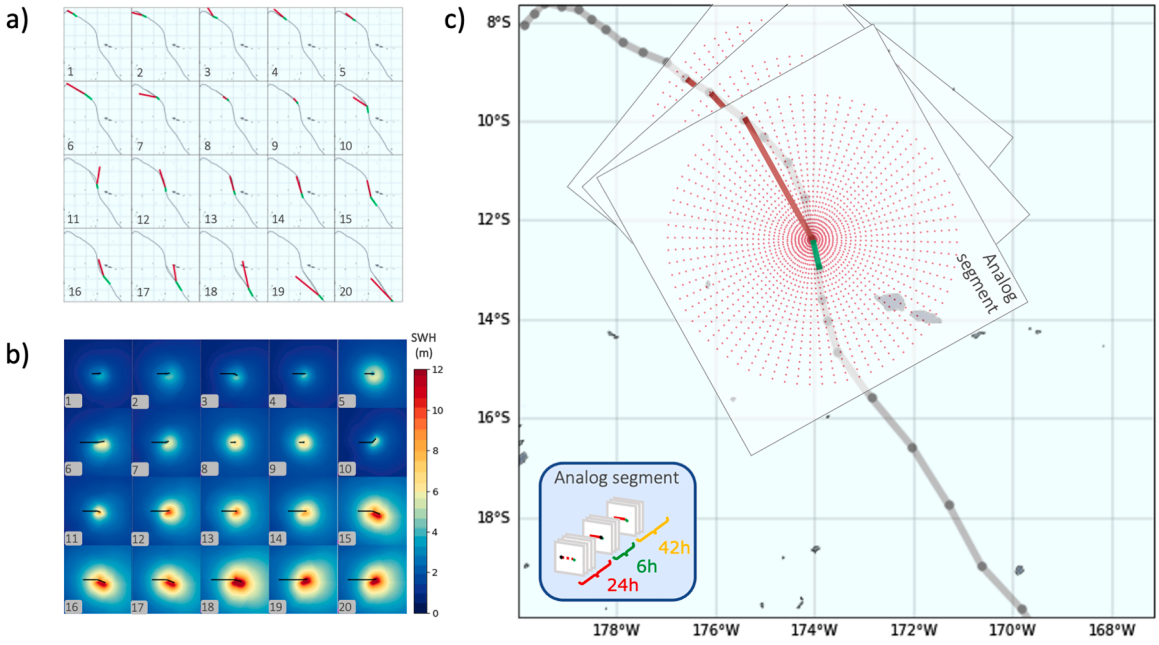


Fig. 7. TC Ofa track (a) analog segments; (b) the associated analog MSWH fields from the precomputed library; and (c) sketch of the stop-motion ensemble with subsequent geo-projections of the analog polar grid directional wave spectra over the study area.

