Development of an Automated Real-Time Intelligent Information System for Early Warning and Preparedness of Offshore Oil and Gas Operations off the Coast of Israel

DARTIS

Israel Ministry of Science & Technology

Bundesministerium für Bildung und Forschung

DECEMBER 31, 2022

Israel Oceanographic and Limnological Research (IOLR). Research & Technology Centre, University of Kiel (FTZ). Aerospace Centre (DLR). Sea & Sun Technology GmbH (SST)

Table of Contents

List of Figures

List of tables

Abbreviations.

IOLR: Israel Oceanographic and Limnological Research.
FTZ: Research & Technology Centre, University of Kiel.
DLR: German Aerospace Centre.
SST: Sea & Sun Technology GmbH.
DARTIS: Development of an Automated Real-Time Intelligent Information System
BMBF: Federal Ministry of Education and Research of Germany

Summary

1. Scientific goals

DARTIS stands for Development of an Automated Real-Time Intelligent Information System for Early Warning and Preparedness of Offshore Oil and Gas Operations off the Coast of Israel. This project aimed to improve the development of an automated, real-time intelligent information system for supporting oil and gas operations off the coast of Israel. For this, we developed several strategies to improve wave prediction in terms of accuracy and reduction of the computational time to produce faster results and reduce the response time in hazard situations.

At the Israel Oceanographic and Limnological Research (IOLR), an existing operational model for wave predictions based on WAM and SWAN existed. Also, these wave forecasts are integrated with an existing oil spill model

To achieve the objectives, we added several methods based on data assimilation of wave measurements and neural networks.

The objectives are:

- Improve the accuracy of the existing operational wave model system.
- Develop strategies to Improve the speed of providing wave forecasts using artificial neural networks.
- Integrate the developed strategy into the existing IOLR operational system.

2. Partners

The scientific partners are:

Israel Oceanographic and Limnological Research (IOLR): Dr. Isaac Gertman., Mr. Ron Goldman, Dr. Elad Dakar.

German Aerospace Centre (DLR): Dr. rer. nat. S. Jacobsen. S. Singha Ph.D. , Msc. Y. Yang.

Sea & Sun Technology GbmH (SST): Prof. H. Schelwat.

Research & Technology Centre, University of Kiel (FTZ): Prof. Dr. R. Mayerle, Dr. Rer. Nat. J.M. Fernández Jaramillo, Msc. D. Restrepo Álvarez.

3. Achieved results

Ensemble Kalman Filter (EnKf) and 4Dvar data assimilation have been implemented for the wave height hindcast and prediction for the coastal areas of Israel. Both methods have improved the results obtained by the existing standalone model prediction system. It was developed using the wave model SWAN, and we implemented the algorithms in an external Python code that automates the processes and the optimization algorithms.

We implemented several strategies to use neural networks to predict wave height. These strategies aimed to improve the quality, the speed up, or both for the wave forecasts. These include a hybrid model that uses traditional numerical models for wave forecast and meteorology, with observations

to enhance prediction. Also, a neural network model directly uses meteorological forecasts to produce spatial wave forecasts and can produce faster results than traditional models. Also, as a by-product, we obtain an autoencoder to compress meteorological and wave models. This autoencoder brings files up to 25 times smaller without a significant quality loss.

As a result of the 4DVar assimilation, it was possible to find a multiplication coefficient variable in the time to adjust the winds from the atmospheric model to improve the wave forecast. A neural network model was developed to estimate such coefficient in advance and provide an improved wind field to SWAN. This model can improve the spatial prognosis without incurring the high computational effort of the data assimilation.

All the developed strategies have been implemented to work in the operational system at IOLR. They are available online at https://www.ocean.org.il/.

1. Introduction

Accurately predicting ocean waves is essential for marine activities, such as shipping, fishing, offshore operations, and recreational water sports. Wave forecasting has traditionally relied on numerical models that use mathematical equations to simulate wave behavior. However, these models must be revised to accurately capture wave height due to their complex and nonlinear nature, especially in regions with rapidly changing coastal bathymetry and ocean currents.

Recent advancements in data assimilation and artificial neural networks (ANNs) offer promising solutions to improve wave forecasting. Data assimilation is a technique that combines observational data with numerical models to obtain more accurate predictions. ANNs are powerful machine learning algorithms that can learn the complex relationships between input data and output predictions, making them well-suited for wave forecasting.

This scientific report will present several methods for improving wave forecasting by integrating data assimilation and ANNs. We implemented different algorithms to produce near real-time wave forecasting with the existing model and the new proposed techniques with data assimilation and ANNs to improve current results. We will also present some results demonstrating the effectiveness of these in improving wave forecasting accuracy and performance.

2. Investigation area.

This project was developed in the Levantine subregion of the Eastern Mediterranean (see Figure 1), focusing on the zone of Israel. This region has diverse landscapes and a Mediterranean climate, with hot, dry summers and mild, wet winters. Large reserves of oil and gas have been discovered in recent years in the Levant Basin, which stretches from the coast of Israel northward to Cyprus and westward to the coast of Lebanon. Also, this region of the Mediterranean Sea is a major shipping route connecting Europe, the Middle East, and North Africa. The easy access to these significant routes allows Israel to become an essential hub for international commerce. It is crucial to assess the risks originated from waves in this region due to the strategic significance for the economy of Israel of oil and gas as natural resources and the dangers of oil spills generated during the exploration and extraction of oil and gas, as well from the transit of cargo ships due to the vicinity of the Suez channel and other ports in the area.

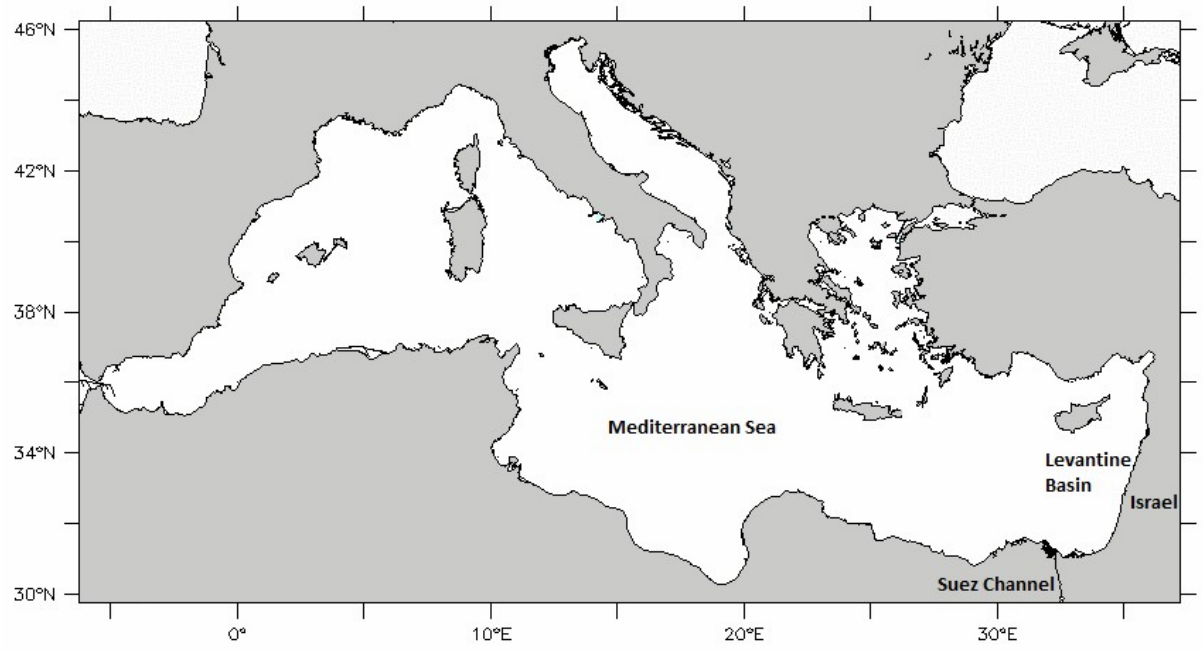

Figure 1 Mediterranean Sea and the Levantine basin.

3. Available data and models

3.1. Meteorological and wave models

This system delivers quasi-real-time continuous wave predictions in the Easter Mediterranean Sea (Gertman, Murashkovsky, Levin, Kallos, & Rosen, 2006). The forecasts are built upon Simulating Waves Nearshore (SWAN) model (Booij, Ris, & Holthuijsen, 1999) (see Fig. 2) based on previous research using the WAM-Cycle 4 (Gertman, Rosen, Kariel, & Raskin, 2000), which is forced by winds from the Skiron model (Kallos, et al., 1997), and a fine-gridded model for the eastern Mediterranean Sea is nested in a coarse-gridded model for the entire Mediterranean Sea (Gertman, Murashkovsky, Levin, Kallos, & Rosen, 2006). The Skiron meteorological model is a regional weather forecast model for the Mediterranean Sea. The Skiron model is run daily at the University of Athens and produces Grib files with a 0.1 degree resolution and a forecast period of 5 days (https://openskiron.org/en)

