Latin American Journal of Aquatic Research, 50(4): 573-595, 2022
DOI: 10.3856/vol50-issue4-fulltext-2811

Research Article

A deep water and nearshore wave height calibration of the ECOWAVES hindcasting database

Diego Becerra[1], Matías Quezada[1] & Humberto Díaz[2]
[1]Departamento de Oceanografía Física y Modelamiento Matemático de EcoTecnos S.A.
Viña del Mar, Chile
[2]EcoTecnos S.A., Viña del Mar, Chile
Corresponding author: Diego Becerra (dbecerra@ecotecnos.cl)

ABSTRACT. Significant wave height (SWH) in shallow waters is assessed by generating two wave *hindcasts*; the first uses ERA-Interim wind fields and the second one from ERA5 to quantify the improvement of the ERA5 surface winds on the SWH representativeness, both in deep and shallow waters along the Chilean coastline. Additionally, wind field predictions from the Global Forecast System (GFS) were used to assess the representativeness of shallow waters. Oceanographic buoys were used to validate SWH in deep waters, while Acoustic Doppler Current Profiler (ADCPs) was equipped to measure waves in shallow waters. Energy spectrums coupling Wavewatch III and Simulating Waves Nearshore (SWAN) models were transferred to evaluate the performance of shallow water simulations. In general, the SWH from both wave hindcasts showed good performance. Nonetheless, those forced by ERA5 presented a better qualitative comparison of sea state temporal variability, which increased the correlation coefficients (>0.9), coefficients of determination (>0.8), and minor errors (RMSE, MAE, and BIAS) compared to oceanographic buoys and ADCPs. Additionally, in simulations forced by GFS, the temporal variability of the waves in shallow waters was successfully reproduced. Nevertheless, an increase in the RMSE, MAE, and BIAS error was statistically verified compared to ERA-Interim and ERA 5.

Keywords: *hindcast*; forecast; Wavewatch III; shallow water; surface waves height

INTRODUCTION

Worldwide knowledge regarding the open ocean wave and its generation from wind fields has significantly increased in recent years. It has evolved from the classic parametric models developed over 70 years ago (Arthur 1947, Sverdrup & Munk 1947, Bretschneider 1951) to different computing tools based on state equations (Group 1988, Resio & Perrie 1989, Tolman 1991), which enable accurately define the spectral fields of the waves.

Long-term characteristics of the waves nearshore are one of the vital aspects for the development of projects and studies related to coastal, maritime, and port engineering because temporal and spatial variability of the significant wave height (SWH) has major implications for different natural processes (Komar 1997, Fan et al. 2009, Cavaleri et al. 2012) and also for different business and recreational activities carried out on the coastline (Jiang et al. 2016, Winckler et al. 2017). When wave heights of massive energy occur on the coast, those events are called extreme waves. Depending on the coast configuration, they might cause major structural and socioeconomic impacts (Winckler et al. 2017), frequently one of the conditions for designing maritime engineering projects (Goda 1988).

In order to satisfy the requirement of having available the spectral characteristics of the long-term waves, it is a conventional practice to rely on deep water waves databases produced by *hindcast* models and then numerically transfer them to the site of interest. Nowadays, it is possible to find diverse ocean waves databases around the globe, such as those of different agencies and investigations centers around the world, namely: the National Oceanic and Atmospheric

Corresponding editor: Marcel Ramos

Administration (NOAA), Oceanweather (OWI), Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER), European Centre for Medium-Range Weather Forecasts (ECMWF), Centre for Environment Fisheries and Aquaculture Science (CEFAS), among others. Nevertheless, the scientific efforts to better understand the wave's characteristics are limited on the southern Pacific coast, especially in Chile.

For the Chilean coastline, some examples of database generation regarding deep water waves may be found in the literature. For example, Fournier et al. (2004), applying WAVAD the numerical modeling for global generation (Resio & Perrie 1989), built a *hindcast* database with 40 years of two-dimensional spectra in deep water for the entire South Pacific Ocean. It was subsequently updated in 2005 and 2006, including the Pacific coasts of the USA and Canada, and additionally modifying the numerical model used for the *hindcast* simulation, corresponding to Wavewatch III^{TM} (Tolman 1991).

Efforts after Fournier et al. (2004) were presented by Aguirre et al. (2017) and Beyá et al. (2017), who applying the numerical model Wavewatch III^{TM} developed waves climatology studies in the entire Pacific Ocean basin. Beyá et al. (2017) developed and validated a wave *hindcast* called the Chilean Wave Atlas database in deep water, extending for 36 years (1979-2015). On their part, Aguirre et al. (2017) oriented their study to the description of the wave climatology in deep waters, highlighting the identification of the influence of the southern winds in the significant heights seasonal pattern and also the dependence on the annual cycle due to the seasonal variability of the atmospheric coastal low-level jets off Peru and Central Chile.

Nowadays, different university organizations, research centers, and public entities such as the Hydrographic and Oceanographic Service of the Chilean Navy (SHOA by its Spanish acronym) have conducted efforts to increase the number of oceanographic buoys in different coastal towns in order to measure and understand the seasonal and inter-annual variability of the waves in different latitudes. Nonetheless, the information generated is brief, incomplete, and restricted. Thus, an updated, validated, calibrated wave database is required to characterize wave conditions in a study site. The most recent public and open access wave *hindcast* (on the Chilean coastline) covers the period up to 2015 and corresponds to Beyá et al. (2017). Currently, none of them is validated in shallow waters, which according to these authors' experience, it is of essential importance in order to quantify not only the existence of the more energetic events but also to assess the performance of the deep-water database as the boundary condition for a nearshore wave model which approximates waves towards a certain study site. Thus, considering the energy spectra coupling determined by the global generation model Wavewatch IIITM it is transferred to the specific comparison site using the numerical model of wave propagation Simulating Waves Nearshore (SWAN; Booij et al. 1999). Based on the latter, this article aims to conduct the initial estimates of shallow waters validation for SWH along the Chilean coastline by comparing simulated data to available instrumental records.

Specific objectives of this study are:

- Validate *hindcast* in deep and shallow waters.
- Statistically quantify the differences in forcing the wave model with ERA5 winds *vs.* ERA-Interim winds in the representativeness of the wave conditions on the Chilean coastline.
- Compare each location's wave forecast forced by Global Forecast System (GFS) and the hindcast in shallow waters with field information available.

MATERIALS AND METHODS

Model and observations

Deep waters: wave generation

Beyá et al. (2017), in the search for the best atmospheric reanalysis, showed that ERA-Interim from the ECMWF corresponds to the forcing which best reproduced wave conditions on the Chilean coastline. Nevertheless, a new version of the atmospheric reanalysis was launched called ERA5 (also developed by ECMWF), which corresponds to the new version of ERA-Interim.

Differences between the ERA5 database and ERA-Interim include spatial and temporal structure aspects. The vertical description increased from 60 to 137 levels, while the horizontal resolution increased by reducing the cell sizes from 80 to 31 km. ERA5 assimilated a greater set of observational data than in the previous reanalysis, among other aspects highlighted by Hersbach (2020).

According to the above wave generation and propagation model, Wavewatch III^{TM} version 6.07 (WW3) was used, a third-generation numerical model widely used by the scientific community worldwide to study wave conditions in deep waters. The information related to the sea surface is contained in the energetic

variance spectrum E (f, θ), which distributes the energy in the frequency domain *f* y direction θ (Tolman 2019).

Through WW3, two wave *hindcasts* were conducted: the first, ECOWAVES, from January 1979 to August 2019 (~40 years). The second, ECOWAVES 2.0, included from January 1979 to December 2020 (~41 years). The reason for building ECOWAVES 2.0 is due to the improved performance of the ERA5 atmospheric analysis reported by Belmonte & Stoffelen (2019), who compared the observations of winds from the Advanced Scatterometer (ASCAT) provided by NOAA with both ERA5 and ERA-Interim, concluding that ERA5 performs better than its predecessor ERA-Interim. Similar results were obtained by Gelaro et al. (2017), who compared MERRA-2 and ERA5 for their use in wind energy, determining that ERA5 works better. Ramon et al. (2019), meanwhile, to identify the atmospheric reanalysis products that best represent the characteristics of surface wind speeds, analyzed the five latest generations of global reanalysis, among which they considered ERA5, ERA-Interim, JRA55, MERRA-2, and R1, finding that ERA5 was able to reproduce with a better performance the variability observed by meteorological stations. In the same way, Hersbach et al. (2020) highlight the advantages ERA5 has compared to ERA-Interim, and the improvements achieved in the representativeness of wave conditions. Therefore, ECOWAVES 2.0 was carried out to quantify the differences in the wave spectral fields due to the updated and improved base of surface winds that ERA5 has (Hersbach et al. 2020).

Both databases have a 3 h temporal resolution for statistical parameters across the Pacific Ocean basin. The energy spectra extracted from the model also have a temporal resolution of 3 h.

In order to create the digital model of the seabed, it was used both the bathymetry information and obstructions masks to represent the island's presence from the ETOPOv2 database provided by NOAA through the National Center for Environmental Information. This information was processed for the construction of the numerical domain by using the Gridgen generation grids algorithm (Chawla & Tolman 2013). A spatial resolution of 1°×1° was considered, in the numerical domain construction, covering 64°N-64°S and 60°-110°W (Fig. 1a). This configuration was applied for both ECOWAVES and ECOWAVES 2.0.

For ECOWAVES, wind fields at 10 m height and the sea ice concentration were obtained from the atmospheric reanalysis of ERA-Interim (Dee et al. 2011), which were used to force wave simulations with a spatial resolution of 0.5°×0.5° and temporal resolution of 6 h. In ECOWAVES 2.0, surface wind fields and sea ice were obtained from the atmospheric reanalysis of ERA5 (Hersbach 2020) with the same temporal and spatial resolution used for ECOWAVES.

The spectral domain of both wave *hindcasts* was discretized in 29 frequencies from 0.0345 to 0.4975 Hz, increased by 10%, and 24 directions regularly spaced by 15°.

It is worth mentioning that the source term compilations of the Wavewatch III model used to create ECOWAVES and ECOWAVES 2.0 were developed according to the recommendation proposed by IFREMER in the IOWAGA Project (Ardhuin 2011) and also used by third parties in the construction of the wave database for the Chilean coast, as the case of Aguirre et al. (2017) and Beyá et al. (2017). These last publications show that the combination of physical parameters in the "Switch_Ifremer1" file presents the most representative outputs of the reality on the Chilean coastline. More relevant parameterizations used are shown in Table 1.

In the calibration stage of ECOWAVES and ECOWAVES 2.0, Beyá et al. (2017) recommendations were considered, which consisted in modifying the dimensionless parameters related to the packages of source terms associated with the processes of growth and dissipation of wave energy due to the wind. For growth (Eq. 1) it was considered $\beta_{max} = 1.52$ (Ardhuin et al. 2011) and $Z_{0,max} = 1.002$. Meanwhile, for dissipation (Eq. 6) it was considered $C_{ds}^{sat} = -0.000022$ and $B_r = 0.0009$.

$$S_{in}(k,\theta) = \frac{\rho_a}{\rho_w}\frac{\beta_{max}}{k} e^z Z^4 \left(\frac{u_*}{C} + z_\alpha\right) cos^{Pin}(\theta - \theta_u)\sigma N(k,\theta) \quad (1)$$

where ρ_a and ρ_w correspond to the densities of air and water, β_{max} is a constant dimensionless growth parameter (calibration parameter), κ is the Von Kármán constant, and Pin is a constant equal to 2 that controls the directional distribution of S_{in} (Tolman 2019). σ is the intrinsic wave frequency, C is the phase celerity of a specific component, u_* is the wind shear speed, z_α is the Charnock number, which enables to quantify roughness of the sea surface, θ is the wave propagation direction, θ_u is the wind direction and $\sigma N(k,\theta)$ is the energy density of a wave component (Ardhuin et al. 2010). The effective wave age Z = log (μ) where Janssen (1989) gives μ and corrected for intermediate water depths (Ardhuin et al. 2010), so that:

$$Z = \log(kz_1) + \frac{\kappa}{\left[\cos(\theta-\theta_u)\left(\frac{u_*}{C}+z_\alpha\right)\right]} \quad (2)$$

where z_1 is the roughness length modified by the stress t_w supported by the wave, and z_α is an adjustment

Table 1. Summary of the parameterizations included in Switch_IFREMER1.