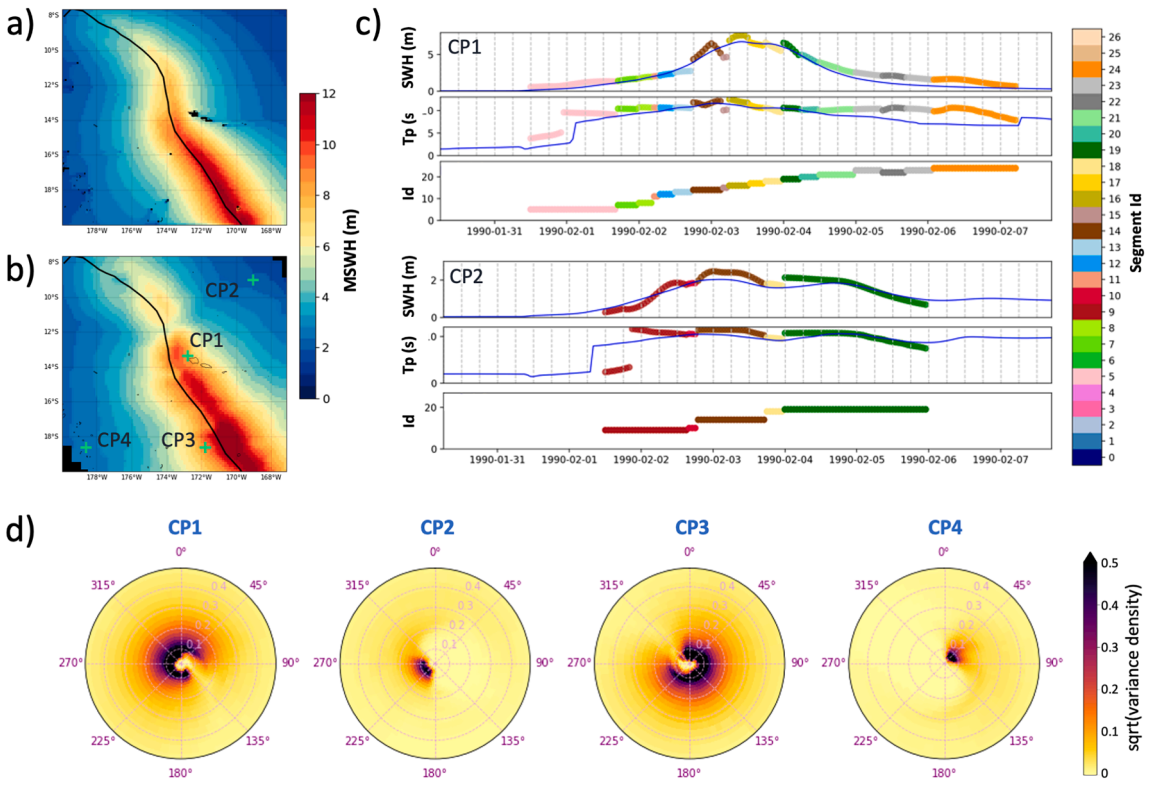


Fig. 8. TC Ofa application results. (a) Dynamical modeling; (b) stop-motion modeling, (c) time series of ensembled SWH and peak period for control points CP1, CP2; and (d) directional wave spectra for control points CP1, CP2, CP3, CP4.

shown).

5. Bulk bias correction

To evaluate the skill of SHyTCWaves in producing consistent good approximations of TC-induced waves, the model has been applied to several historical TCs and compared against satellite data. After comparison, a bulk bias correction of the metamodel is proposed in order to improve its accuracy. This correction attempts to reduce all the uncertainties associated to the parameters of the tracks, as well as the parameters of the wave model for extreme conditions. For this purpose, the metamodel is applied to a subset of historical TCs from all ocean basins (Fig. 9a), for which the IMOS satellite data was available, thus comparing the SWH at longitudinal profiles (Fig. 9b). The procedure that is followed to determine the subset of TCs was the application of an MDA over the variables longitude, latitude, speed and direction. Furthermore, this is performed for all category groups (0, 1, 2, 3, 4, 5) according to the Saffir-Simpson pressure scale with a total of 500 tracks. On the other hand, satellite data is available from 1985 to present, and we evaluated contemporary data within a radius of 300km and a time lapse of +-2h

The longitudinal profile comparison shown in Fig. 9b manifested a good agreement in some cases whereas especially the low intensity categories were underestimated by SHyTCWaves. These errors may be explained by different reasons. On one hand, the wind drag coefficient in the wave model that caps the wave growth for high wind speeds may result in lacking energy for weaker TCs. Lastly, satellite data may incorporate some uncertainty of its own. Moreover, when the geometry storm tracks were closely examined, some errors were naturally accountable to the non-existent land mask effect (overestimation) and to some convoluted tracks which add wave interaction processes to account for.

In order to reduce the sources of uncertainty, the calibration subset was filtered by taking only events with historical records of RMW (not estimated), and at tropical latitudes (up to 25°); satellite data was filtered with flags to differentiate outliers. After examining the profiles comparison, it was found that there was a consistent relationship between the disagreement and the storm intensity. Therefore, the pressure parameter was corrected as P*=P+dP with a linear fitting of dP as a function of intensity categories, which has proven to improve the infra/overestimation at low/high categories in tropical latitudes as it is illustrated by the QQplot shown in Fig. 9c. It can be seen that lower tail is successfully corrected after calibration while the upper tail remains with some bias. For our particular application of SHyTCWaves, the bias correction factor dP is {-17, -15, -12.5, -7, 2.5, 10}mbar for central pressure {1015, 990, 972, 954, 932, 880}mbar corresponding to the center of categories 0, 1, 2, 3, 4 and 5 respectively. Fig. 9b shows the final (black dots) versus uncorrected (grey dots) SHyTCWaves longitudinal profiles comparison, with the subplots background color indicating the storm intensity (0: green, 1: yellow, 2: brown, 3: red, 4: purple, 5: black), the vertical axis representing SWH from 0–15m, with altimeter (red dots). Generally, there is a good agreement in the maximum and evolution of SWH in the majority of cases. To conclude, several historical TCs which are well-known in different oceans have been plotted in Fig. 9d to illustrates the SHyTCWaves estimates of SWH at 0.5° resolution., with an average of 5 min of computation each.