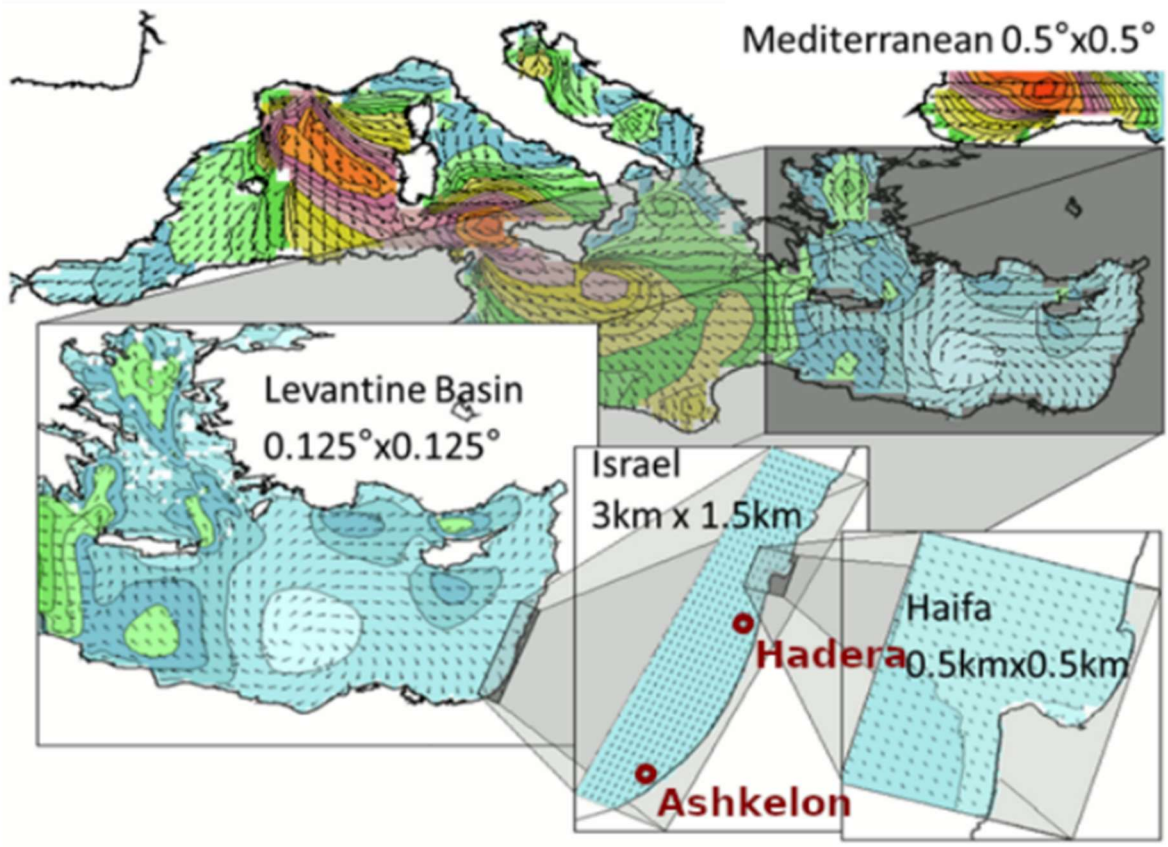

Figure 2. Existing models at IOLR.

3.2. Measurement stations.

Data from up to six measurement stations (see Fig 3.) have been available for assimilation; two of them in Israel (Hadera and Ashkelon) were already available at IOLR, and four additional in Greece operated by the Hellenic Centre for Marine Research and provided by COPERNICUS (www.copernicus.eu) (Copernicus, doi:10.48670/moi-00044). The stations in Greece are located at Athos in the Aegean Sea, Saron in the Saronic Gulf, and two North of the island of Crete (Heraklion and 61277) in the Sea of Crete.

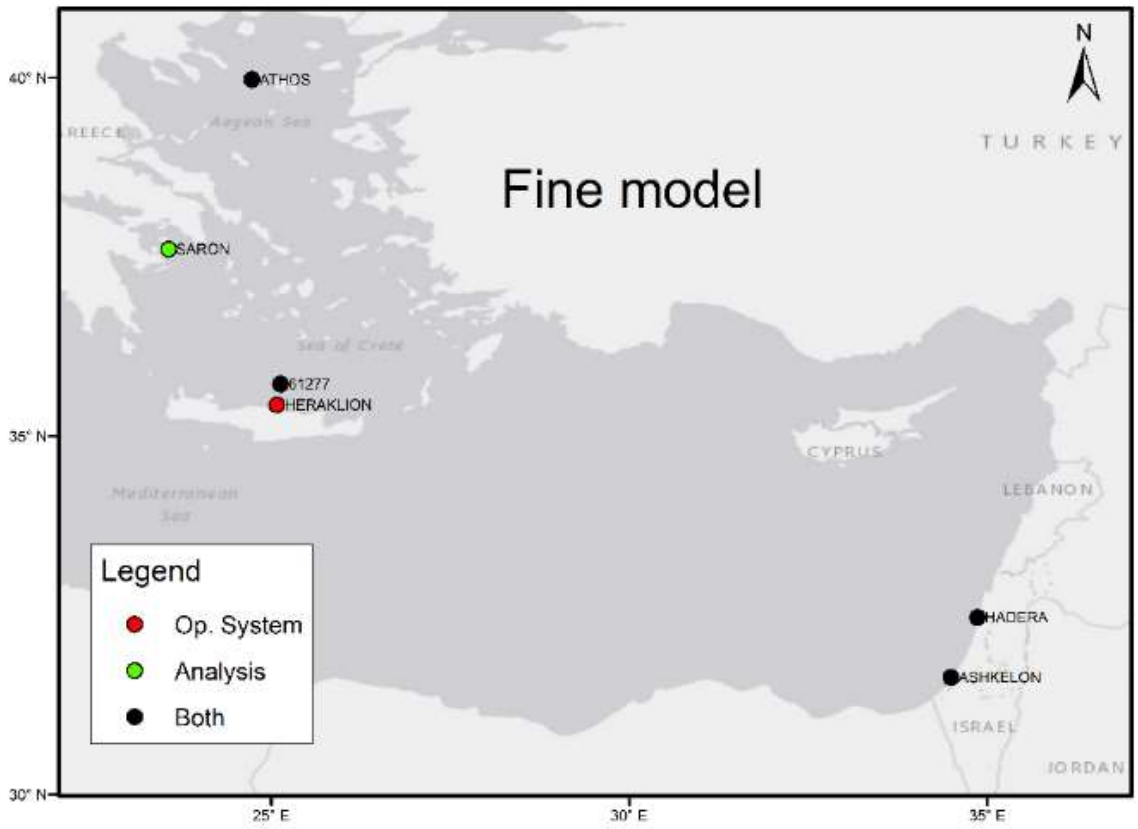

Figure 3. Available measurement stations for wave heights in the area of the Levantine model.

3.3. Bathymetry

Bathymetrical information has been obtained from measurements carried out by IOLR and the global terrain model "General Bathymetric Chart of the Oceans" (GEBCO www.gebco.net). GEBCO is a global bathymetric dataset that provides information about the topography and shape of the seafloor (Weatherall et. al. 2015, Jacobsson et. al. 2012). This dataset is a compilation of ocean depth measurements taken by various sources, including ships, satellites, and airborne surveys. The bathymetry is provided in an interpolated grid with a resolution of 15 arc seconds (around 500 m). Figure 4 shows the terrain model from GEBCO for the region of the Mediterranean Sea.

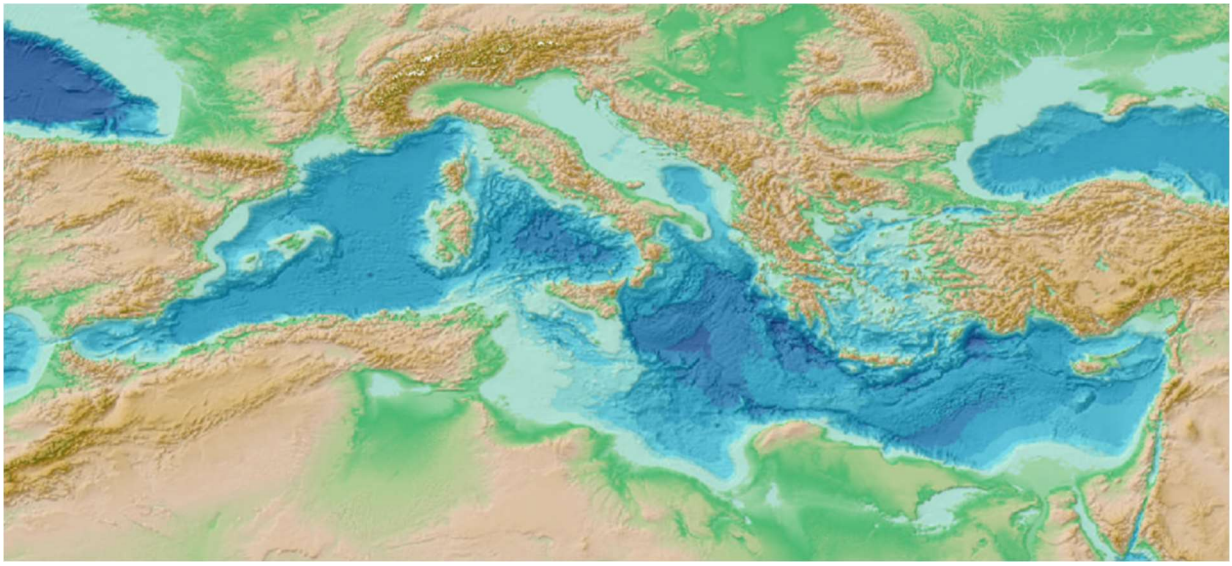

Figure 4 GEBCO terrain model for the region of the Mediterranean Sea.