Process	Switch	Observation
Nonlinear energy dissipation and growing due to wind	ST4	Ardhuin et al. (2010)
Hardware model	DIST	Distributed memory with message passing interface.
Message passing protocol	MPI	
Bottom friction	BT4	Bottom friction graph SHOWEX Ardhuin et al. (2003)
Nonlinear interactions	NL1	Discrete Interaction Approximation (DIA) Hasselmann et al. (1985)
Lineal growing of energy due to wind	SEED	Tolman (2019)
Garden Sprinkler effect Alleviating technique	PR3	Averaging Technique Tolman (2002)
Propagation scheme	UQ	Third-order propagation scheme Leonard (1975), Davis & More (1982).
Rupture induced by depth	DB1	Battjes & Janssen (1978)

parameter of the age of the wave. z_1 is implicitly defined by:

$$U_{10} = \frac{u_*}{\kappa}\log\left(\frac{z_u}{z_1}\right) \quad (3)$$

$$z_0 = min\left\{\alpha_0\frac{\tau}{g}, Z_{0,max}\right\} \quad (4)$$

$$z_1 = \frac{z_0}{\sqrt{1-\tau_w/\tau}} \quad (5)$$

where z_u is the height at which the wind velocity is specified (typically 10 m).

The physical parameterization of wave growth imposes a maximum or upper limit of the Charnock number Z_{0max} to control the efforts (overestimated) that result when the transfer of *momentum* during storms is quantified (Ardhuin et al. 2010).

For the physical parameterization of wave dissipation by *white capping* proposed by Ardhuin et al. (2010), the concept of spectral saturation threshold B_r was introduced. This threshold represents a dimensionless energy level that indicates the beginning of dissipation.

$$S_{ds}(k,\theta) = \sigma\frac{C_{ds}^{sat}}{B_r^2}[\delta_d \max\{B(k)-B_r\}^2 + (1-\delta_d)\max\{B'(k,\theta)-B_r,0\}^2]\,N(k,\theta) \quad (6)$$

where $B'(k,\theta)$, and $B(k)$ correspond to the directional and isotropic spectral saturation, respectively. C_{ds}^{sat} is a dimensionless calibration constant, and δ_d is a weighting factor that allows the user to control the contribution of directional and isotropic wave dissipation (Ardhuin et al. 2010).

Thirty-eight nodes (virtual buoys) were extracted for each wave *hindcast* with energy spectra located off the central-south Pacific coast from 6°N to 56°S every 2° latitude (Fig. 1b), opposite the insular regions of Chile. It should be noted that the temporal resolution of the energy spectra was 3 h.

Additionally, *forecasts* simulations were carried out by forcing the WW3 model with wind fields extracted from the GFS, using temporal and spatial resolutions of 6 h and 0.5°×0.5°, respectively. The simulations were conducted two months before the records were measured by each available Acoustic Doppler Current Profiler (ADCP) instrument for this research. It is important to note that due to the discontinuities of the GFS wind database, it was impossible to run a single case for all the ADCP locations.

The general purpose of this simulation was to verify if the surface winds, without being subjected to reanalysis, adequately reproduce the characteristics of the waves in shallow waters at different latitudes of the Chilean coast. It is important to highlight that the numerical simulations of deep water forced by GFS winds were named WW3 GFS (from now on).

It should be noted that as a computational resource to carry out the numerical simulations of the WW3 model, 120 cores with 48 GB of RAM were used in the cluster belonging to the National Laboratory for High-Performance Computing (NLHPC) (https://www.nlhpc.cl/) located in the facilities of the Faculty of Physical Sciences and Mathematics from the University of Chile.

Shallow water model: energy transfer

The SWAN model developed by the Delft University of Technology corresponds to a spectral model for wave propagation that solves the transport equation for the spectral wave action density. This equation considers source terms and parameterizations that explain wave-wave interactions, generation, and dissipation in deep and shallow waters (Booij et al. 1999, Rogers et al. 2006, Sang-Ho et al. 2009, SWAN

team 2019). The spectral action balance equation in Cartesian coordinates is shown according to Eq. 7:

$$\frac{\partial N}{\partial t} + \frac{\partial C_{g,x} N}{\partial x} + \frac{\partial C_{g,y} N}{\partial y} + \frac{\partial C_{g,\sigma} N}{\partial \sigma} + \frac{\partial C_{g,\theta} N}{\partial \theta} = \frac{S}{\sigma} \quad (7)$$

where σ is the relative frequency (the frequency of wave measured from a reference frame), N is the wave action density, which corresponds to the quotient between the energy density and the relative frequency (N = E / σ), θ is the direction of the waves, C_g is the speed of propagation of wave action in space (x, y, σ, θ), and S is the total source/sink terms expressed as wave energy density. In deep water, the right side of Equation 7 is dominated by three terms, $S \sim S_{in} + S_{nl} + S_{ds}$, corresponding to wind input, four-wave nonlinear interactions, and dissipation, respectively (SWAN Team 2019).

Additionally, SWAN uses typical formulations for wave growth by surface wind action, *whitecapping* and nonlinear quadruple interactions, similar to WW3 (Rogers et al. 2006, SWAN team 2019). The most important difference compared to the WW3 model is that SWAN includes physical processes that occur in shallow waters, for example, shoaling, dissipation of energy by bottom friction, breaking due to limited depth, refraction, reflection, and diffraction in some cases (Rogers et al. 2006, SWAN Team 2019).

SWAN simulates wave diffraction by incorporating the directional turning rate term δ_E (Eq. 8), which was developed by Holthuijsen et al. (2003) using the mild-slope equation and the spectral energy balance, resulting in:

$$\delta_E = \frac{\nabla(CC_{g*}\nabla\sqrt{E})}{k^2 CC_g\sqrt{E}} \quad (8)$$

Equation 8 allows the simulation of energy transfer in coastal areas where diffraction is important, such as wave fields around geophysical features (e.g. rocks, islands) or coastal structures. Where k is the wave number, C is the celerity of the wave, Cg is the group celerity, and E is the energy of waves and ∇ corresponds to the velocities gradient. In this study, SWAN version 40.31 was used.

SWAN was configured based on two nested grids with different spatial resolutions. The first was based on square elements, structured at 200 m per side, while the detail domain was discretized in square elements of 50 m per side to minimize calculation times. Therefore, the grid presents lower resolution in areas of less spatial variability and a higher resolution near shore. The spatial domains of the lower resolution grids were defined so that the boundaries are in deep water (Fig. 1c), so the spectral outputs of the WW3 generation model are valid as a boundary condition of domains built in SWAN. In Figure 1c, it is presented the latitudinal distribution of the domains used in this study, which have been distributed to approximate the swell from deep waters towards shallow waters; for each location in which local measurements of ADCP were available, they are described in section 2.1.3.

ECOWAVES, ECOWAVES 2.0, and WW3 GFS deep-sea spectra were used as boundary conditions. However, it is important to note that, in this modeling of wave propagation to shallow waters, the surface wind was not considered, so all the source terms related to the action of the wind were deactivated. The wind was not included in the wave propagation simulation because there was no high-resolution wind database for each of the implemented meshes. It is important to mention that this topic will be the subject of future research by the present authors.

In situ observations

Measurements of deep and shallow waters were available, broadly described below, to perform the comparison and subsequent statistical analysis of the model results.

In deep waters, a total of 19 oceanographic buoys (Fig. 2a) were freely accessible from the National Data Buoy Center (NDBC) belonging to the NOAA, which are located in different geographical sites distributed mainly in the northeast Pacific Ocean (17 buoys) and southeast (2 buoys). Additionally, spectral data from 2 buoys were available from a wave energy evaluation study off the coast of Chile (CORFO-INNOVA 2009) and also used by Beyá et al. (2017). These buoys located in intermediate waters off the Chilean coast (Fig. 2a) were used as field backgrounds for comparison with the numerical results. The name of the two buoys will be abbreviated as follows; buoy 1 will be C-INNOVA 1, and buoy 2 will be C-INNOVA 2 (Table 2). Table 2 presents general information on the buoys used to validate the wave hindcasts.

Additionally, 18 measurements of the waves in shallow water to a depth of fewer than 20 m were available (Table 3). These wave measurements were conducted by using ADCP instruments. The geographical distribution of the buoys and the latitudinal distribution of the ADCPs along the Chilean coastline are shown (Fig. 2b).

Validation

Two validation procedures were conducted to the generated wave databases: one in deep waters and another in shallow waters. Each of the validation stages

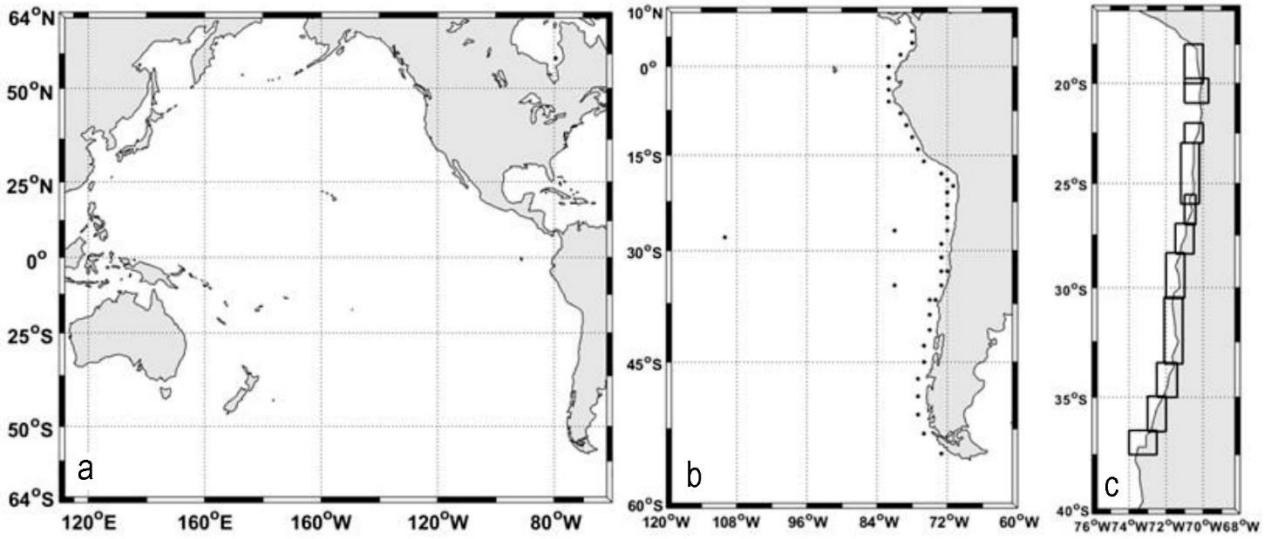

Figure 1. a) The numerical domain of WAVEWATCH III model, b) location of virtual nodes with energy spectra of both wave *hindcast*, c) numerical domains (black rectangles) used in SWAN.

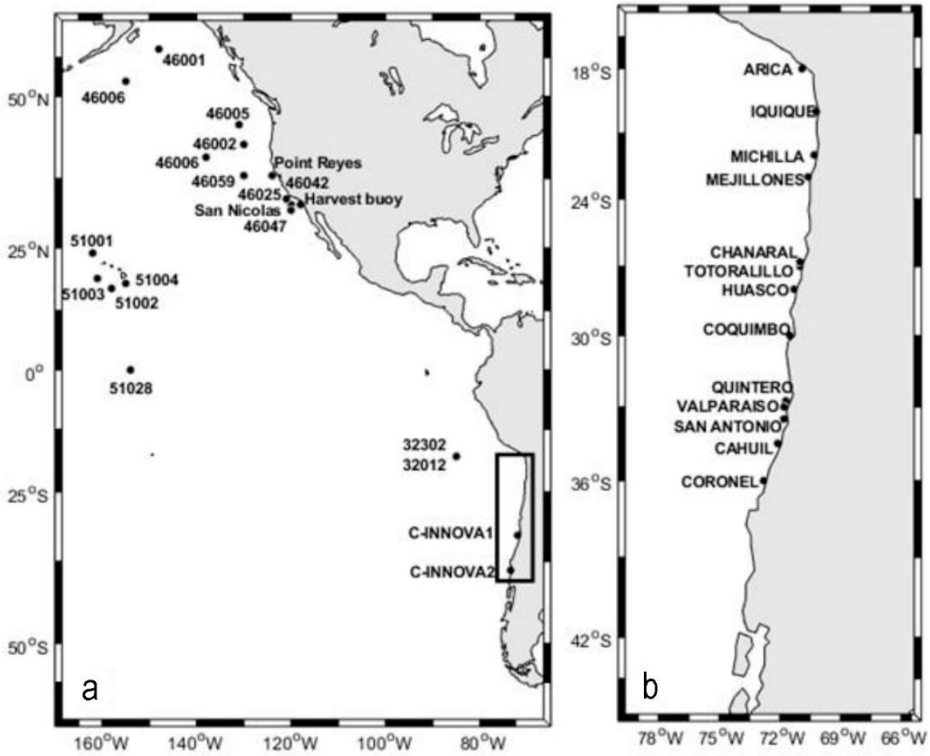

Figure 2. a) Oceanographic buoys location, and b) Acoustic Doppler Current Profiler.

is aimed at different purposes. In the case of deep waters, to verify that wave conditions obtained directly from the WW3 model (with different configurations) were comparable to instrumental records and that might be used as a boundary condition in order to serve the second purpose of the validation, that is, to generate a wave approximation model towards shallow waters.