6. Summary and conclusions

This paper examines and demonstrates the concept of the stop-motion approach to hybrid tropical cyclone wave modelling in deep waters, where the influence of depth related dissipation and refraction can be considered negligible. The approach successfully approximates dynamic wave growth and subsequent propagation of storm waves as the linear superposition of sequential 6 h periods of wind conditions, provided that initial wave conditions are appropriately informed. This means that any TC track can be split into smaller segments of 6 h, and together with the immediately preceding 24 h period to account for

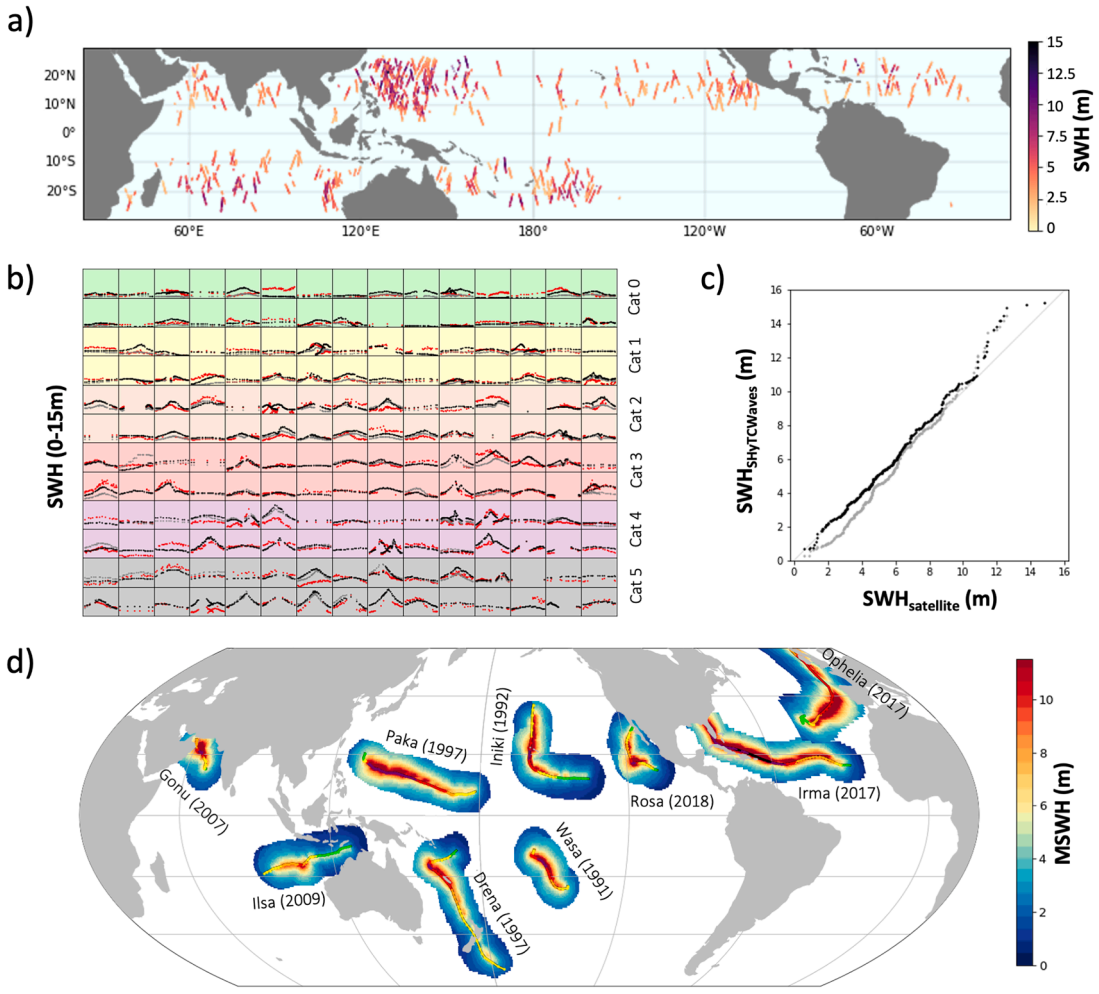


Fig. 9. (a) Satellite altimeter longitudinal profiles of SWH; (b) comparison of longitudinal profiles of historical TC tracks with satellite altimeter (red dots) and SHyTCWaves SWH (grey/black dots for prior/after bulk correction); (c) QQplot of SWH for prior/after bulk correction (grey/black dots); (d) SHyTCWaves application for several TCs