4. Data Assimilation

In this module, four main activities have been proposed: a) identification of the available wave data, b) validation of the existing wave models, c) implementation of the EnKF & 4Dvar schemes, and d) assessment of the performance of the various scale of the model (EMS & Coastal Model). The identification of data and validation of existing models (Mediterranean, Levantine Basin, and Israel domains, see Fig 2) has been fully accomplished. The two data assimilation schemes have been developed by adjusting the boundary conditions and the wind field applied to the domain.

4.1. Implementation of the assimilation schemes

Initially, a 4DVar scheme was implemented directly within the SWAN source code, with good results in the smallest domain where the influence of boundary conditions is more relevant than in the other domains. We faced problems with this implementation and could not significantly improve the wave predictions only with the wind field. This forced us to discard this approach and restart the activity by implementing an external code that prepares the inputs for the simulations, calculates the waves with SWAN, and then performs the optimization steps. With this strategy, an EnKF scheme was developed (Restrepo et. al., unpublished) using an external script in Python. This implementation was much more effective in improving the predictions throughout the three model domains by tuning both wind and boundary conditions. Besides, the scheme enabled us to capture the peaks of the storm in better agreement with the observations. Due to the problems faced in implementing the 4DVar within the SWAN code, a new implementation has been done using external scripts. It proved to be able to improve the results by changing wind conditions. The results of reconstructing the storms from this new implementation of the 4DVar have errors in the same order of magnitude as those obtained from the EnKF assimilation.

4.2. Hindcast results for EnKF and 4DVar.

Figure 5. presents the results of the two assimilations implemented with Python for the Hindcast as time series, here, it is possible to observe how the standard SWAN model tends to have much higher peaks of significative wave height, while assimilation results represent much better observations. Figure 6 presents the scatter plot comparing each method with the observations; in this figure, a perfect fit is represented by a line of 45 degrees. For the standard wave model that was in use before the start of the project, we can observe that it provides a consistent overprediction of the wave height. On the other hand, both assimilation techniques developed in this project have adjusted the observation and modeled outputs to a closer value. Table 1 also presents some statistics for the tested methods. Both methods can improve the model results at the measurement station.

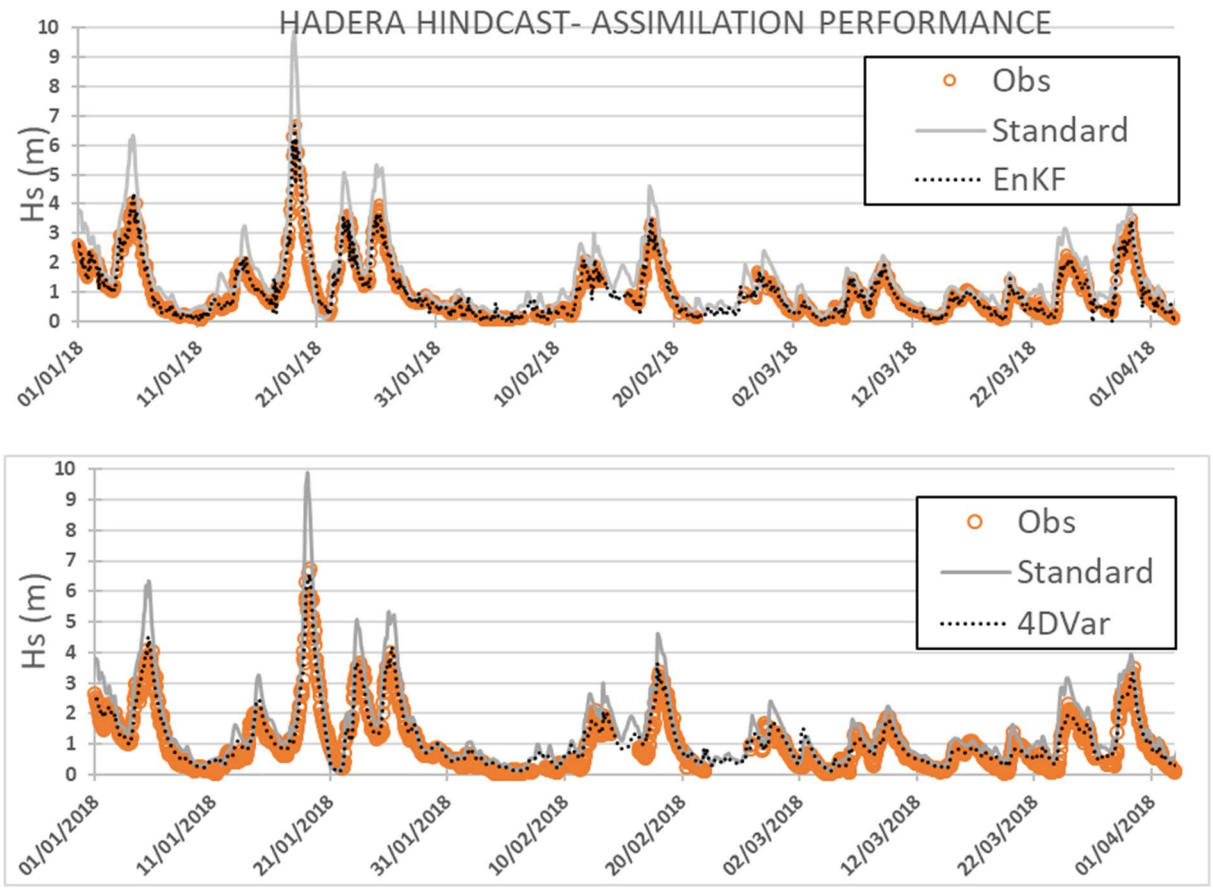

Figure 5. a) compares the standard SWAN model without assimilation, observations, and EnKF data assimilation at Hadera. b) the same comparison, but it includes the 4DVar data assimilation.

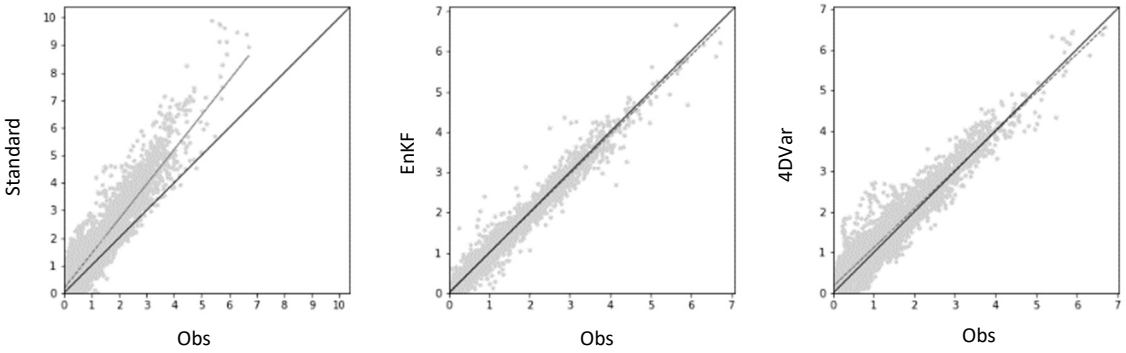

Figure 6. Scatter plot comparing the significant wave height observations (in meters) at Hadera with the outputs of a) standard SWAN, b) EnKF assimilation, and c) 4DVar assimilation.

Table 1. statistics for the SWAN model without assimilation and the two assimilation schemes.

Run type	Bias(m)	R	MAE (m)	RMSE (m)
SWAN	0.364	0.961	0.404	0.577
EnKF	0.000	0.998	0.061	0.094
4DVar	0.122	0.984	0.197	0.294

4.3. Forecast results

Both schemes have been tested for the forecast to improve the quality of the predictions. The EnKf has been able to show improvements only for a short period, while the 4DVar was able to improve results even one day in advance. Figure 7 compares the predictions made with the standard SWAN and 4DVar for predictions of 6 and 24 hours in advance. Predictions for 6 hours are much better than the ones created with the standard SWAN. The forecasts for 24 hours have less accuracy than the ones for 6 hours, but they still are better than the results from the standard SWAN

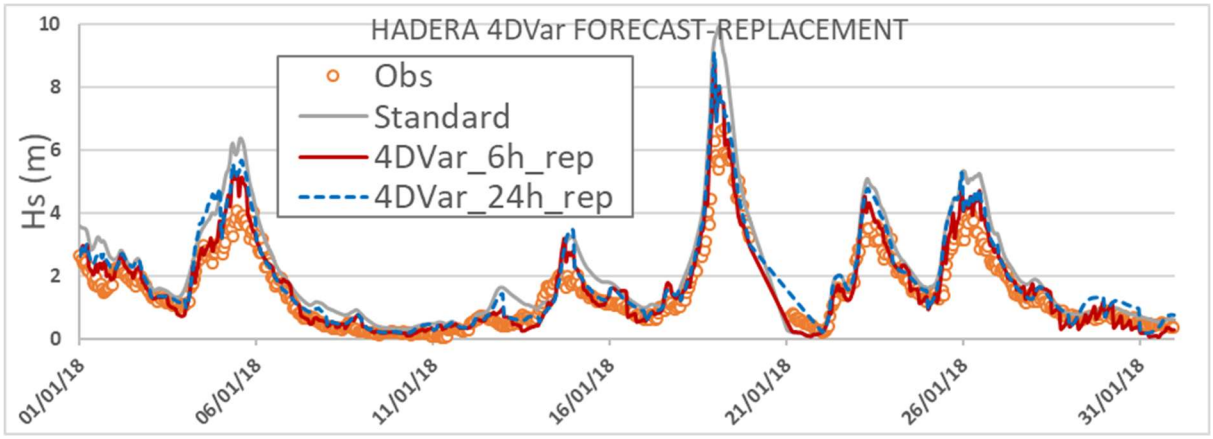

Figure 7. Forecast comparison of the standard SWAN model with the 4DVar for 6 and 24 hours.