Validation in deep water was conducted by extracting energy spectra from ECOWAVES and ECOWAVES 2.0 in the same geographical coordinates of the oceanographic buoys (or nearby), which are indicated, and for the same periods in which each of the monitoring stations recorded the wave conditions (Table 2). Subsequently, different statistical parameters

Table 2. Buoys data used to validate wave *hindcast* in the Pacific Ocean, in intermediate and deep waters. *Indicate buoys placed in the southeastern Pacific.

Buoy	Latitude (°S)	Longitude (°W)	Start time	End time
Buoy 46001	56°18.0'	148°10.2'	1979-06-01	2005-08-31
Buoy 46005	46°3.00'	131°1.20'	1979-06-01	2004-12-26
Buoy 46006	40°50.4'	137°30.0'	1981-08-08	2005-08-31
Buoy 46025	33°45.0'	119°4.80'	1982-04-21	2005-07-31
Buoy 46042	36°45.0'	122°25.2'	1988-01-01	2005-08-31
Buoy 46047	32°25.8'	119°31.8'	1999-06-01	2005-08-31
Buoy 46059	37°58.8'	130°0.00'	1994-10-19	2005-08-31
Buoy 46066	52°12.0'	155°24.0'	2000-05-13	2005-08-31
Buoy 51001	23°25.8'	162°12.6'	1981-02-11	2005-08-31
Buoy 51002	17°08.4'	157°47.4'	1984-09-07	2005-08-31
Buoy 51003	19°09.6'	160°44.4'	1984-11-01	2005-08-31
Buoy 51004	17°31.2'	152°28.8'	1985-02-13	2005-08-31
Buoy 51028	00°12.0'	153°54.0'	1997-11-01	2005-08-31
Buoy Harvest	34°27.6'	120°42.0'	1998-03-20	2005-08-31
Buoy Point Loma	32°37.8'	117°27.0'	1996-05-01	2005-08-31
Buoy Point Reyes	37°57.0'	123°28.2'	1996-12-07	2005-08-31
Buoy San Nicolás	33°13.2'	119°49.8'	1999-09-02	2005-08-31
32302*	18°00.0'	85°00.00'	1986-02-01	1989-12-31
32012*	19°42.0'	85°07.00'	2007-10-29	2017-12-31
C-INNOVA1*	33°13.0'	71°82.00'	2011-06-16	2013-12-21
C-INNOVA2*	39°55.0'	73°40.00'	2012-11-21	2013-07-30

were calculated to determine the performance of the simulations in contrast to measurements. Statistical parameters recommended by Williams & Esteves (2017) were used in this study (Table 4), where M_i and O_i correspond to modeled and observed data in time, accordingly, and N is the quantity of data matching both time series, Cov (O, M) is the covariance between O and M, and S_o and S_m are the standard deviations of O and M, accordingly.

Furthermore, in order to assess SWH in shallow waters, coupling between WW3 and SWAN models was conducted. This process consisted of selecting the energy spectra in deep waters from ECOWAVES, ECOWAVES 2.0, and WW3 GFS and rearranging them automatically to be used as a boundary condition of the SWAN model to transfer the waves to shallow waters considering the coastline and high-resolution bathymetries. The high-resolution bathymetries were built from the digitization of the SHOA nautical charts in the Global Mapper software. They were later interpolated to build the grid used in the SWAN coastal model.

This spectral transfer method enabled us to ascertain and understand the spatial variability on the study site and verify the validity of using the database in deep waters as a boundary condition for developing numerical studies for coastal wave characterization.

Due to the short length of ADCPs records (Table 3) and the spectral matrixes having 29 frequencies per 24 directions, transferring spectra into SWAN did not generate substantial computing costs, and the validation process for 18 coastal stations was achieved.

Performance of simulations in shallow waters was assessed through qualitative comparisons, examining whether the simulations represent the physics involved in the wave height temporal variation and quantitatively through the calculation of statistical parameters previously mentioned in the deep water validation procedure. Additionally, through boxplot, medians, 25 and 75% percentiles, outliers from the measured data, and those obtained from the numerical simulation were compared. It is important to highlight that SWAN was used in stationary mode, meaning that temporal variation in the wave propagation was not considered. Thus, temporal corrections, described below, were made to the model results to perform a proper comparison against ADCP measurements.

In simulations coupled with wave hindcast, an advance in the arrival of 9 to 12 h was observed on average. In simulations coupled with the forecast, an average advance of 12 to 15 h in the arrival was observed. The cross-correlation was calculated to quantify the lag between the series.

Table 3. Acoustic Doppler Current Profiler date used to validate simulations (*hindcast* and *forecast*) in shallow waters.

Site name	Latitude (°S)	Longitude (°W)	Start time	End time	Depth (m)
Arica (winter)	18°29.987'	70°19.771'	2017-07-05	2017-08-08	17.8
Arica (summer)	18°29.987'	70°19.771'	2018-01-09	2018-02-14	17.8
Port Iquique (winter)	20°12.053'	70°09.071'	2015-07-02	2015-08-05	16.3
Port Iquique (summer)	20°12.053'	70°09.071'	2015-01-17	2015-03-02	16.5
Offshore Iquique (winter)	20°12.823'	70°09.675'	2015-07-02	2015-08-05	17.2
Offshore Iquique (summer)	20°12.823'	70°09.675'	2015-01-17	2015-03-02	17.2
Michilla	22°43.479'	70°17.494'	2007-01-05	2007-02-11	12.3
Mejillones	23°04.528'	70°24.033'	2009-04-21	2009-05-27	14.5
Chañaral	26°19.789'	70°39.326'	2014-06-27	2014-07-28	19.3
Totoralillo	26°51.047'	70°48.952'	2004-08-30	2004-10-08	10.0
Huasco	28°28.232'	71°14.648'	2006-02-18	2006-03-25	11.0
Coquimbo	29°57.090'	71°21.917'	2018-08-21	2018-09-22	28.1
Quintero	32°45.100'	71°29.358'	2019-06-20	2019-07-18	15.2
Valparaíso	33°00.802'	71°33.706'	2018-06-28	2018-07-19	16.4
San Antonio	33°34.925'	71°37.482'	2006-08-04	2006-09-06	13.1
Cahuil (spring)	34°28.845'	72°02.125'	2016-08-09	2016-09-21	17.6
Cahuil (winter)	34°28.845'	72°02.125'	2017-06-13	2017-07-17	18.2
Coronel Biobío	36°58.099'	73°10.712'	2016-07-06	2016-08-27	14.1

Table 4. Statistical parameters used to assess simulation performance.

Statistical parameters	Formulae
Correlation coefficient	$R = \frac{\sum_{i=1}^{N}\left(\left(M_i - media(M)\right) * \left(O_i - media(O)\right)\right)}{N * S_o * S_m}$
Determination coefficient	$R^2 = \frac{Cov(O,M)^2}{S_o^2 S_m^2}$
Root-mean-square-error	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)^2}$
BIAS	$BIAS = \frac{1}{N}\sum_{i=1}^{N} M_i - O_i$
Predictive ability	$SS = 1 - \sqrt{\frac{\frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)^2}{\frac{1}{N}\sum_{i=1}^{N} {O_i}^2}}$
Scatter index	$SI = \frac{RMSE}{\frac{1}{N}\sum_{i=1}^{N} O_i}$

In essence, three coupling for each study site were conducted:

1) ECOWAVES + SWAN

2) ECOWAVES 2.0 + SWAN

3) WW3 GFS + SWAN

Table 5. Sites with no surface wind GFS available. ADCPs: Acoustic Doppler Current Profiler.

ADCPs	Start time	End time
OFFSHORE Iquique (winter)	2015-07-02	2015-08-05
PORT Iquique (winter)	2015-07-02	2015-08-05
Huasco	2006-02-18	2006-03-25
San Antonio	2006-08-04	2006-09-06
Cahuil (spring)	2016-08-09	2016-08-09
Coronel	2016-09-21	2016-09-21

Nonetheless, there were periods in which no surface wind GFS was available (Table 5); thus, in these sites, coupling was not possible, and consequently, nor their corresponding statistical analysis. Statistical performance of simulated SWH was assessed with ECOWAVES and ECOWAVES 2.0 in deep waters, and the performance resulting from the simulations coupled with SWAN against ADCPs measurements.

RESULTS

Deep water validation

A summary of the statistical parameters calculated to evaluate the performance of both *hindcast* databases in the central-north Pacific is presented (Table 6). It is observed that ECOWAVES can statistically represent the SWH obtaining linear correlations (R) values varying from 0.85 to 0.96 and coefficients of determination (R^2), indicating that ECOWAVES accounts for more than 70% of the measurements (Table 6). Additionally, RMSE and MAE lower than 60 cm were observed, which indicates a good agreement.

However, SWH results forced by ERA5 (ECOWAVES 2.0) showed a significant increase in R, R^2, and predictive ability factor (S.S.), and also a reduction of RMSE, MAE and BIAS errors (Table 6). For example, the Harvest buoy statistical agreement represented an increase in R from 0.85 to 0.93 and an RMSE reduction from 0.64 to 0.34 m, which indicates that the best agreements in deep waters were achieved with the forced wave model with surface wind fields of ERA5.

Results obtained for the North Pacific Ocean were associated with the NOAA denomination for buoys localization. Grouping them in a Taylor diagram (Taylor 2001) for the far north and Alaska (Fig. 3), south-west quadrant (Fig. 4), and Hawaii proximity (Fig. 5). In each of the illustrations, the highest recommended RMSE (Williams & Esteves 2017) has been incorporated (dashed red line) for each comparison conducted.

In the far north and Alaska (Fig. 3), both databases generated by the *hindcast* technique (ECOWAVES and ECOWAVES 2.0) obtained correlation coefficient values equal to or higher than 0.95 and RMSE lower than the maximum recommended, also showing a better performance of ECOWAVES 2.0 in all compared buoys. The difference in each simulation RMSE and the maximum recommended RMSE was about 2.5 m for each analyzed site, which shows the great accuracy in wave height estimation which both ECOWAVES and ECOWAVES 2.0 obtained for the far north and Alaska in the North Pacific Ocean.

Standard deviations of numerical simulations associated with each instrumental measurement were presented in the same range for each database compared in the far north and Alaska. For example, buoy 46001, instrumentally, showed standard deviations of wave heights close to 1.4 m, while numerical simulations 1.3 m for ECOWAVES and 1.2 m for ECOWAVES 2.0. Except for ECOWAVES in buoy 46066, the standard deviation of the simulated data presented lower magnitudes than the measurements, which would reflect that the numerical modeling developed would present lower fluctuations of wave heights compared to the evidenced in measurements.

The results obtained for the lower latitudes of the northern hemisphere (Fig. 4) presented statistical behaviors similar to those described previously for the far north. They satisfactorily represented the measurements available for the four buoys analyzed (46025, 46042, 46047, and 46059). In this quadrant, it could be observed that the standard deviation of the wave heights, both instrumental and numerical, were less than the unit (except for buoy 46059), which identifies the least fluctuation of the wave heights for the analyzed period in comparison with what happened in the far north and Alaska. In all buoys south of the North Pacific, RMSE is lower than those recommended by Williams & Esteves (2017); that is, the errors obtained in the numerical modeling of both ECOWAVES and ECOWAVES 2.0 were low enough to consider that the numerical model is calibrated.

Table 6. Statistical parameters of ECOWAVES and ECOWAVES 2.0 *vs.* NDBC buoys in the north-central Pacific Ocean. RMSE: root-mean-square-error, SI: scatter index, MAE: mean absolute error, BIAS: the average of the error.