warm-up prior the target segment, the contributions of each successive shorter events can be used to reconstruct the full original storm track with reasonably accurate results. This framework is based on the performance of high-fidelity wave model (SWAN) simulations, with the difference that shorter periods are simulated instead of the complete event. Nevertheless, it has been shown that the reconstruction by taking the maximum values of SWH from all the active segment events at each spatio-temporal coordinate provides a very good agreement compared with the dynamical control case. Thus, the information of the continuum transfer of energy from high to low frequencies across the numerical domain and the simulation period is preserved, even in the context of a TC which is a time-varying moving system which generate local seas and swell trains that propagate in all directions. Emphasis was placed on evaluating several sensitivity tests to determine the optimum configuration in terms of the necessary warm-up duration and its parameterization. Furthermore, the final SHyTCWaves parameterization of segment events is characterized by *Pmin, Wmax, RMW, Vmean, dP, dW, dR, dV, dθ* and *lat*. Moreover, the 24 h warm-up segment is composed of the four 6 h segments leading to the target segment, and it is assimilated as one single segment over a "great circle" or geodesic path between the start/end coordinate of the first/last segment. This geodesic segment is parameterized in turn by the intensity and size parameters (Pmin, Wmax, RMW) of the last 6 h warm-up segment which immediately precedes the target segment, as it demonstrated more accurate results in view of the error metrics evaluated and compared among three criteria of parameterizations.

Additionally, a metamodel was built taking advantage of the stop-motion approach and the hybrid method of simplifying a high dimensional problem with the aid of statistical and selection techniques which has shown good results in previous studies. Therefore, the SHyTCWaves metamodel is vastly computational less expensive than fully dynamically simulated TC wave model predictions. While it is unable to directly account for shallow-water wave transformations, it is capable of providing directional wave spectra which can be used by coastal downscaling models. It makes use of a library of 5,000 precomputed subset of parameterized segments well distributed over the parametric data space, to reconstruct a close-to-real storm track by geo-projecting the closest analog cases of the library over subsequent 6 h segments of the original TC track. The developed metamodel provides a methodology that allows to obtain wave parameters (SWH, peak period, direction) with good approximations in a matter of 5–10 min on a single central processing unit (CPU) depending on the length and duration of the TC. Furthermore, results over a set of historical storms showed that the parameterization needed to be calibrated as it was identified that low intensity storms consistently underestimated altimeter data collected at concurrent longitudinal profiles. Therefore, the pressure parameter was calibrated with an additional differential of pressure as a function of the storm category, improving the methodology validation in the lower tail of the distribution, while the upper tail remains with some bias.

A comprehensive framework has been described; it is worth noting that it is ready to use globally in deep waters with no land obstacles, however its applicability is currently limited to offshore estimates not including the refraction and shade of energy due to obstacles such as island territories or coastlines. In the context of dynamical or hybrid downscaling into nested numerical models, SHyTCWaves can provide good and fast approximations of the directional wave spectra that may be incorporated as hydraulic boundary conditions especially around oceanic islands and archipelagoes. The precomputed library has been stored with full and degraded resolution which can be used depending on the application and spatial extent of the storm; such databases are 4Tb and 500Gb respectively. There is potential research that can be pursued to improve and incorporate certain aspects of the methodology. The application of an obstacle mask should be considered to remove the wave spectral energy which would not reach in real conditions due to the existence of land obstacles near the target location, which would enlarge even more the method's applicability. The number of cases that constitute the library was not subjected to a sensitivity analysis to determine the optimum number, therefore this remains to be addressed, as well as the resolution used to define the library polar grid and spectral frequency and direction resolution, to optimize the method's efficiency. On the other hand, the use of synthetic cyclones may be the scope of further research in order to generate a library of increased plausible cases due to climate change and future projections. An additional advantage of this methodology is that its low computational cost enables to increase greatly future scenario modeling, thus translating in improved probabilistic estimates of storm wave impacts at both weather and climate timescales.