4.4. Computational cost of assimilation techniques.

In addition to improving the prediction quality, one important factor to consider is the computational cost required to obtain such results. Table 2 presents the required computational time needed for the machine used for testing. Wall time represents the real-time that requires the calculation of the scenario. The equivalent computational time is the time that will be required if all the tasks are calculated using one single processor. The maximum number of parallel SWAN models represents how many are evaluated simultaneously. This number can take a value from 1 for the standard SWAN model, where only one model is evaluated, to the number of members required to produce the ensemble in the EnKF (in this case 5), or the number of model results used in the optimization of the 4DVar (here 8). The number of repetitions refers to the number of times the model or the assimilation process is repeated to cover the desired period. And finally, the number of cores represents the power capacity required to obtain the results in the presented wall time. The total time will be increased if a

machine has fewer cores than the one defined in this column. From these values, it is possible that with a machine with 64 cores, even the most expensive method in terms of computational power will require around half an hour. Considering the improvement in the results, it is worth using the 4DVar in the operational system to provide forecasts regularly. If fewer cores are available, the required time could increase to values that will not be useful for an operational system. In this case, with one core, the prediction for one day will require more than 30 hours to be completed.

Table 2. Computational time and core resources used by the different model and assimilation configurations.

	Wall time [min]	**Max number of parallel Swan models**	**Total cores**	**Number of repetitions [1]**	**Equivalent computational time [min]**
hindcast Standard	4	1	8	1	32 (0,5h)
hindcast EnKF	4	5	40	1	160 (2,7h)
hindcast 4DVar	32 avg.	8	64	1	1152 (19,2h)
forecast no replacement	4	1	8	1	32 (0,5h) [2]
forecast 6 hours with replacement	5 avg.	8	64	4	1056 (17,6h) [2]
forecast 12 hours with replacement	14 avg.	8	64	2	1568 (26,1h) [2]
forecast 24 hours with replacement	32 avg.	8	64	1	1824 (30,4h) [2]

1. For an equivalent comparison, it is assumed a total hindcast or forecast of 24 hours
2. All forecasts require the previous 24-hour hindcast.
3. Forecast with replacement. Runs blocks of 3 hours of the standard swan and 4DVar assimilation

5. Artificial Intelligence

This module comprises the following tasks: a) Data preparation for the ANNs, b) Selection of crucial stations for the derivation of the ANNs, c) design and training of the neural networks, d) testing the ANNs for several scenarios, e) improving the accuracy of wave predictions, and e) validation with measurements.

Data preparation and selection of stations have been made. The autoencoder's design and training based on Deep Convolutional Neural Networks (DCNN) has been the most time-consuming task as a proper selection of the number of layers. The depth of each DCNN requires expensive computing testing. All the necessary code for the training and forecast based on DCNN has been implemented.

5.1. Autoencoder depth convolutional neural networks (DCNN) for atmospheric and wave height models

An autoencoder for the wave and meteorological models has been successfully trained (Fernández Jaramillo et. al., 2023). Fig. 8 shows a scheme of the architecture of de DCNN. The information at the output should be as close as possible to the information feed at the input. The green layer represents the encoded version of the data, which is supposed to be equivalent to the input and output of the encoder. The benefit is that the encoded version allows compressing of the data and reduces the freedom degrees of the problem. Table 3, using different levels of compression reached for the meteorology, shows that a reduction of 10 times is already providing good quality results; beyond this compression still, the results could be improved as the training was still showing improvements, just that because limitations in time and the costly computational effort required for that level of compression, it was not possible to improve the results further in the frame of the project. Table 4 shows the results for the wave height model, the autoencoder is able to produce a compression of 25 times with an average error of less than 10 cm.

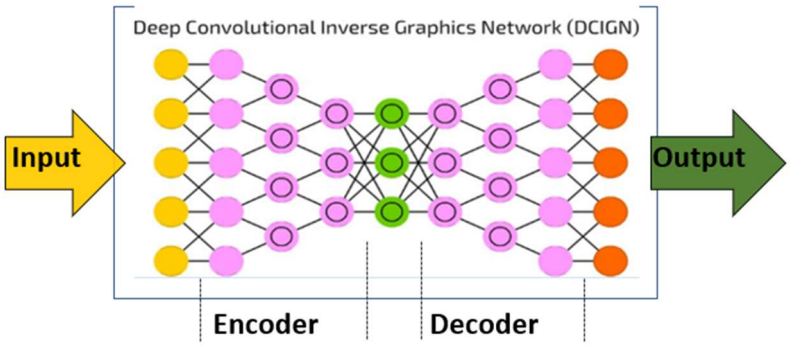

Figure 8. DCNN Autoencoder architecture.

Table 3. Different versions of the meteorological DCNN autoencoder,

Number of intermediate layers	Layer side size (cells)	Layer Depth	Compression Factor	U component difference (m/s)	V component difference (m/s)	Pressure difference (atm)
(Input) 0	512x512	3	---	---	---	---
1	256x256	32	0,4	0,17	0,17	0,0001
2	128x128	48	1,0	0,38	0,37	0,0002
3	64x64	100	1,9	0,58	0,53	0,0003
4	32x32	190	4,0	0,78	0,76	0,0004
5	16x16	310	9,9	1,21	1,14	0,0007
6	8x8	460	26,7	1,60	1,55	0,0011
7	4x4	650	75,6	2,25	2,07	0,0023
8	2x2	850	231,3	2,69	2,88	0,0035
9	1x1	1070	735,0	3,06	2,86	0,0047

Table 4. Different versions of Wave Height DCNN autoencoder,

Number of intermediate layers	Layer side size (cells)	Layer Depth	Compression Factor	Hs difference (m)
Input) 0	128x128	1	---	---
1	64x64	32	0.13	0,02
2	32x32	48	0.33	0.02
3	16x16	100	0.64	0.,03
4	8x8	190	1.34	0,04
5	4x4	310	3.30	0.05
6	2x2	460	8.90	0,05
7	1x1	650	25.20	0,06

Figure 9. presents the DCNN autoencoder for waves. With a compression of 25 times, it can reproduce the wavefield with good quality in spatial and temporal dimensions. Figures 10 and 11 present the DCNN autoencoder for the pressure field and the U wind component with a compression factor of 10. As can be seen, it also keeps the information and can reproduce the pressure and wind field satisfactorily.

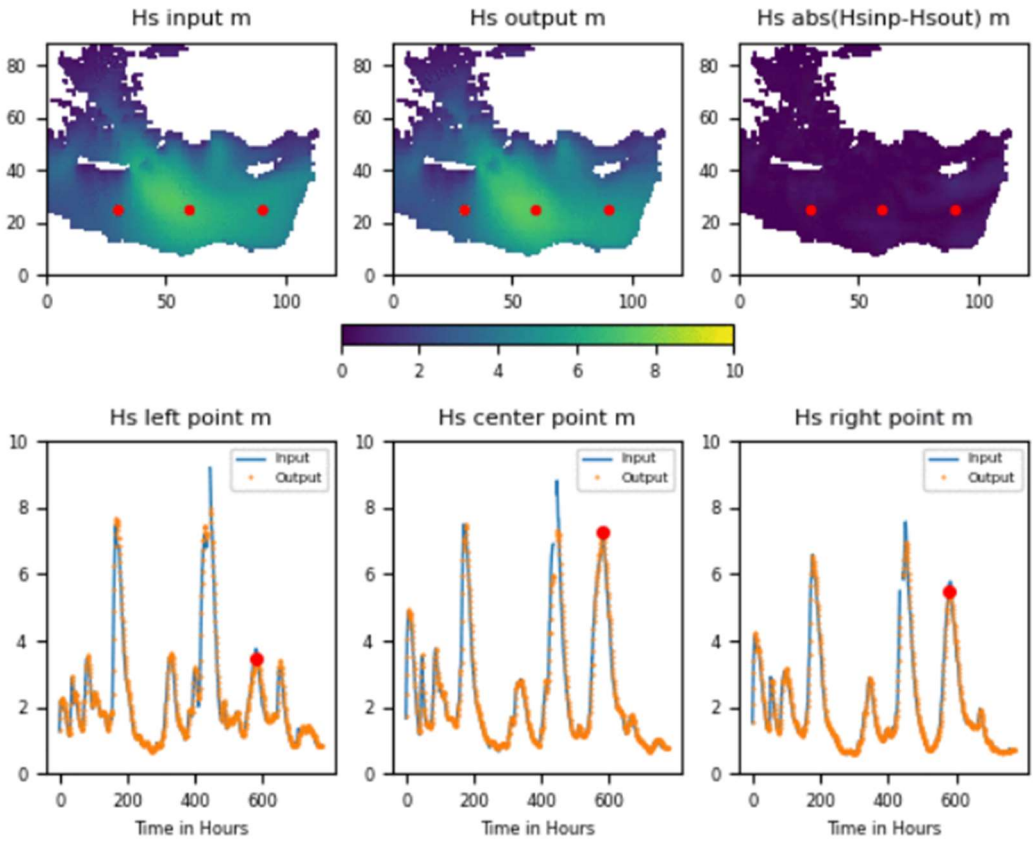

Figure 9. Wave height (in meters) DCNN autoencoder. In the upper row are presented input, output, and difference maps simultaneously (the encoded model is 54 times smaller than the input). In the lower row, the input and output as time series at three different locations.