Buoy name	SWH [m] ECOWAVES								SWH [m] ECOWAVES 2.0						
	Data	R	R^2	RMSE	SI	SS	MAE	BIAS	R	R^2	RMSE	SI	SS	MAE	BIAS
Buoy 46001	69188	0.94	0.88	0.49	0.18	0.84	0.36	0.07	0.96	0.93	0.42	0.15	0.87	0.30	-0.03
Buoy 46005	63171	0.96	0.92	0.42	0.15	0.87	0.31	0.11	0.97	0.94	0.38	0.14	0.88	0.28	0.06
Buoy 46006	55051	0.96	0.93	0.42	0.15	0.87	0.31	0.05	0.97	0.94	0.40	0.14	0.88	0.29	0.02
Buoy 46025	59016	0.86	0.73	0.30	0.25	0.77	0.24	0.14	0.87	0.76	0.26	0.22	0.80	0.21	0.09
Buoy 46042	46987	0.93	0.86	0.37	0.17	0.85	0.28	0.01	0.93	0.87	0.37	0.17	0.85	0.28	0.14
Buoy 46047	17178	0.91	0.83	0.43	0.20	0.82	0.32	-0.15	0.93	0.87	0.35	0.16	0.85	0.27	-0.01
Buoy 46059	29612	0.96	0.92	0.38	0.14	0.87	0.28	0.02	0.97	0.94	0.34	0.12	0.89	0.25	0.01
Buoy 46066	12616	0.94	0.88	0.53	0.18	0.84	0.38	0.26	0.96	0.92	0.43	0.15	0.87	0.32	0.17
Buoy 51001	57838	0.93	0.86	0.39	0.16	0.85	0.27	-0.16	0.93	0.87	0.37	0.15	0.86	0.25	-0.10
Buoy 51002	53220	0.90	0.82	0.30	0.12	0.88	0.22	-0.12	0.92	0.85	0.26	0,11	0.90	0.19	-0.01
Buoy 51003	51927	0.90	0.80	0.35	0.16	0.85	0.26	-0.16	0.92	0.85	0.28	0.12	0.88	0.20	-0.05
Buoy 51004	49104	0.89	0.80	0.31	0.13	0.87	0.21	-0.10	0.91	0.83	0.27	0.11	0.89	0.19	0.01
Buoy 51028	19060	0.87	0.75	0.23	0.12	0.88	0.18	0.13	0.85	0.72	0.27	0.14	0.86	0.23	0.20
Buoy Harvest	20648	0.85	0.73	0.64	0.30	0.72	0.18	0.13	0.93	0.87	0.34	0.16	0.85	0.27	0.14
Buoy Point Loma	12453	0.87	0.76	0.27	0.22	0.79	0.48	-0.41	0.92	0.84	0.29	0.23	0.78	0.25	0.22
Buoy Point Reyes	23393	0.92	0.84	0.41	0.16	0.85	0.23	0.19	0.94	0.89	0.36	0.15	0.86	0.29	0.15
Buoy San Nicolás	13986	0.90	0.71	0.36	0.17	0.84	0.31	0.01	0.94	0.89	0.29	0.14	0.87	0.23	0.10

Wave conditions near Hawaii, which corresponds to the central area of the North Pacific, like the previously described areas, showed a high statistical agreement between the numerical simulations and the buoys' measurements. The correlation coefficients were in all cases higher than 0.90, with RMSE below the recommended maximum and standard deviations slightly lower than those obtained from the measured data.

This highly concordant behavior in deep waters for the northern hemisphere of the Pacific Ocean is a descriptor that both the ECOWAVES and ECOWAVES 2.0 *hindcast* databases show the statistical characteristics of the waves related to their wave heights.

Additionally, the results in the southeastern Pacific also showed a good statistical fit of ECOWAVES with R greater than 0.88 and R^2 greater than 75%, which indicates a strong statistical agreement (Table 7). However, ECOWAVES 2.0 presented an increase in statistical performance since it manages to increase correlations (>0.91) and reduces errors, which implies a better data-to-data agreement of the model (Table 7).

It should be noted that the resulting BIAS with ECOWAVES 2.0 was adjusted to the maximum limit recommended (<0.15 m) by Williams & Esteves (2017) in the statistical guidelines of the calibration standards for wave models.

By summarizing the information in Table 7 in a Taylor diagram (Fig. 6) and comparing it with the recommended RMSE limits suggested by Williams & Esteves (2017), it can be observed that, in all the buoys in the southeastern Pacific, the RMSE obtained by modeling are lower than the maximum limit suggested in the literature for wave modeling. Additionally, the closeness of the points (green, blue, and black) denotes the high statistical similarity between the compared databases and the determined for the North Pacific. It is expected that the databases built from ECOWAVES and ECOWAVES 2.0 will allow us to know the main characteristics of wave heights in deep waters.

The scatter plots between ECOWAVES and the four buoys in the southeastern Pacific are shown in Figure 7. It is observed that ECOWAVES succeeds in statistically reproducing the SWH with high accuracy, which can be increased even more when the wind fields from ERA5 are used as forcing since not only the correlation coefficients are increased, but also the RMSE, MAE, and BIAS errors decreased (Fig. 8).

Thus, if the results in deep waters are considered, it can be observed that both ECOWAVES and ECOWAVES 2.0 correspond to databases that statistically represent the swell in deep waters to be used as boundary conditions of spectral transferring models and whose purposes are the propagation of the waves towards shallow waters.

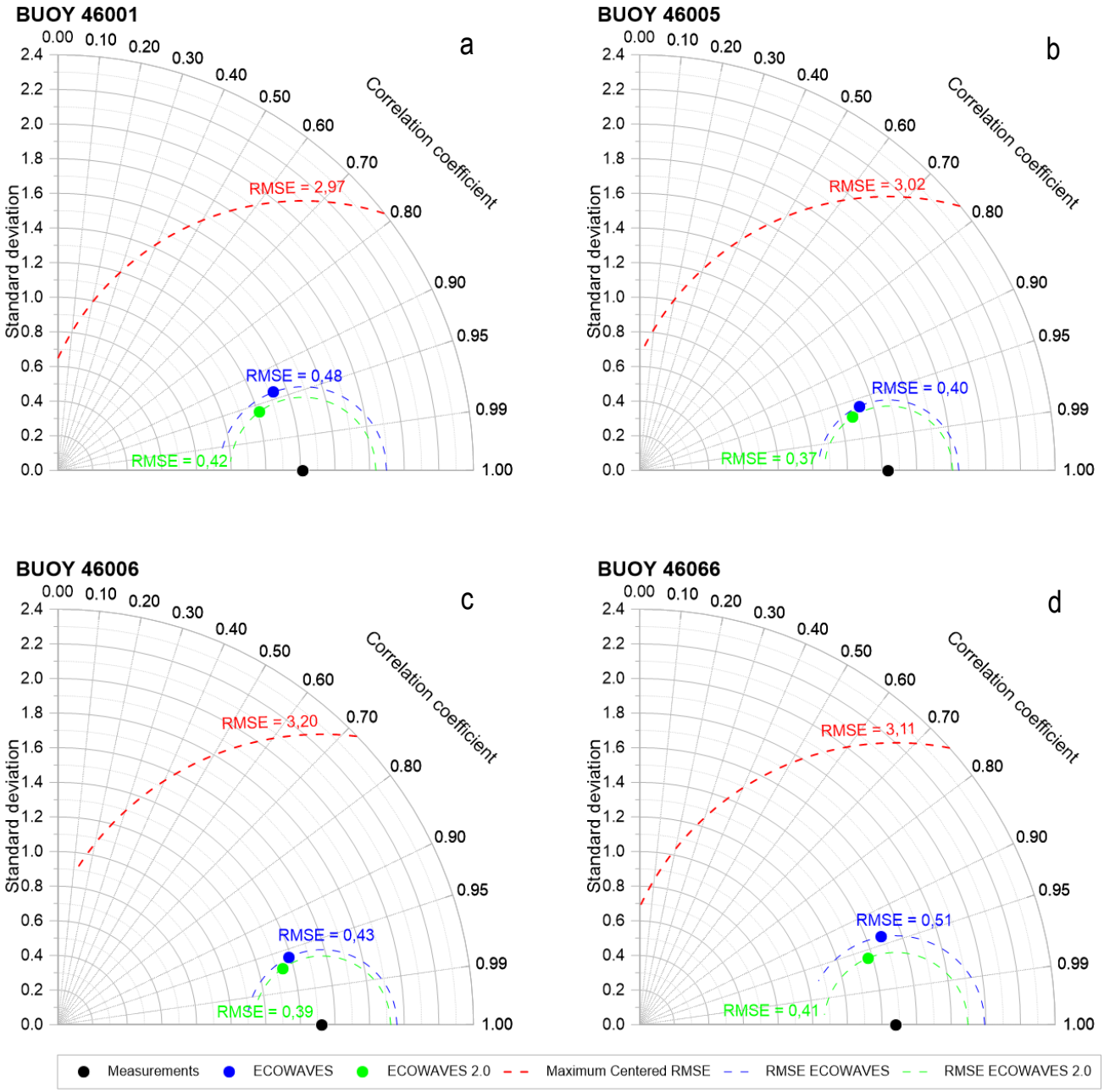

Figure 3. Diagram of the statistical pattern comparison between the measured and simulated data, for the NOAA distribution of buoys in the north and Alaska quadrant, in the North Pacific Ocean. a) Buoy 46001 *vs*. ECOWAVES and ECOWAVES 2.0, b) buoy 46005 *vs*. ECOWAVES and ECOWAVES 2.0, c) buoy 46006 *vs*. ECOWAVES and ECOWAVES 2.0, and d) buoy 46066 *vs*. ECOWAVES and ECOWAVES 2.0.

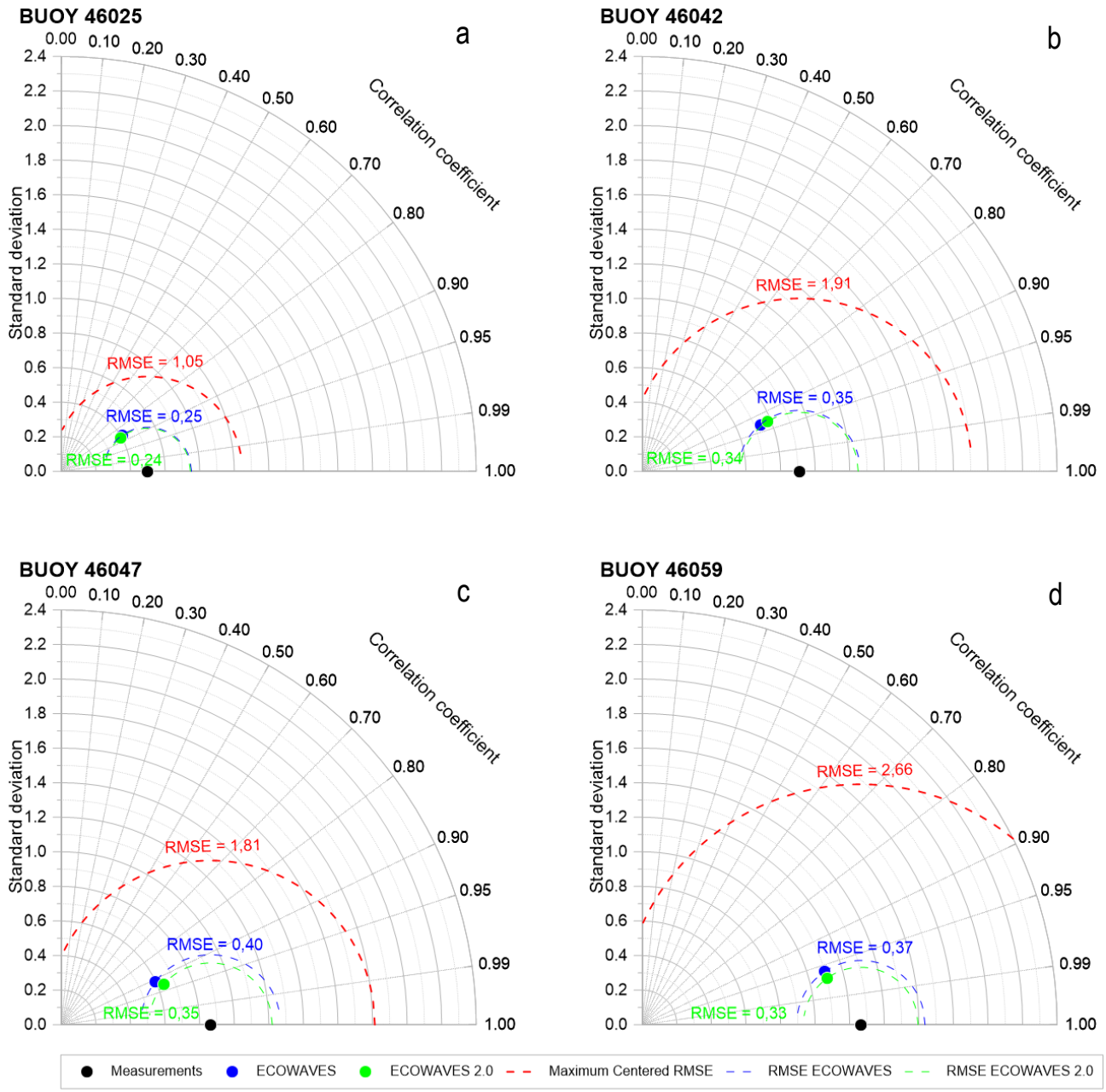

Figure 4. Comparison diagram of the statistical pattern between measured and simulated data for the NOAA buoy distribution in the southwest quadrant of the North Pacific Ocean. a) Buoy 46025 *vs*. ECOWAVES and ECOWAVES 2.0, b) buoy 46042 *vs*. ECOWAVES and ECOWAVES 2.0, c) buoy 46047 *vs*. ECOWAVES and ECOWAVES 2.0, and d) buoy 46059 *vs*. ECOWAVES and ECOWAVES 2.0.