CRediT authorship contribution statement

Sara O. van Vloten: Conceptualization, Methodology, Software, Validation, Visualization, Writing – original draft. **Laura Cagigal:** Funding acquisition, Methodology, Supervision, Writing – review & editing. **Beatriz Pérez-Díaz:** Supervision, Writing – review & editing. **Ron Hoeke:** Supervision, Writing – review & editing. **Fernando J. Méndez:** Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, 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.

Data availability

Data will be made available on request.

Acknowledgments

The authors would like to acknowledge the funding from the projects CE4Wind (CPP2022-010118 MCIN/AEI/10.13039/501100011033 and European Union - Next GenerationEU/PRTR), Perfect-Storm (2023/TCN/003 - Goverment of Cantabria/FEDER, UE) and MyFlood (PLEC2022-009362 - MCIN/AEI/10.13039/501100011033 and European Union - Next GenerationEU/PRTR). Ron Hoeke's contributions were supported by the CSIRO's Research Office and in-kind contributions from CSIRO Environment. Laura Cagigal acknowledges the funding from the Juan de la Cierva – Formación FJC2021-046933-I/MCIN/AEI/10.13039/501100011033 and the European Union "NextGenerationEU"/PRTR.

References

Bakker, T.M, Antolínez, J.A.A., Leijnse, T.W.B., Pearson, S.G., Giardino, A., 2022. Estimating tropical cyclone-induced wind, waves, and surge: A general methodology based on representative tracks. Coast. Eng. 176 (2022), 104154 https://doi.org/10.1016/j.coastaleng.2022.104154.

Cagigal, L., Mendez, F.J., van Vloten, S.O., Rueda, A., Coco, G., 2022. Wind wave footprint of tropical cyclones from satellite data. Int. J. Climatol. 43 (1), 372–381. https://doi.org/10.1002/joc.7764.

Camus, P., Mendez, F.J., Medina, R., 2011. A hybrid efficient method to downscale wave climate to coastal areas. Coast. Eng. 58 (9), 851–862. https://doi.org/10.1016/j.coastaleng.2011.05.007.

Emanuel, K., 2003. Tropical cyclones. Annu. Rev. Earth Planet Sci. 31 (1), 75–104. https://doi.org/10.1146/annurev.earth.31.100901.141259.

Fleming, J., Fulcher, C., Luettich, R., Estrade, B., Allen, G., Winer, H., 2008. A real time storm surge forecasting system using ADCIRC. In: Proceedings of the Estuarine and Coastal Modeling 10th International Conference. https://doi.org/10.1061/40990(324)48.

Grossmann-Matheson, G., Young, I.R., Alves, J., Meucci, A., 2023. Development and validation of a parametric tropical cyclone wave height prediction model. Ocean Eng. 283, 115353 https://doi.org/10.1016/j.oceaneng.2023.115353.

Hasselmann, K., Sell, W., Ross, D.B., Müller, P., 1976. A parametric wave prediction model. J. Phys. Oceanogr. 6, 200–228. https://doi.org/10.1175/1520-0485(1976)006<0200:APWPM>2.0.CO;2.

Harper, B. A., J. D. Kepert, and J. D. Ginger, 2010: Guidelines for converting between various wind averaging periods in tropical cyclone conditions. World Meteorological Organization, TCP Sub-Project Rep., WMO/TD-1555, 54 pp.

Hoeke, R.K., McInnes, K.L., O'Grady, J.G., 2015. Wind and wave setup contributions to extreme sea levels at a tropical high island: a stochastic cyclone simulation study for Apia, Samoa. J. Mar. Sci. Eng. 3, 1117–1135. https://doi.org/10.3390/jmse3031117.

Holland, G.J., 1980. An analytic model of the wind and pressure profiles in hurricanes. Mon. Wea. Rev. 108 (8), 1212–1218. https://doi.org/10.1175/1520-0493(1980)108<1212:AAMOTW>2.0.CO;2.

Hwang, P.A., 2011. A note on the ocean surface roughness spectrum. J. Atmos. Oceanic Technol. 28, 436–443. https://doi.org/10.1175/2010JTECHO812.1.