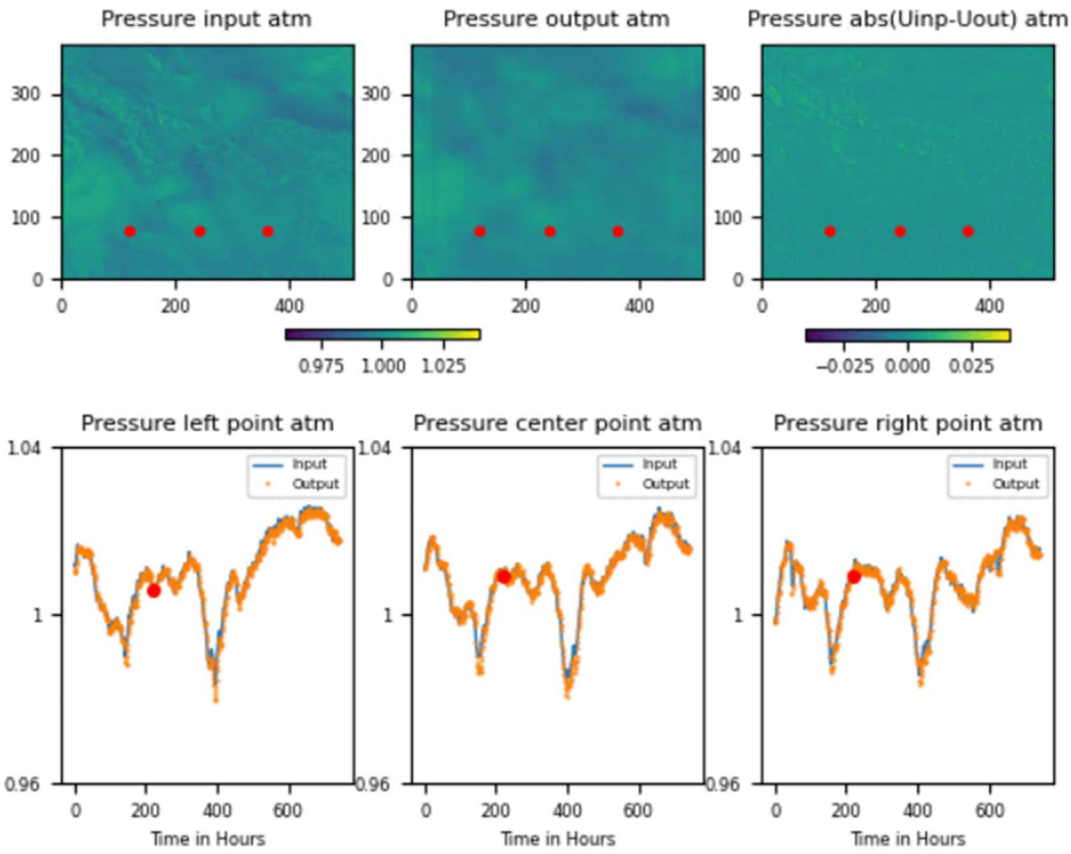

Figure 10. Pressure in atmospheres. The input, the output, and the difference in the autoencoder for a compression factor of 26 are presented in the upper row. The lower row shows the time series for the input and output pressures at three different locations.

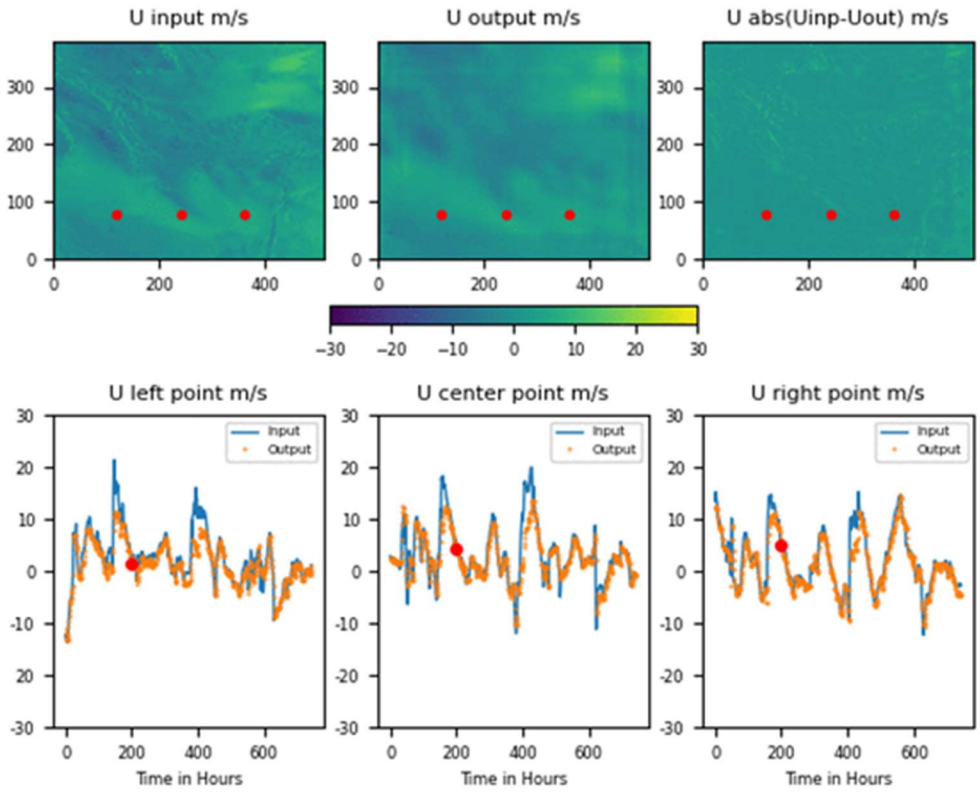

Figure 11. U component of the wind in m/s. In the upper row are presented the input, the output, and the difference in the autoencoder for a compression factor of 26. The input, the output, and the difference in the autoencoder for a compression factor of 26 are presented in the upper row in the lower row, are presented the time series for the input and the output pressures at three different locations.

5.2. Meteorological to spatial wave height depth convolutional neural network

Once the autoencoder for meteorology and wave height is defined, it is possible to connect the meteorology model's encoder with the wave model's decoder by using an intermediary fully connected artificial neural network. This is possible because as it was mentioned before, the encoded meteorology contains the information of the original input and from the encoded version of the wave height model, it is possible to reconstruct the complete domain of wave height. Figure 12 shows the results of using the Meteorology to Wave height DCNN. This figure in the upper row shows first the target output provided by the assimilation of wave heights developed in the previous working package, then the output from the neural network and the differences in the domain between both. At the bottom is the time series for January 2017 at three locations (Hadera, Sarom, and Athos). The neural network model captures the patterns and provides a good result. Figure 13 shows the time series for wave height at Hadera station for January 2017. This figure shows the measurements, the results from the 4DVar data assimilation, and the neural network model. Despite the fact that the results from data assimilation are better than those using the neural network model, the results from

the last are still reasonable but obtained with a minimal computational cost when compared with the data assimilation.