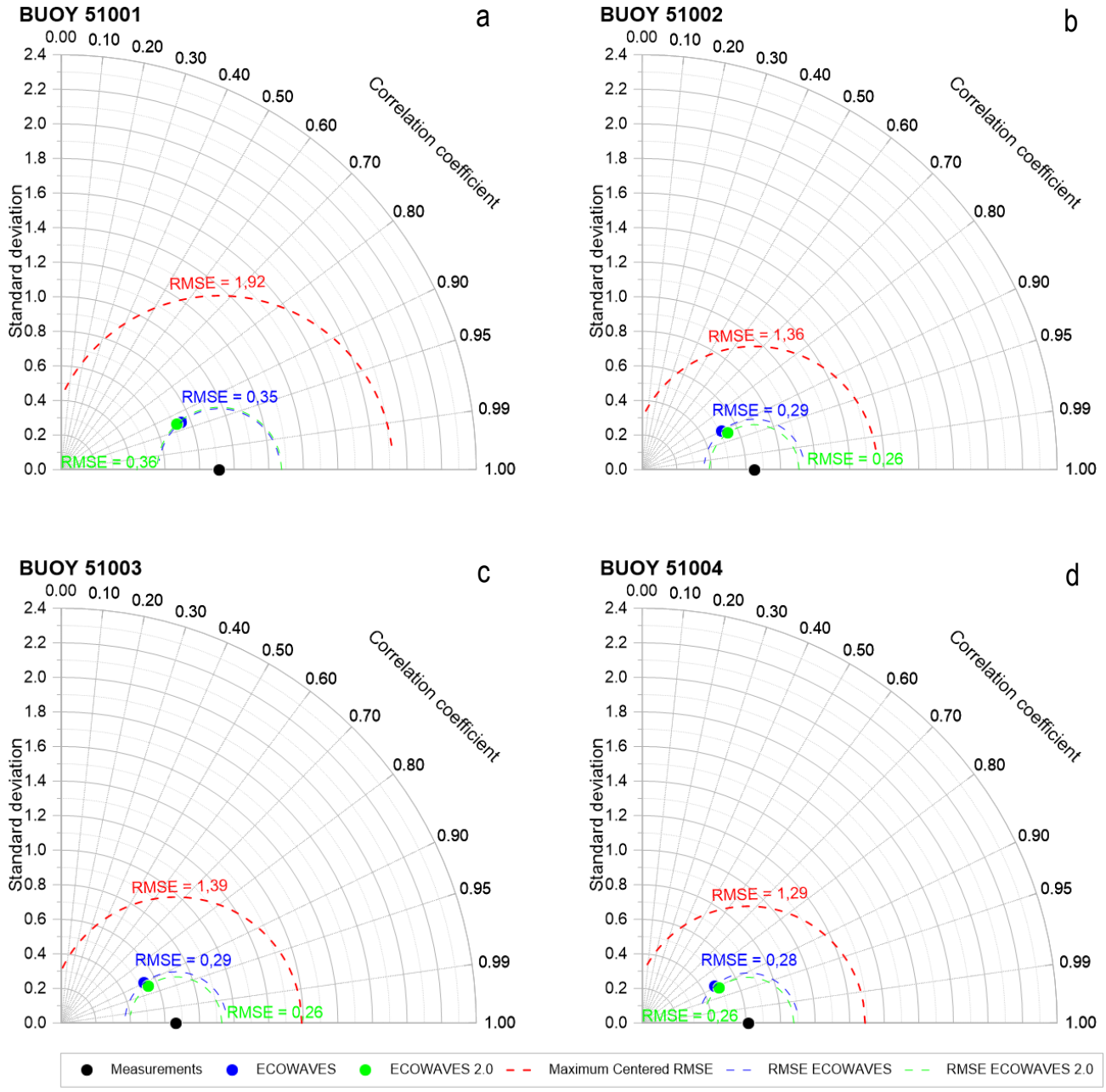

Figure 5. Comparison diagram of the statistical pattern between measured and simulated data for the NOAA buoy distribution in Hawaii of the North Pacific Ocean. a) Buoy 51001 *vs*. ECOWAVES and ECOWAVES 2.0, b) buoy 51002 *vs*. ECOWAVES and ECOWAVES 2.0, c) buoy 51003 *vs*. ECOWAVES and ECOWAVES 2.0, and d) buoy 51004 *vs*. ECOWAVES and ECOWAVES 2.0.

Table 7. Statistical parameters between ECOWAVES and ECOWAVES 2.0 *vs.* southeastern Pacific buoys. RMSE: root-mean-square-error, SI: scatter index, MAE: mean absolute error, BIAS: the average of the error.

Buoy name	SWH [m] ECOWAVES								SWH [m] ECOWAVES 2.0						
	Data	R	R^2	RMSE	SI	SS	MAE	BIAS	R	R^2	RMSE	SI	SS	MAE	BIAS
NDBC 32302	11440	0.91	0.83	0.26	0.12	0.88	0.20	0.08	0.93	0.86	0.24	0.11	0.89	0.19	0.08
NDBC 32012	29730	0.93	0.87	0.28	0.12	0.88	0.23	0.14	0.94	0.88	0.25	0.11	0.89	0.20	0.11
C-INNOVA 1	4726	0.88	0.77	0.40	0.17	0.84	0.30	-0.16	0.91	0.83	0.33	0.14	0.87	0.24	-0.06
C-INNOVA 2	1625	0.93	0.86	0.45	0.18	0.83	0.37	0.28	0.94	0.89	0.35	0.14	0.87	0.27	0.15

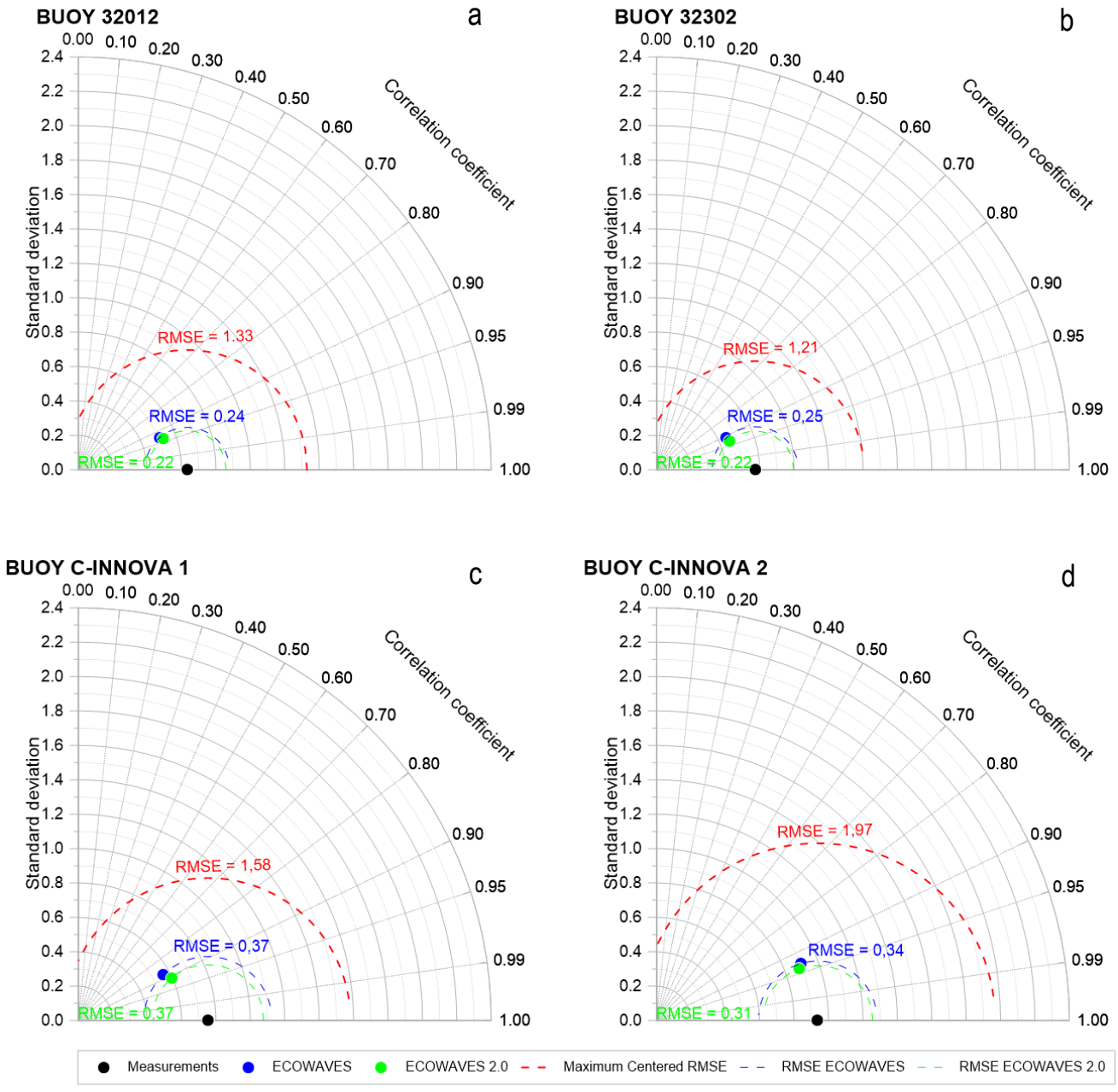

Figure 6. Comparison diagram of the statistical pattern between measured and simulated data for the distribution of NOAA and C-INNOVA buoys on the Chilean coastline. a) Buoy 32012 *vs.* ECOWAVES and ECOWAVES 2.0, b) buoy 32302 *vs.* ECOWAVES and ECOWAVES 2.0, c) buoy C-INNOVA1 *vs.* ECOWAVES and ECOWAVES 2.0, and d) buoy C-INNOVA2 *vs.* ECOWAVES and ECOWAVES 2.0.

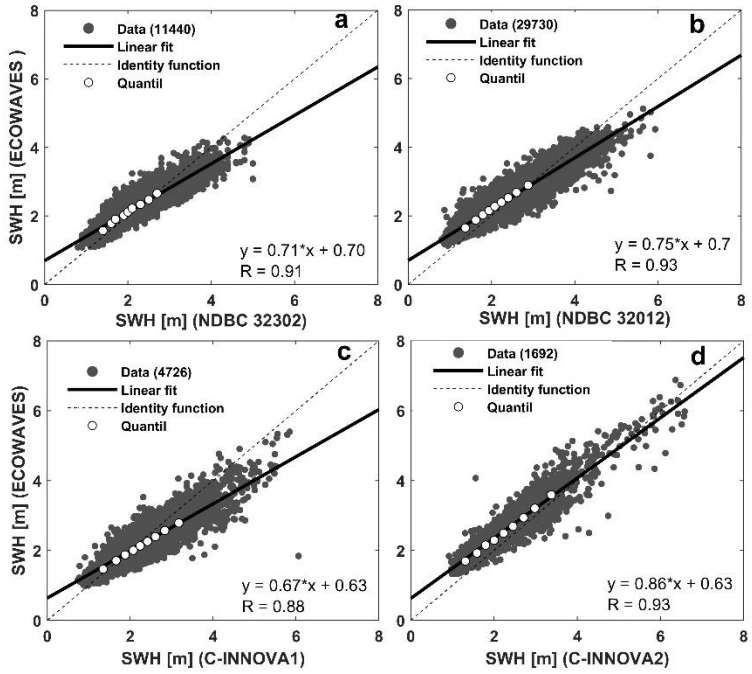

Figure 7. Scatter plots of buoys a) NDBC 32302 *vs.* ECOWAVES, b) NDBC 32012 *vs.* ECOWAVES, c) C-INNOVA 1 *vs.* ECOWAVES, and d) C-INNOVA2 *vs.* ECOWAVES. White points correspond to quantiles (from 0.05 to 0.85).