Irish, J.L., Resio, D.T., Cialone, M.A., 2009. A surge response function approach to coastal hazard assessment. Part 2: quantification of spatial attributes of response functions. Nat. Hazards 51, 183–205.

Jia, G., Taflanidis, A.A., Nadal-Caraballo, N.C., Melby, J.A., Kennedy, A.B., Smith, J.M., 2016. Surrogate modeling for peak or time-dependent storm surge prediction over an extended coastal region using an existing database of synthetic storms. Nat. Hazards 81 (2), 909–938. https://doi.org/10.1007/s11069-015-2111-1.

Kim, SW., Melby, J.A., Nadal-Caraballo, N.C., et al., 2015. A time-dependent surrogate model for storm surge prediction based on an artificial neural network using high-fidelity synthetic hurricane modeling. Nat Hazards 76, 565–585. https://doi.org/10.1007/s11069-014-1508-6.

Knaff, J., Longmore, S., Demaria, R., Molenar, D., 2015. Improved tropical-cyclone flight-level wind estimates using routine infrared satellite reconnaissance. J. Appl. Meteor. Climatol. 54 (2), 463–478. https://doi.org/10.1175/JAMC-D-14-0112.1.

Knapp, K., Kruk, M., Levinson, D., Diamond, H., & Neumann, C., 2010. The international best track archive for climate stewardship (IBTrACS). Project: Overview of Methods and Indian Ocean Statistics. 10.1175/2009BAMS2755.1.

Knapp, K.R., Diamond, H.J., Kossin, J.P., Kruk, M.C., Schreck, C.J. 2018. International best track archive for climate stewardship (IBTrACS) project, version 4. NOAA National Centers for Environmental Information. 10.25921/82ty-9e16.

Kudryavtsev, V., Yurovskaya, M., Chapron, B., 2021. 2D parametric model for surface wave development under varying wind field in space and time. J. Geophys. Res. Oceans 126, e2020JC016915. https://doi.org/10.1029/2020JC016915.

Kyprioti, A.P., Taflanidis, A.A., Nadal-Caraballo, N.C., Cambpell, M., 2021. Storm hazard analysis over extended geospatial grids utilizing surrogate models. Coast. Eng. 168 (2021), 103855 https://doi.org/10.1016/j.coastaleng.2021.103855.

Ready, S., Woodcock, F., 1992. The South Pacific and southeast Indian Ocean tropical cyclone season 1989–1990. Aust. Meteorol. Mag. 40, 111–121.

Resio, D.T., Westerink, J.J., 2008. Modeling the physics of storm surges. Phys. Today 61 (9), 33–38. https://doi.org/10.1063/1.2982120.

Resio, D.T., Irish, J., Cialone, M., 2009. A surge response function approach to coastal hazard assessment – part 1: basic concepts. Nat. Hazards 51 (1), 163–182. https://doi.org/10.1007/s11069-009-9379-y.

Ribal, A., Young, I.R., 2019. 33 years of globally calibrated wave height and wind speed data based on altimeter observations. Sci. Data 6 (1), 1–15. https://doi.org/10.1038/s41597-019-0083-9.

Ricondo, A., Cagigal, L., Rueda, A., Ripoll, N., Mendez, F.J., 2023. HyWaves: hybrid downscaling of multimodal wave-climate in small Pacific Islands. Ocean Model. 184 https://doi.org/10.1016/j.ocemod.2023.102210.

Rogers, W.E., Babanin, A.V., Wang, D.W., 2012. Observation-consistent input and whitecapping dissipation in a model for wind- generated surface waves: description and simple calculations. J. Atmos. Ocean. Technol. 29 (9), 1329–1346.

Tamizi, A., Young, I.R., 2020. The spatial distribution of ocean waves in tropical cyclones. J. Phys. Oceanogr. 50 (8), 2123–2139. https://doi.org/10.1175/JPO-D-20-0020.1.

Van Vloten, S.O., Cagigal, L., Rueda, A., Ripoll, N., Méndez, F.J., 2022. HyTCWaves: a hybrid model for downscaling tropical cyclone induced extreme waves climate. Ocean Model. 178 (2022), 102100 https://doi.org/10.1016/j.ocemod.2022.102100.

Young, I., 1988. Parametric hurricane wave prediction model. J. Waterw. Port Coast. Ocean Eng. 114 (5) https://doi.org/10.1061/(ASCE)0733-950X(1988)114:5(637).