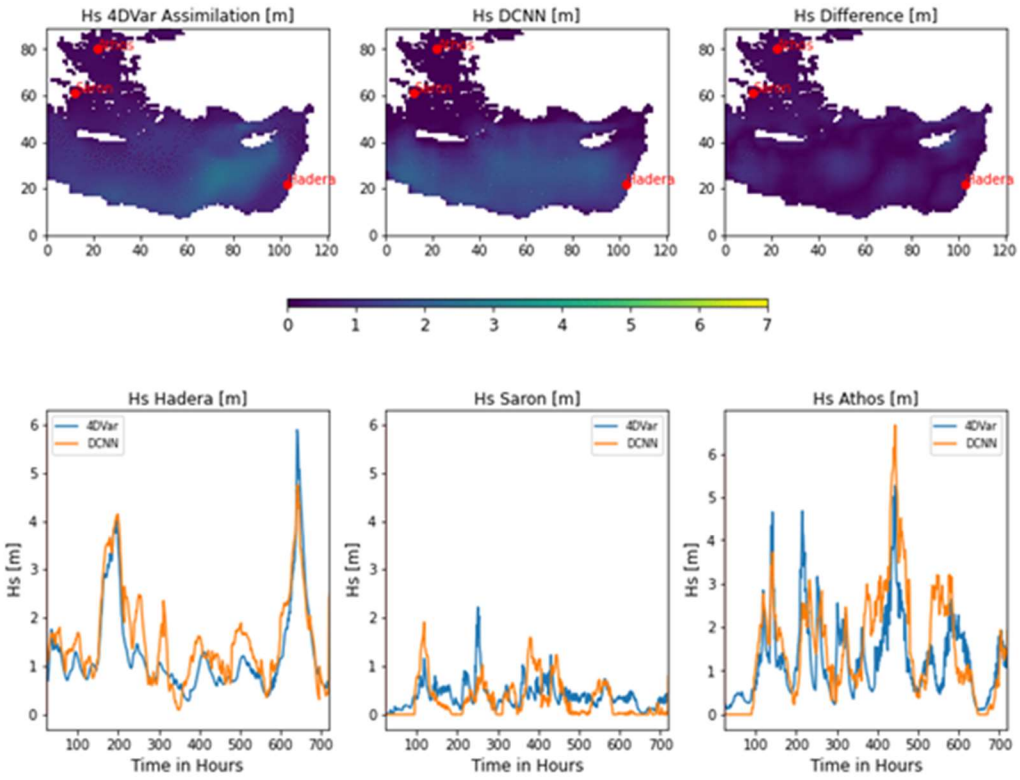

Figure 12. Meteorological to wave DCNN. The time series are for January 2017

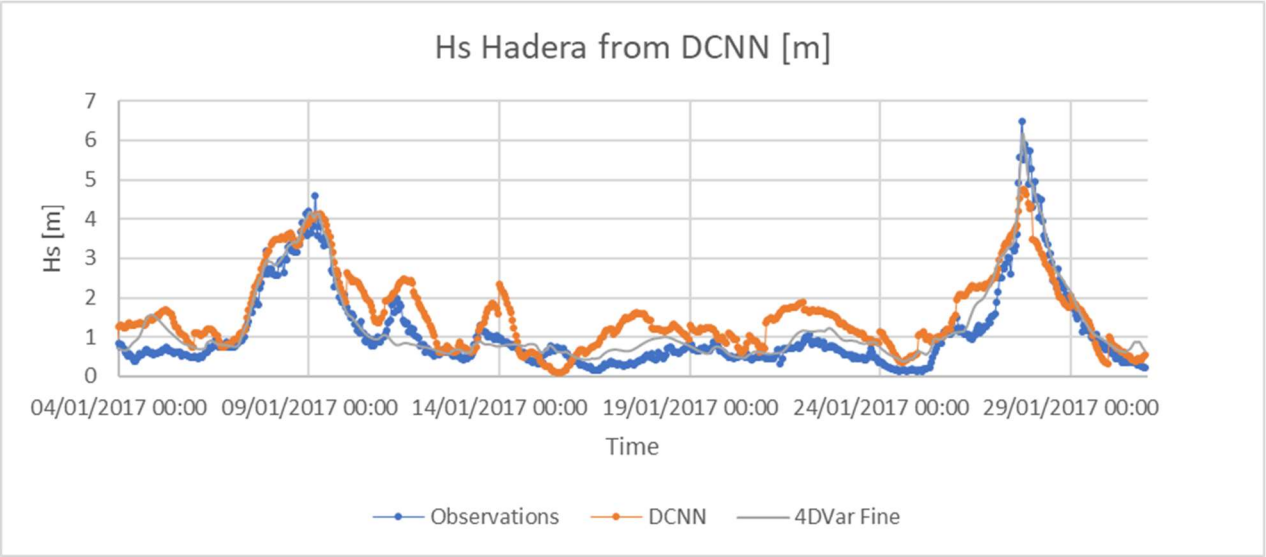

Figure 13. Wave height at Hadera for January 2017. Wave height observations, deep convolutional network model, and 4DVar assimilation.

Table 5 compares the wave height results of the depth convolutional neural network model with the standard and 4DVar assimilation. Results from the DCNN are in between both methods. This DCNN model can significantly improve the standard model taking only 20 seconds using one single core for a prediction of 24 hours, without incurring the high extra computational effort that is required by the assimilation methods that can reach up to 30 hours with the same computing power (see Table 2).

5.3. Local wave high prediction hybrid model

A local hybrid ANN model to improve the prediction of waves by combining the standard wave predictions with meteorological models in one single position has been developed and tested with good results (Dakar et. al. 2023). For this neural network model, the input of meteorology was considered at several time steps giving a better chance to capture the pattern of movement of the wind, wave height measurements, and also model outputs from the standard SWAN (see Fig. 14). The use of the wave model and observations allows the ANN to learn how much the model must be adjusted according to the meteorological conditions.

The performance of this hybrid model is closer to the one obtained by the 4DVar, being able to improve the results compared with the standard SWAN and the DCNN. The computational cost of this hybrid model is the sum of the evaluation of the standard SWAN model (4 min by using 8 cores and up to 32 min. Using one single core) plus other 20 seconds for 24 hours forecast. The limitation is that the output is only available for one single point. However, when the interest is localized in one single station, the improvements obtained with this method are excellent, with a minimal extra computational effort compared with the 4DVar assimilation.

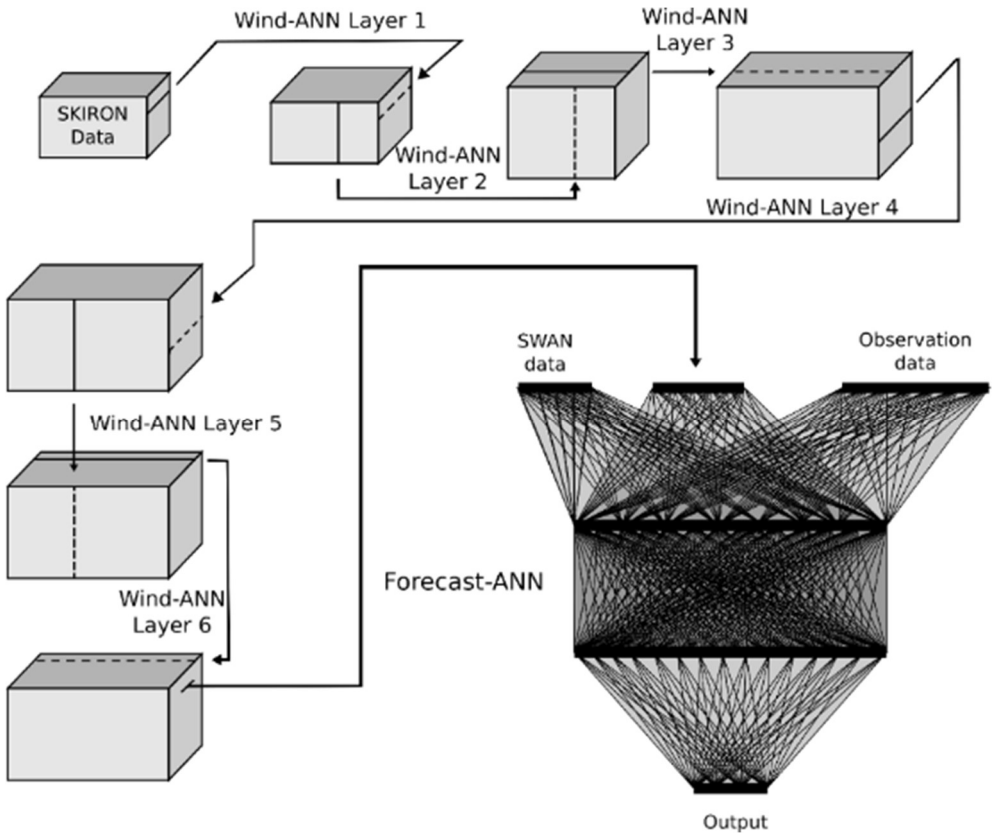

Figure 14: A schematic chart of the ANN structure and data flow. Each block denotes a 3D tensor of the wind dataset; the wind dataset is a 4D tensor. Each line inside a face of a block represents a vector within the tensor. Each layer sums along a specific dimension of the tensor. The forecast-ANN input concatenates the Wind-ANN output, Observation data time series, and SWAN data time series.

5.4. Comparison of results from physical-based and neural network predictions

Table 5 compares the statistical results from some of the used models for wave height calculation. Here we have the standard SWAN, the 4DVar assimilation, the DCNN meteorological to wave model, and Hadera's hybrid SWAN-ANN local model. The 4DVar and the local hybrid models improve in a significant way the results from the standard model in any of the parameters used for the comparison. It shows the advantage of methods that include the assimilation of real observations to increase the quality of the results. The DCNN has not been designed to consider real-time observations but for learning the improved version of the numerical model after the assimilation of previous sea states. DCNN model produces wave heights with a quality intermediate to the standard SWAN and the methods including assimilation. DCNN improves the standard model in all the parameters except for the regression coefficient. From this, we can conclude that the best results are obtained from the 4DVar, which comes with a high computational cost. The local hybrid model requires a computational effort similar to the standard model, but its high-quality results are only available for one location. Finally, DCNN produces an intermediate improvement but with the advantage of a much reduced computational effort and can produce outputs much faster than any of the other methods.

Table 5. Statistical comparison of the standard wave model, the 4DVar assimilation, Depth Convolutional Neural Network, and the local hybrid ANN model.

	Standard (no assimilation)	**4DVar (hindcast)**	**DCNN**	**Local Hybrid ANN**
Samples	648	648	648	576
Bias [m]	-0,72	-0,16	-0,44	0,13
R	0,93	0,96	0,89	0,96
MAE [m]	0,73	0,25	0,55	0,33
RMSE [m]	0,88	0,33	0,65	0,49

6. Improve Wind Drag Coefficient

In working package one, the data assimilation was developed by using a wind factor that allows the wind speed adjustment so the SWAN model can adjust better to the wave height measurements. As for each assimilation window from the 4DVar, we obtained one wind factor representing the optimal factor for such a window. Figure 15 shows one example of the factors obtained for January 2017. At the beginning of the project, we planned to test if it was possible to find a simple function to apply this correction as part of the wind drag coefficient. Still, we find that the delivered coefficients show some oscillations that could have been more challenging to provide a simple function.

The other alternative that was investigated consisted of developing an artificial neural network that captures the relation between meteorology and wind factors. The developed neural network reuses the encoders for the wind generated in the previous section and connects 13 of them spaced every two hours to cover a total period of 24 hours before estimating a wind factor (see Fig 16). During the training of this network, the rapid oscillations of the factor from the assimilation still create problems for the convergence of an optimal model, to reduce these problems, it was required to apply a moving average to filter high-frequency oscillations (see Fig. 15).