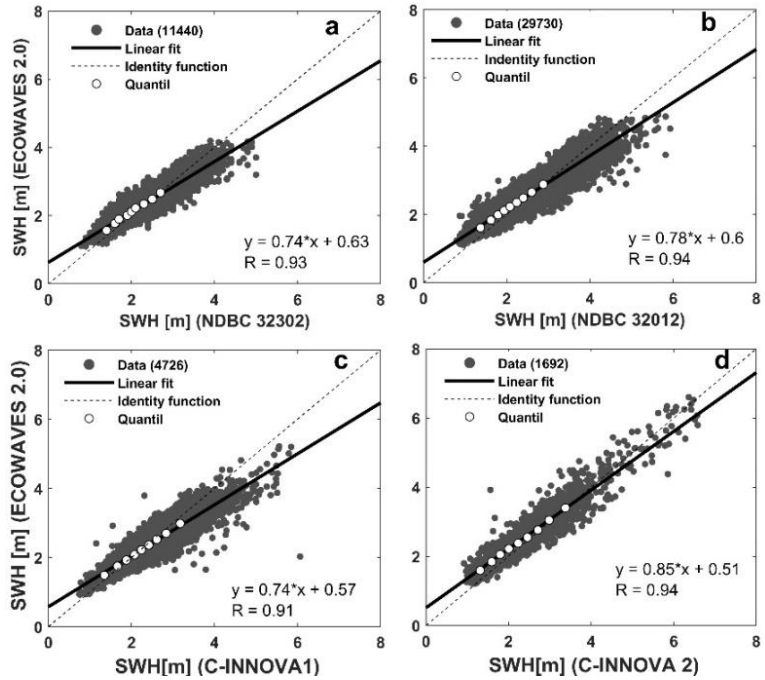

Figure 8. Scatter plots of buoys a) NDBC 32302 *vs.* ECOWAVES 2.0, b) NDBC 32012 *vs.* ECOWAVES 2.0, c) C-INNOVA 1 *vs.* ECOWAVES 2.0 and d) C-INNOVA2 *vs.* ECOWAVES 2.0. White points correspond to quantiles (from 0.05 to 0.85).

Shallow waters validation

A summary of the resulting statistics between couplings ECOWAVES+SWAN and ECOWAVES 2.0+SWAN *vs.* the ADCPs records is shown (Table 8). It is observed that the results from ECOWAVES simulations are strongly linearly correlated with the ADCP measurements, with magnitudes ranging from 0.75 to 0.95. Additionally, the R^2 indicates that, on average, the simulations manage to represent more than 75% of the measurements. The highest BIAS values are obtained in the comparisons with ADCP OFFSHORE in Iquique during summer and winter and Chañaral, with magnitudes of 0.28, 0.32, and 0.38 m, respectively. The rest of the BIAS is less than 0.16 m (in absolute value), which indicates a good agreement in the magnitudes of the wave heights. Additionally, the RMSE and MAE errors are less than 0.45 m, where the maximum differences occur during highly energetic sea states that the models cannot reproduce with high accuracy.

Then, it is evidenced that the results with ECOWAVES 2.0 show better statistical performance in most measurements. Increasing linear correlations demons-trate the performance increase from 0.77 to 0.96, a decrease in SI from 0.12 to 0.34, and a decrease in RMSE, MAE, and BIAS errors. Although the differences between the ECOWAVES 2.0 statistics in contrast to ECOWAVES are dimensionally small, the statistical agreement from data to data is significant (Table 8).

In general, it is observed that both wave *hindcasts* reproduce with great accuracy the wave height in shallow waters for different periods, different bays, and different latitudes. Nevertheless, ECOWAVES 2.0 is the best *hindcast* to reproduce the wave height in shallow waters.

Subsequently, in the case of the wave simulations forced by *forecast* winds (WW3 GFS), it was observed statistical representativeness in the temporal variability of wave heights but less accuracy than in the previous cases (Table 9), with an average BIAS of 0.5 m, significant reduction of R (0.67 to 0.96) and also increase in errors (RMSE and MAE).

Additionally, the performance of the average simulated SWH compared to the ADCPs records also may be observed in the SWH mean values of dispersion figures for all coastal stations (Fig. 9). Where it is evidenced that the mean SWH determined with wave *hindcast* is statistically strongly adjusted to the one to one line along Chile. Nonetheless, the better statistically adjusted *hindcast* (S.S. = 0.84) corresponds to the forced by ERA5 (Fig. 9b).

Also, in the case of simulations forced by *forecast* wind (Fig. 9c), it is observed that even though representativeness in the temporal variability data todata (R = 0.91) exists, there is a significant overestimation of (~0.5 m) and a decrease in the S.S.

In order to describe in more detail the comparison variability of the modeled wave against conducted measurements, the analysis in the far north, central, and south of the country was deepened, considering for such purpose, the records of Arica, Quintero, and Coronel, accordingly.

Arica

In this year's season, the most energetic waves on the Chilean coastline occur. Two measurement campaigns are available for this location. Results of the wave height analysis during winter are presented.

In general, it is observed that three simulations succeed in reproducing the temporary variability of wave heights in the measurement period. Nonetheless, those adjusted statistically better are forced by reanalysis winds (ECOWAVES and ECOWAVES 2.0). ECOWAVES has an R of 0.92, with RMSE, MAE, and BIAS of 0.26, 0.22, and 0.11 m, respectively (Figs. 10a,d), showing a robust statistical agreement.

Additionally, ECOWAVES 2.0 presents an increase of 0.94 in R and a drop in RMSE, MAE, and BIAS of 0.21, 0.15, and -0.06 m, accordingly (Figs. 10b,d). The wave model forced by wind fields from ERA5 presents a better statistical agreement to wave heights than forcing with ERA-Interim data. It is worth mentioning that, in this case, an average underestimation of 0.006 m in the ADCP measurements, in contrast to ECOWAVES, where an average overestimation equal to 0.11 m exists.

Besides that, when using *forecast* winds to force wave model, it is expected to decrease the percentage of convergence of the statistical agreement. Nonetheless, it is important to highlight that even when overestimations exist (BIAS = 0.8 m) in wave heightsand a decrease in the data-to-data correlation (R = 0.78), the model successfully reproduces the variability in the events as more energetic (Figs. 10c,d). Thus, the purpose of the extremes events *forecast* would correctly ascertain the occurrence prediction. Nevertheless, its ability to predict the energy (as a wave height function) is limited and tends to overestimate the wave height in each event.

The boxplots (Fig. 10e) show that ECOWAVES and ECOWAVES 2.0 successfully represent over 75% of the data measured by ADCP. Also, ECOWAVES best

Table 8. Summary of the statistical result between Acoustic Doppler Current Profiler (ADCPs) and wave *hindcast*. RMSE: root-mean-square-error, SI: scatter index, MAE: mean absolute error, BIAS: the average of the error.

ADCPs	SWH [m] ECOWAVES + SWAN							SWH [m] ECOWAVES 2.0 + SWAN					
	Data	R	R^2	RMSE	SI	MAE	BIAS	R	R^2	RMSE	SI	MAE	BIAS
Arica (winter)	276	0.92	0.84	0.26	0.17	0.22	0.11	0.94	0.89	0.21	0.14	0.15	-0.06
Arica (summer)	290	0.83	0.68	0.18	0.15	0.14	0.06	0.87	0.75	0.17	0.14	0.13	-0.07
OFFSHORE Iquique (summer)	341	0.95	0.90	0.30	0.26	0.28	0.28	0.96	0.93	0.20	0.17	0.18	0.17
PORT Iquique (summer)	341	0.93	0.87	0.15	0.31	0.14	0.14	0.92	0.84	0.14	0.29	0.12	0.12
OFFSHORE Iquique (winter)	274	0.91	0.83	0.39	0.26	0.34	0.32	0.93	0.87	0.29	0.19	0.25	0.22
PORT Iquique (winter)	274	0.87	0.76	0.14	0.30	0.11	0.07	0.85	0.72	0.13	0.26	0.11	0.02
Michilla	293	0.75	0.57	0.13	0.12	0.10	0.05	0.78	0.60	0.13	0.12	0.10	-0.05
Mejillones	286	0.75	0.57	0.05	0.21	0.04	-0.01	0.82	0.67	0.04	0.18	0.03	-0.02
Chañaral	251	0.89	0.80	0.45	0.33	0.39	0.38	0.92	0.84	0.40	0.30	0.34	0.33
Totoralillo	311	0.79	0.62	0.11	0.25	0.09	0.02	0.77	0.59	0.11	0.25	0.08	0
Huasco	279	0.88	0.78	0.05	0.11	0.04	-0.02	0.88	0.78	0.06	0.14	0.05	-0.04
Coquimbo	260	0.88	0.78	0.27	0.18	0.23	0.12	0.91	0.83	0.21	0.14	0.17	-0.02
Quintero	223	0.86	0.74	0.17	0.19	0.14	0.07	0.87	0.76	0.13	0.15	0.10	0.05
Valparaíso	167	0.95	0.90	0.28	0.26	0.21	-0.16	0.93	0.86	0.36	0.34	0.25	-0.20
San Antonio	264	0.93	0.87	0.17	0.13	0.13	0.05	0.95	0.90	0.15	0.11	0.12	0.04
Cahuil (spring)	343	0.90	0.82	0.29	0.16	0.22	-0.07	0.92	0.86	0.27	0.14	0.20	-0.08
Cahuil (winter)	193	0.85	0.73	0.36	0.17	0.30	-0.07	0.85	0.73	0.38	0.19	0.30	-0.17
Coronel	418	0.80	0.64	0.23	0.22	0.17	0.05	0.84	0.71	0.20	0.19	0.15	0.01

Table 9. Summary of the statistical results between Acoustic Doppler Current Profiler (ADCPs) and WW3 GFS. (-): indicates no GFS data available. RMSE: root-mean-square-error, SI: scatter index, MAE: mean absolute error, BIAS: the average of the error.

ADCPs	SWH [m] WW3 GFS + SWAN						
	Data	R	R^2	RMSE	SI	MAE	BIAS
Arica (winter)	276	0.81	0.65	0.89	0.60	0.82	0.81
Arica (summer)	290	0.69	0.48	0.49	0.41	0.46	0.46
OFFSHORE Iquique (summer)	341	0.92	0.83	0.71	0.62	0.69	0.69
PORT Iquique (summer)	341	0.79	0.63	0.58	1.22	0.55	0.55
OFFSHORE Iquique (winter)	-	-	-	-	-	-	-
PORT Iquique (winter)	-	-	-	-	-	-	-
Michilla	293	0.71	0.50	0.50	0.47	0.47	0.47
Mejillones	286	0.74	0.54	0.13	0.60	0.12	0.12
Chañaral	251	0.81	0.65	1.26	0.94	1.14	1.14
Totoralillo	252	0.67	0.44	0.26	0.62	0.23	0.23
Huasco	-	-	-	-	-	-	-
Coquimbo	260	0.76	0.57	0.99	0.67	0.91	0.91
Quintero	223	0.71	0.50	0.51	0.55	0.45	0.45
Valparaíso	167	0.96	0.93	0.31	0.29	0.23	-0.22
San Antonio	-	-	-	-	-	-	-
Cahuil (spring)	-	-	-	-	-	-	-
Cahuil (winter)	193	0.82	0.66	0.92	0.45	0.75	0.75
Coronel	-	-	-	-	-	-	-

represents some of the most energetic events but underestimates those less energetic ones. While ECOWAVES 2.0 presents a median similar to the measurements and better agreement with less energetic events. In contrast, simulations with WW3 GFS show overestimation in the medians and percentiles.