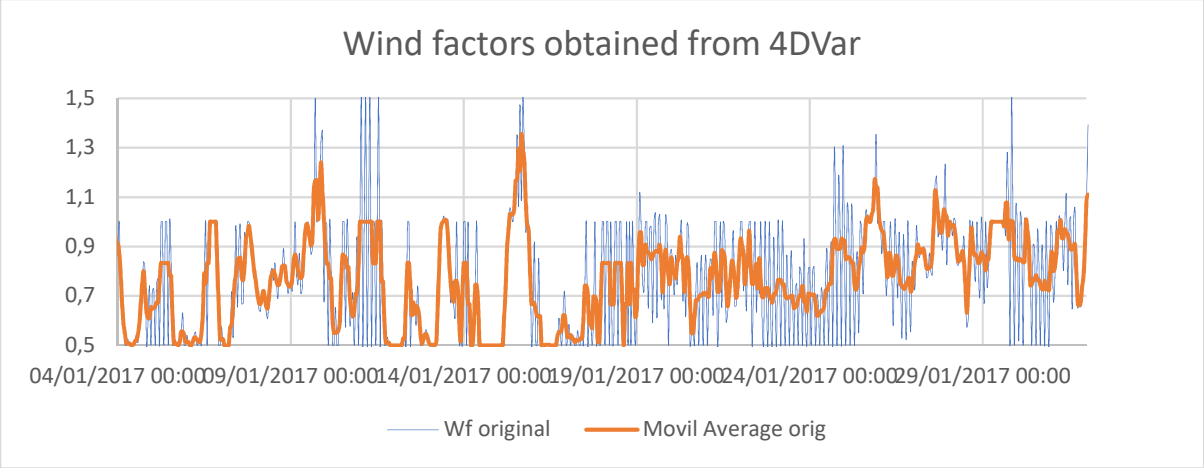

Figure 15. Wind amplification factors were obtained from data assimilation.

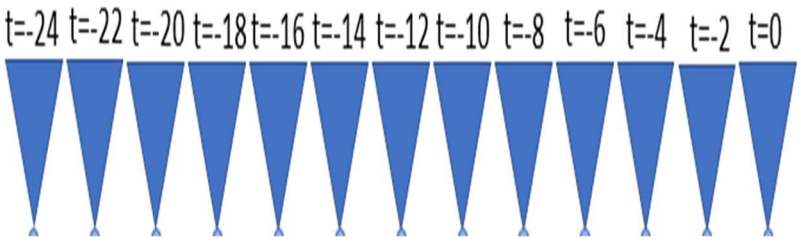

Figure 16. ANN model combines 13 meteorological inputs every two hours to produce the equivalent wind factor to be used in the SWAN model.

For the application of the ANN model, the wind factors are calculated from the meteorological model before the execution of the SWAN model. Then the modified wind fields are used to calculate the wave using the standard SWAN model. It can be seen that the method, in general, can improve wave forecasts, especially during storms (See Fig. 17). This method is not as good as the hybrid local presented in working package two, as is not including measurements or the SWAN as part of the inputs. Modifying this neural network that includes measurements could improve this strategy even more in the future.

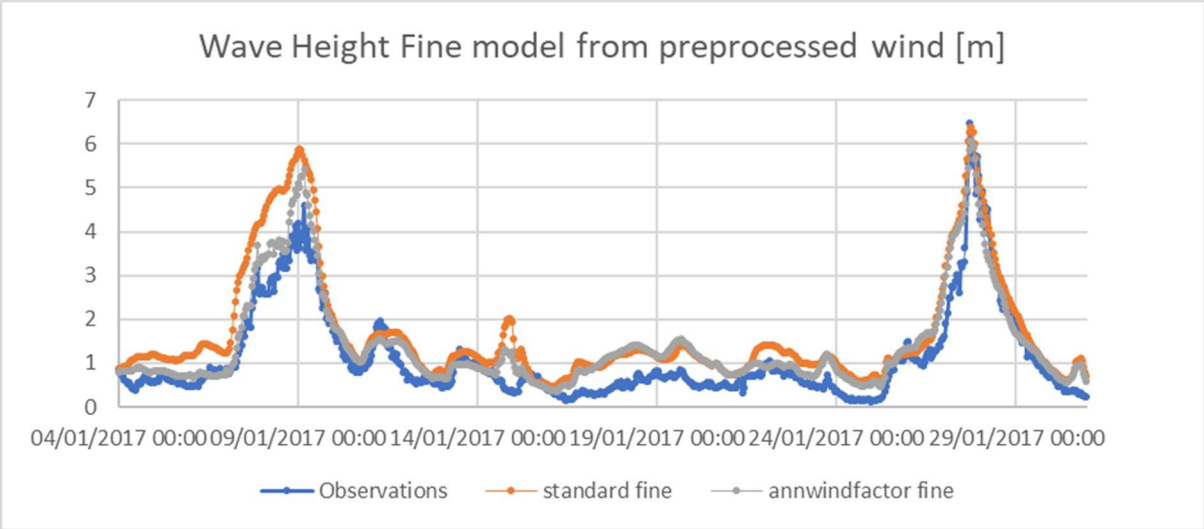

Figure 17. Comparison of the wave height at Hadera. Blue represents the measurements, orange is the output from the fine model of SWAN, and grey is the values obtained by using the modified wind factor obtained from the ANN.

7. Comparison of all the developed methods.

When the standard SWAN model was available at the beginning of the project at IOLR, it is possible to see (see Fig18.) that all the developed models are able to provide improvements in different grades. Especially the bias, mean root square, and mean average errors (see Table 6). The correlation coefficient is improved in all of them except in the DCNN model.

Each model at the actual development status responds to different requirements of input data (more data, more fragility in case some are missing) and the potential to provide spatial or being restricted to one single position prediction. There is still potential to improve some of these methods by combining the input elements used in some and missing in others. The necessity of using several time steps for the input of the ANNs is evident. That could benefit the quality of the DCNN that even its spatial prediction has difficulties capturing the dynamics of the storm. The wind factor estimation that already considers the time variation could benefit from including a term that involves the error in the prediction as the hybrid model can consider internally, and the local hybrid model could be improved by adding the decoders for meteorology as outputs to provide spatial predictions.

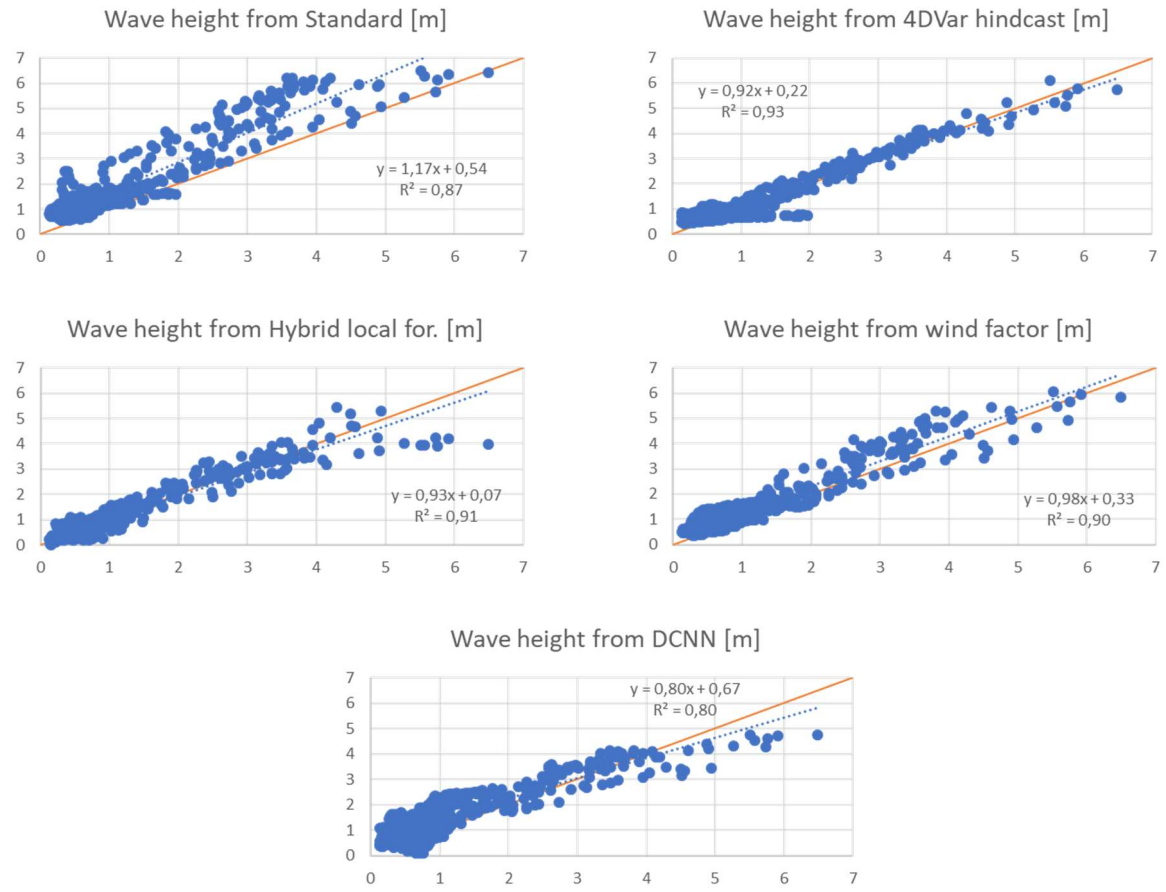

Figure 18. Comparison of wave height from the observations at Hadera measurement station to the different methods developed along the project. Upper left: Standard SWAN model. Upper right: 4DVar assimilation for the hindcast. Center left: Hybrid local model. Center right: Wave height from precalculated wind factors. Bottom: Wave height from the DCNN model.