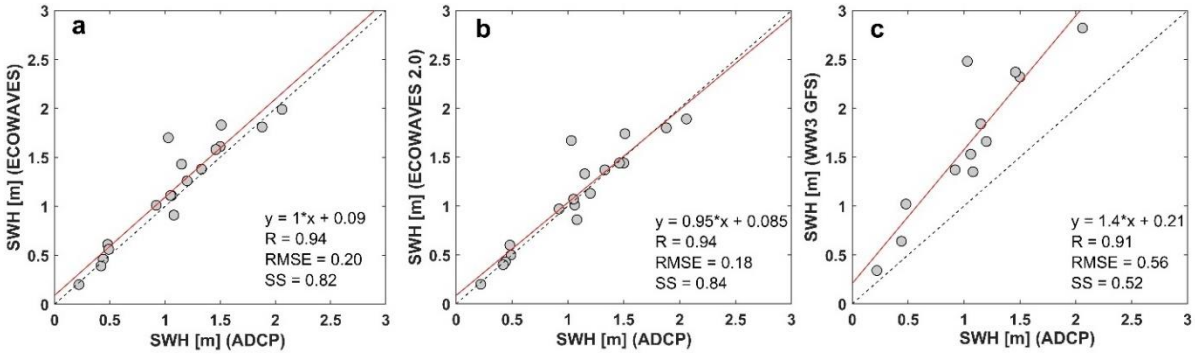

Figure 9. Scatter plots of simulated mean significant heights *vs*. measured. a) ECOWAVES + SWAN *vs*. ADCPs, b) ECOWAVES 2.0 + SWAN *vs*. ADCPs, and c) WW3 GFS + SWAN *vs*. ADCPs. R: correlation, RMSE: root mean square error, SS: skill score.

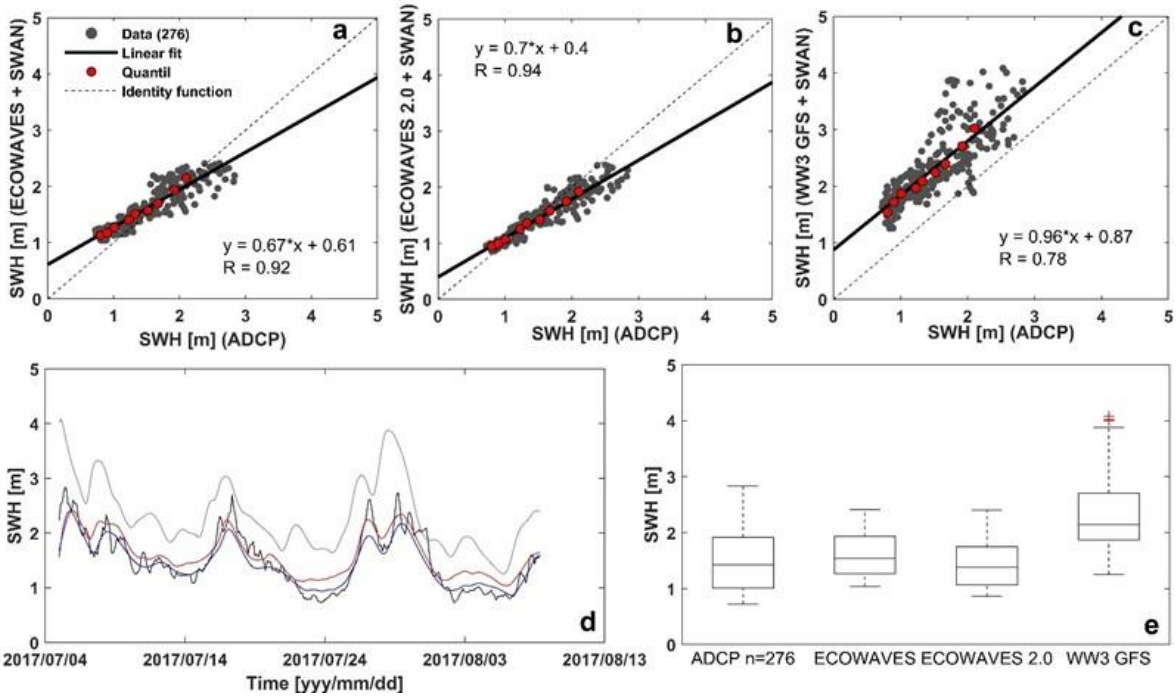

Figure 10. Significant wave height (SWH) scatter plots for a) ECOWAVES *vs.* Acoustic Doppler Current Profiler (ADCP), b) ECOWAVES 2.0 *vs*. ADCP, and c) WW3 GFS *vs*. ADCP; d) time series, in red corresponds to ECOWAVES, blue to ECOWAVES 2.0, grey to WW3 GFS and black to ADCP, e) boxplots.

Quintero

It is observed that the three simulations successfully represent the temporal distribution of the wave heights that are present in shallow waters (Fig. 11d). More specifically, coupling with ECOWAVES presents good data to data performance (R = 0.86). However, it shows an overestimation with BIAS equal to 0.09 m (Fig. 11a). RMSE and MAE errors are 0.16 and 0.15 m, respectively, indicating good performance in shallow waters.

Simulations resulting from coupling with ECOWAVES 2.0 present an improvement regarding ECOWAVES, increasing the R slightly up to 0.87 but decreasing BIAS to 0.05, RMSE to 0.14 m, and MAE to 0.10 m (Fig. 11b). Nevertheless, with ECOWAVES 2.0 exists less representativeness in some specific more energetic events compared to ECOWAVES (Fig. 11d).

Also, results with WW3 GFS successfully represented the wave height variability. Nevertheless, the entire time series overestimates (BIAS = 0.45 m) (Figs. 11c,d). The overestimation is an average of 0.4 m higher

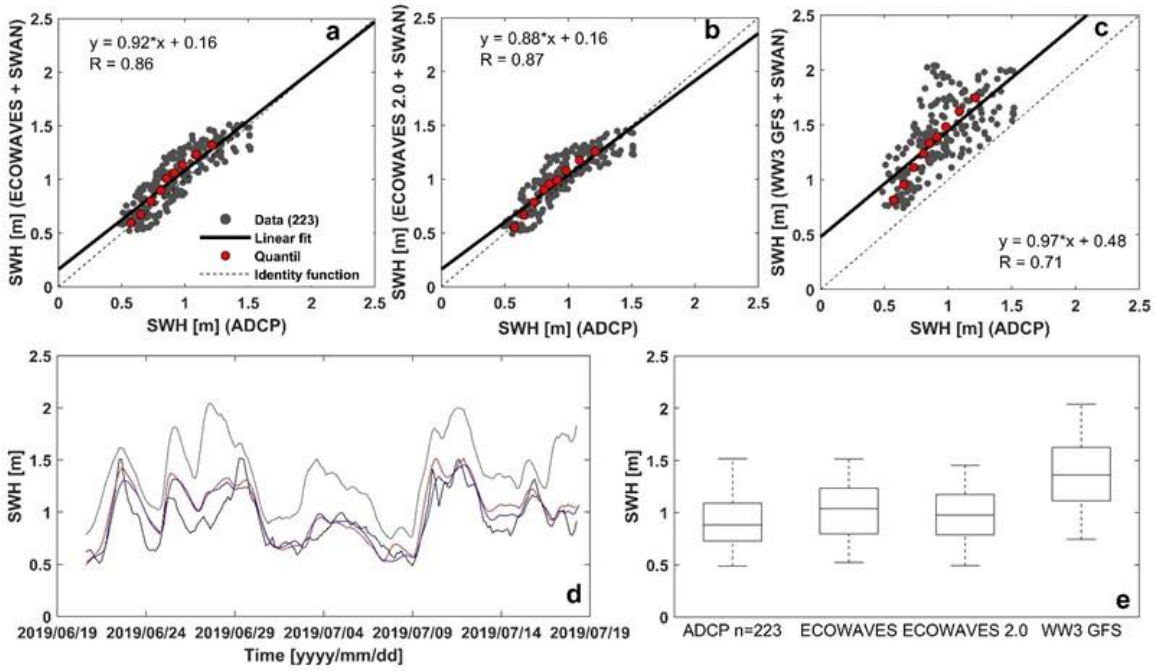

Figure 11. Significant wave height (SWH) scatter plots for a) ECOWAVES *vs.* Acoustic Doppler Current Profiler (ADCP), b) ECOWAVES 2.0 *vs*. ADCP, and c) WW3 GFS *vs*. ADCP; d) time series, in red corresponds to ECOWAVES, blue color to ECOWAVES 2.0, grey color to WW3 GFS and black to ADCP, and e) boxplot.

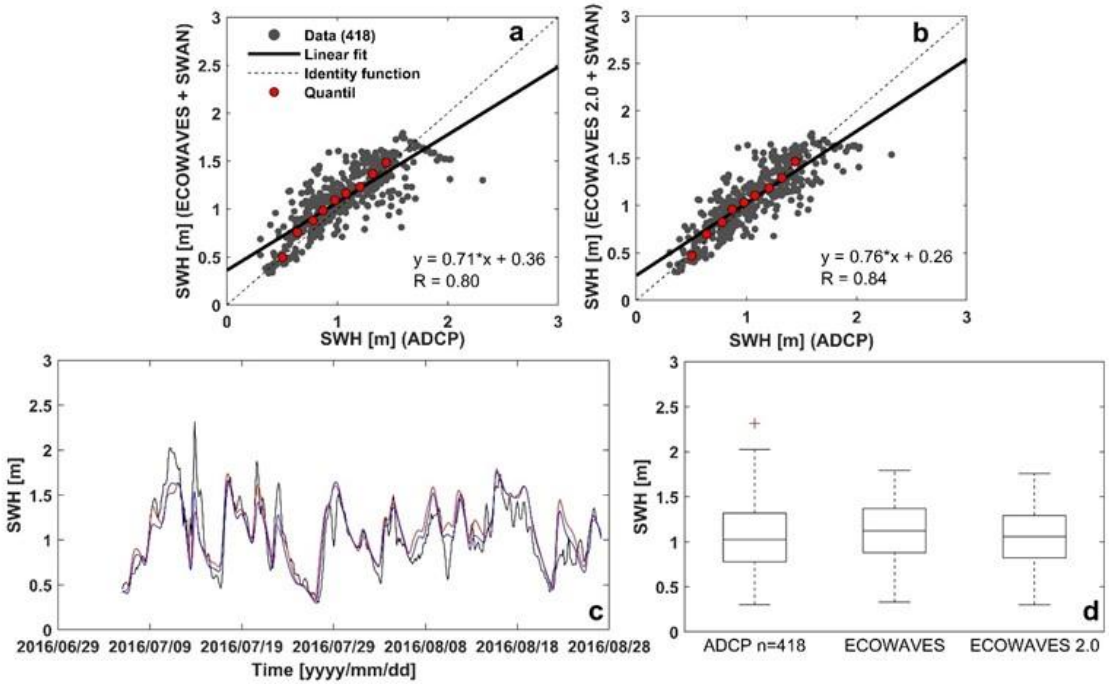

Figure 12. Scatter plots for a) ECOWAVES *vs*. Acoustic Doppler Current Profiler (ADCP), b) ECOWAVES 2.0 *vs.* ADCP, c) time series, in red corresponds to ECOWAVES, blue to ECOWAVES 2.0 and black to ADCP; and d) boxplot.

than in the wave *hindcast*. It is important to highlight that there exist sea states which are not properly represented (Fig. 11d), which involves a decrease in correlation to 0.71, and an increase of RMSE and MAE to 0.51 and 0.45, respectively.

It is shown that with ECOWAVES, better representativeness of some of the most energetic sea states is achieved. With ECOWAVES 2.0, median and percentiles of 25-75% are more similar to measured magnitudes (Fig. 11e). For WW3 GFS, it is evidenced an overestimation in all statistical parameters (Fig. 11e).

Coronel

Results show that strong lineal relations exist in both simulations, with R magnitudes oscillating between 0.80 and 0.84 for ECOWAVES and ECOWAVES 2.0 (Figs. 12a,b).

Regarding media errors, ECOWAVES 2.0 presents BIAS equal to 0.01 m. In contrast, ECOWAVES present BIAS equal to 0.05 m. Lower RMSE and MAE are present in ECOWAVES 2.0 with a difference of two decimals comparing the same rates in ECOWAVES (Table 8).

Additionally, it is evidenced that both simulations successfully represent the wave height variability (Fig. 12d). It is observed that medians, percentiles 25-75%, and minimal values of simulations adjust better with ECOWAVES 2.0. In contrast, ECOWAVES provides better representativeness of some of the most energetic events (Figs. 12c,d).

It is worth mentioning that, as indicated in Table 9, no GFS wind data records for the ADCP measured period were available in this analysis. Therefore, results for such modeling were not included in the comparison.

DISCUSSION

This work presents the validation of significant heights in deep and shallow waters resulting from couplings between the WW3 model forced by different surface winds (ERA-Interim, ERA5, and GFS) and the SWAN model. The parameterization configuration and calibration used for the WW3 model were the same for the three forcings. Therefore, wave simulations depended exclusively on the quality of the entered surface wind forcings. The validation process in deep waters of the ECOWAVES and ECOWAVES 2.0 databases showed satisfactory performance in representing wave heights in both the North and South Pacific. However, among them, the best corresponds to ECOWAVES 2.0, which presented the highest correlation indices, lowest standard deviation, and lowest errors (BIAS, RMSE). If the results obtained from ECOWAVES and ECOWAVES 2.0 are compared with the other wave *hindcasts* elaborated in Chile (Fig. 13), it can be seen that both wave *hindcasts* present a better correlation for the buoys located in the North Pacific compared to the results of the Olas del Pacifico Project (http://www. olasdelpacifico.com/validacion.html).

Additionally, for the NOAA 32012 buoy (in deep waters located in the north of Chile), the best correlation was obtained with ECOWAVES 2.0 (R = 0.94) exceeding those values obtained by Beyá et al. (2017) and Aguirre et al. (2017) (Fig. 13). Furthermore, if C-INNOVA 1 and C-INNOVA 2 correlation coefficients obtained by Beyá et al. (2017) are compared to the *hindcasts* presented in this document, they are dimensionally similar with ECOWAVES.

However, with ECOWAVES 2.0, a significant increase in correlation was observed, showing the improvement of the wave *hindcast* to represent the wave heights in Chile.

It was observed that the hindcast type simulations have a correct assertiveness in the occurrence and magnitude of the events, unlike what happens with the forecast type simulations (Table 9), where a correct assertiveness of the occurrence is observed. However, its ability to forecast the energy is restricted and tends to overestimate the wave heights in each event. Additionally, it was shown that in shallow waters, the simulations managed to reproduce the measured wave heights (Table 8) but with different statistical performances.

Meanwhile, the simulations of forced waves with GFS winds presented the worst statistical performance in comparing the measurements. It was possible to show that the overestimation of the speeds of the surface wind fields generates overestimations of the magnitudes of the wave heights. Therefore, to improve the representativeness of the wave heights, it is necessary to perform a previous adjustment of the surface wind fields.

In particular, for the localities of Arica, Quintero, and Coronel, it was shown that coastal simulations manage to reproduce the variability of wave heights in the north, center, and south of Chile. A high correlation (R > 0.8) was observed in the occurrence of the events, but the best fit in the magnitudes of the events (BIAS < 6 cm) was obtained with ECOWAVES 2.0.

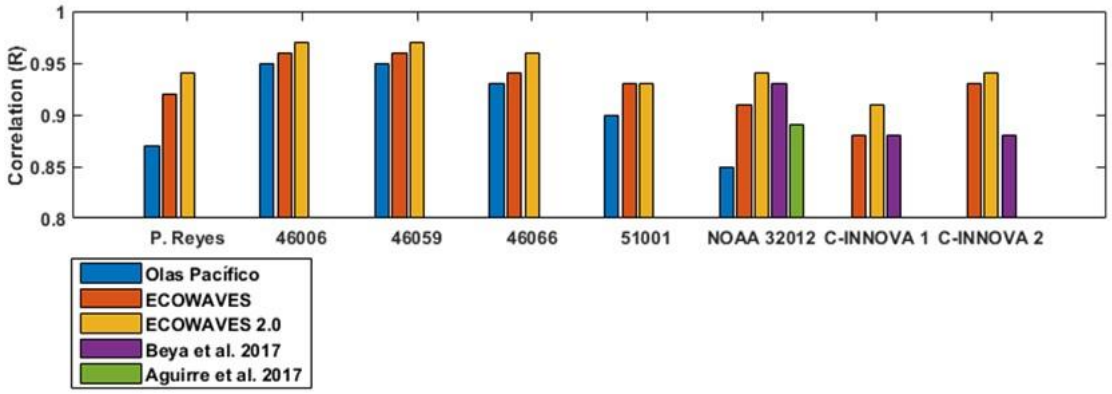

Figure 13. Comparison of linear correlations between different wave *hindcasts* elaborated in Chile.

It is important to highlight that, despite the technical limitations that the SWAN model has to represent coastal processes (for example, reflection and diffraction), together with the different configurations that the ADCPs had and the geographical distribution where they were installed, in all cases, the results obtained are a first step to delve into future research.

CONCLUSION

It is concluded that to build a wave *hindcast* in shallow waters, the wind fields from ERA5 best statistically reproduce wave conditions on the coast of Chile. Additionally, it is concluded that ECOWAVES 2.0 corresponds to a good quality database for research studies and coastal, maritime, and port engineering projects.

ACKNOWLEDGMENTS

We thank the co-workers of the technical team in the Department of Physical Oceanography and Mathematical Modeling of EcoTecnos S.A. for the support and constructive criticism during this work.

REFERENCES

Aguirre, C., Rutllant, J. & Falvey, M. 2017. Wind waves climatology of the southeast Pacific Ocean. International Journal of Climatology, 37: 4288-4301.

Ardhuin, F., Reilly, W., Herbres, T. & Jessen, P. 2003. Swell transformation across the continental shelf. Part I: attenuation and directional broadening. Journal of Physical Oceanography, 33: 1921-1939.

Ardhuin, F., Rogers, E., Babanin, A. & Filipot, J.F. 2010. Semiempirical dissipation source functions for ocean wave. Part I: definition, calibration, and validation. Journal of Physical Oceanography, 40: 1917-1941.

Ardhuin, F., Hanafin, J., Quilfen, Y., Chapron, B., Quwffeulou, P. & Obrebski, M. 2011. Calibration of the "IOWAGA" global wave *hindcast* (1991-2011) using ECMWF and CFRS winds. 12th International Workshop in Wave Hindcasting and Forecasting, Kohala Coast, Hawaii, pp. 1-13.

Arthur, R.S. 1947. Revised wave forecasting graphs and procedure. Wave Report 73. Scripps Institution of Oceanography, San Diego.

Battjes, J. & Janssen, J.P.F.M. 1978. Energy loss and setup due to breaking of random waves. Proceedings of 16th Conference on Coastal Engineering, Hamburg, Germany, pp. 569-587.

Belmonte, M. & Stoffelen, A. 2019. Characterizing ERA-interim and ERA5 surface wind biases using ASCAT. Ocean Science, 15: 831-852.

Beyá, J., Álvarez, M., Gallardo, A., Hidalgo, H. & Winckler, P. 2017. Generation and validation of the Chilean Wave Atlas database. Ocean Modeling, 116: 16-32.

Booij, N., Ris, R. & Holthuijsen, L. 1999. A third-generation wave model for coastal regions 1. Model description and validation. Journal of Geophysical Research, 104: 7649-7666.

Bretschneider, C.L. 1951. Revised wave forecasting relationships. Coastal Engineering Proceedings, 2: 1-5.

Cavaleri, L., Fox-Kemper, B. & Hemer, M. 2012. Wind waves in the coupled climate system. Bulletin of the American Meteorological Society, 93: 1651-1661.

Chawla, A. & Tolman, H. 2013. Automated grid generation for wavewatch III. MMAB contribution, NCEP, NOAA, Washington DC.

Corporación de Fomento de la Producción-INNOVA (CORFO-INNOVA). 2009. Catastro del recurso energético asociado a oleaje para el apoyo a la evaluación de proyectos de generación de energía undimotriz. Proyecto CORFO-INNOVA 09CN14-5718 Hydrochile, INH, PUC, PRDW, UV, UTFSM, Valparaíso.

Davis, R. & More, E. 1982. A numerical study of vortex shedding from rectangles. Journal of Fluid Mechanics, 116: 475-506.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, U., et al. 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137: 553-597.

Fan, Y., Ginis, I. & Hara, T. 2009. The effect of wind-wave-current interaction on air-sea momentum fluxes and ocean response in hurricanes. Journal of Physical Oceanography, 39: 1019-1034.

Fournier, C., Pantoja, C., Resio, D. & Scott, D. 2004. The development of a 40 year wave climate for the entire Chilean coastline. American Society of Civil Engineers, Reston.

Gelaro, R., McCarty, W., Suarez, M., Todling, R., Molod, A., Takacs, L., et al. 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). Journal of Climate, 30: 5419-5454. doi: 10.1175/JCLI-D-16-0758.1

Goda, Y. 1988. On the methodology of selecting design wave heights. Coastal Engineering, 21: 899-913.

Group, T.W. 1988. The WAM model-A third generation ocean wave prediction model. Journal of Physical Oceanography, 18: 1775-1810.

Hasselmann, S., Hasselmann, K., Allender, J. & Barnett, T. 1985. Computations and parametrizations of the nonlinear energy transfer in a gravity-wave spectrum. Part II: parametrizations of the nonlinear energy transfer for applications in wave models. Journal of Physical Oceanography, 15: 1378-1391.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Sabater, M., Nicolas, J., et al. 2020. The ERA5 global reanalysis. Journal of the Royal Meteorological Society, 146: 1999-2049.

Holthuijsen, L., Booij, N., Ris, R., Haagsma, I., Kieftenburg, A., Kriezi, E. & Zijlema, M. 2003. SWAN Cycle-III version 40.20 User Manual. Delft University of Technology, Delft.

Janssen, P. 1989. Wave-induced stress and the drag of air flow over sea waves. Journal of Physical Oceanography, 19: 749-754.

Jiang, H., Babanin, A. & Chen, G. 2016. Event-based validation of swell arrival time. Journal of Physical Oceanography, 46: 3563-3569.

Komar, P. 1997. Beach processes and sedimentation. Prentice-Hall, New Jersey.

Leonard, B. 1975. A stable and accurate convective modeling procedure based on quadratic upstream interpolation. Computational Method Applied Mechanics and Engineering, 19: 59-98.

Ramon, J., Lledó, L., Torralba, V., Soret, A. & Doblas-Reyes, F. 2019. What global reanalysis best represents near-surface winds? Quarterly Journal of the Royal Meteorological Society, 145: 3236-3251.

Resio, D. & Perrie, W. 1989. Implications of an f-4 equilibrium range for wind-generated waves. Journal of Physical Oceanography, 19: 193-204.

Rogers, E., Kaihatu, J., Hsu, L., Jensen, R., Dykes, J. & Holland, T. 2006. Forecasting and hindcasting waves with the SWAN model in the Southern California Bight. International Journal for Coastal, Harbour and Offshore Engineers, 54: 1-15.

Sang-Ho, O., Suh, K., Son, S. & Lee, D. 2009. Performance comparison of spectral wave models based on different governing equations including wave breaking. Journal of Civil Engineering, 13: 75-84. doi: 10.1007/s12205-009-0075-y

Sverdrup, H.U. & Munk, W.H. 1947. Wind, sea and swell: theory of relations for forecasting. United States Navy Department, Washington D.C.

SWAN Team. 2019. SWAN user manual. The Delft University of Technology, Delft.

Taylor, E. 2001. Summarizing multiple aspects of model performance in a single diagram. Journal Geophysical Research, 106: 7183-7192. doi: 10.1029/2000JD900719

Tolman, H.L. 1991. A third-generation model for wind waves on slowly varying, unsteady, and inhomogeneous depths and currents. Journal of Physical Oceanography, 21: 782-797.

Tolman, H.L. 2002. Alleviating the garden sprinkler effect in wind wave models. Ocean Modeling, 4: 269-289.

Tolman, H.L. 2019. User manual and system documentation of WAVEWATCH III version 6.07. Environmental Modeling Center, Marine Modeling and Analysis Branch, Maryland.

Williams, J.J. & Esteves, L.S. 2017. Guidance on setup, calibration, and validation of hydrodynamic, wave, and sediment models for shelf seas and estuaries. Advances in Civil Engineering, 2017: 5251902.

Winckler, P., Contreras-López, M., Campos-Caba, R., Beyá, J. & Molina, M. 2017. El temporal del 8 de agosto de 2015 en las regiones de Valparaíso y Coquimbo, Chile Central. Latin American Journal of Aquatic Research, 45: 622-648. doi: 10.3856/vol45-issue4-fulltext-1

Received: September 9, 2021; Accepted: August 1, 2022