Table 6. Statistical comparison of wave height estimation from different kinds of models in all working packages.

	Standard	**4DVar**	**Hybrid local**	**DCNN**	**DCNN Wf**
Samples	648	648	576	648	648
Bias (m)	-0,72	-0,16	0,13	-0,44	-0,32
R	0,93	0,96	0,96	0,89	0,95
MAE (m)	0,73	0,25	0,33	0,55	0,37
RMSE (m)	0,88	0,33	0,49	0,65	0,47

8. Improve Operational Wave Model East Mediterranean Sea

Along with implementing the different algorithms for data assimilation, neural network models, and hybrid models, the schemes have been developed to minimize the differences with the existing operating system in IOLR. Meteorological information and outputs from the different systems are the same as the ones used in the existing ones. To implement the algorithms, offline data with wave measurements have been used. Due to the assimilation of real-time information, we used an offline dataset during the research. But later, they were integrated into the existing systems operated at IOLR for real-time operation.

All modules have been implemented using containers (see Fig. 19), allowing an easy implementation as we could avoid many issues with the physical infrastructure and software conflicts at the host operative system used by the different partners. Each tool runs isolated, passing only relevant inputs and outputs outside the container. All the results are made available to be displayed to the final users via a web interface.

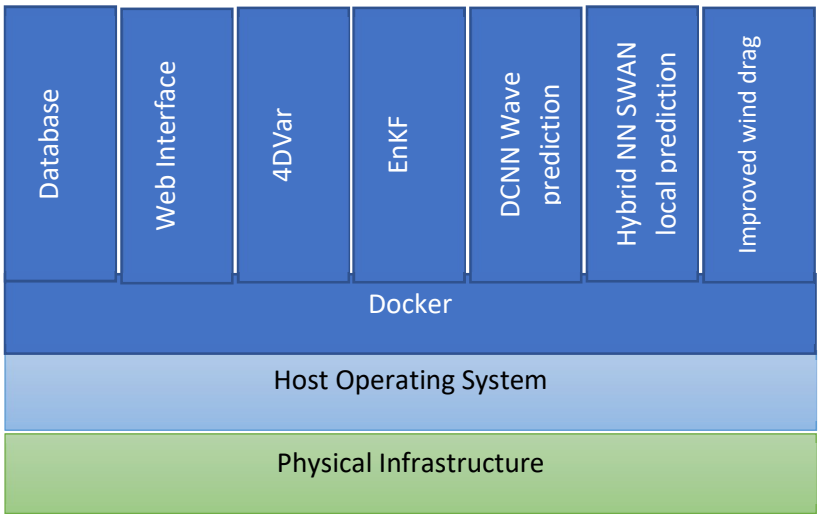

Figure 19. Schema of containers for the different applications developed in the project.

Figure 20 depicts the web interface showing results from four different modules for forecasting waves: the SWAN-4DVar assimilation, DCNN for spatial prediction, the SWAN-ANN hybrid model for local wave, and the Improved wind field SWAN model. This web interface provides an interactive map where information from a defined coordinate can be extracted and presented as time series (See Fig. 21).

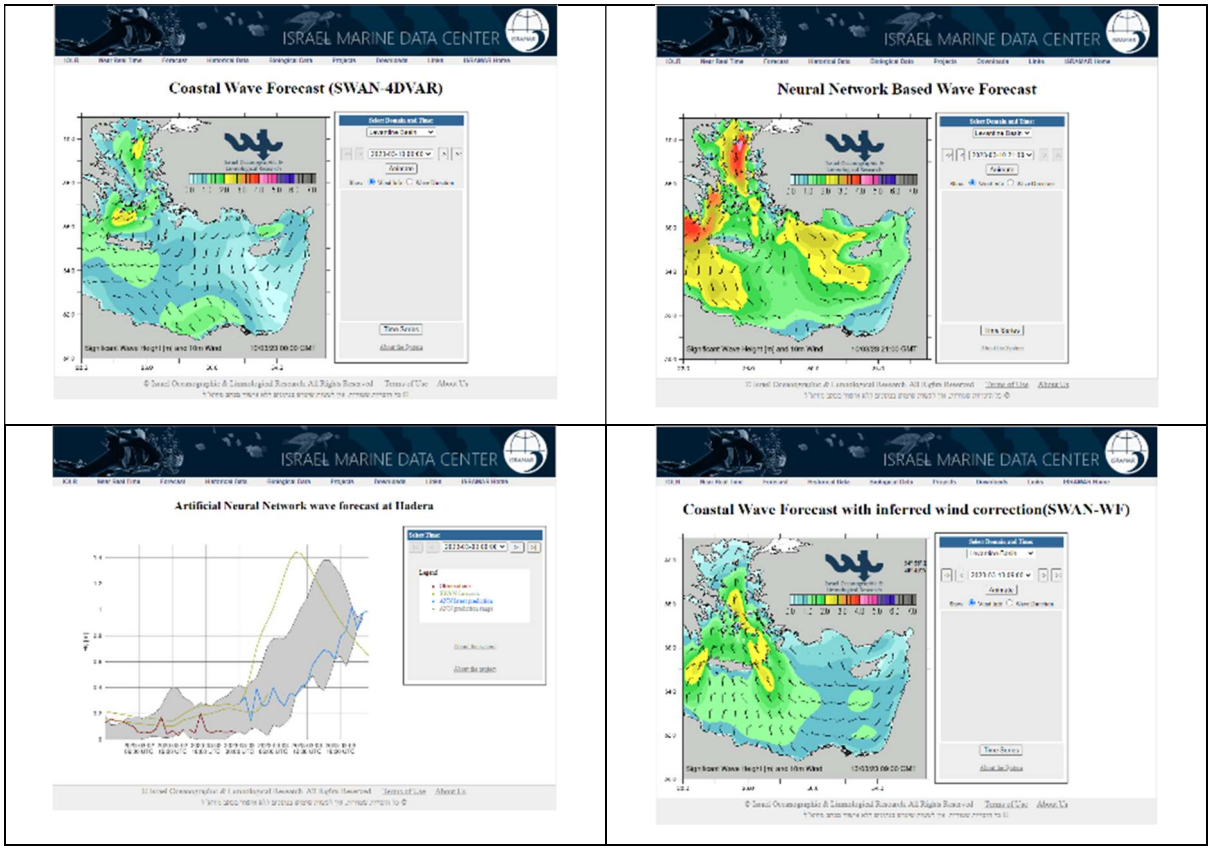

Figure 20. Screenshots of the operational system at IOLR. a) 4DVar wave data, b) DCNN wave model, c) Hybrid local wave model, d) wind improved model.

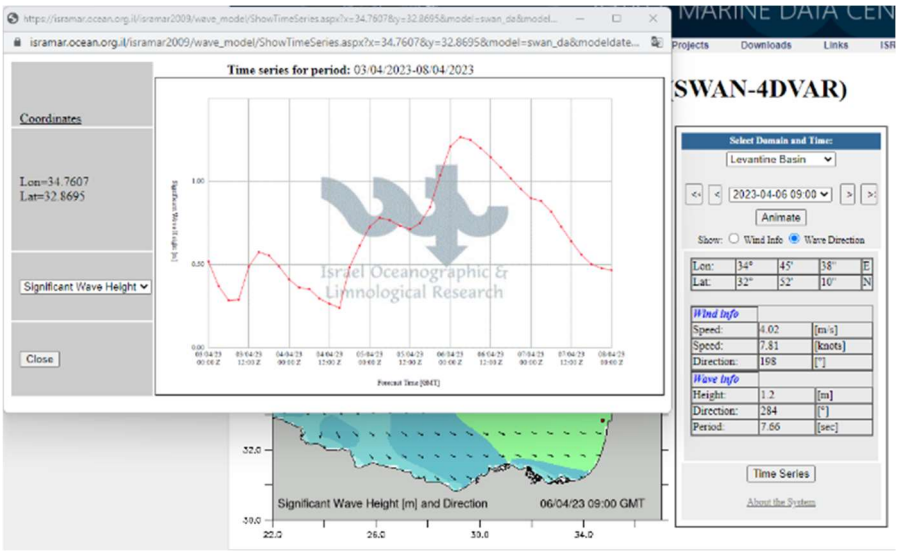

Figure 21 Time series and information for a single point obtained from the web interface.

Conclusions

With this project, we have tested and implemented different techniques for improving wave height estimation in the Levantine region of the Mediterranean Sea. The 4DVar and EnKF have been able to produce very accurate hindcast results, and 4DVar was able to create a significant improvement in the forecast. The disadvantage of these methods is the computational cost that is involved. Novel techniques involving artificial neural networks have increased the quality of the results with a much lower computational effort. Each of the proposed methods involving artificial neural networks has some limitations based on the type of inputs used. These limitations could lead to a future improvement if a new neural network model considers a combination of the virtues of the ones presented here. The hybrid model uses measurements and SWAN estimation, while DCNN or the improvement of the wind field can provide a spatial domain representation instead.

Literature

Booij, N., Ris, R. C., & Holthuijsen, L. H. (1999). A third-generation wave model for coastal regions: 1. Model description and validation. *Journal of geophysical research: Oceans, 104*(C4), 7649-7666.

Copernicus. Mediterranean Sea- In-Situ Near Real Time Observations. Doi: 10.48670/moi-00044

Dakar, E., Fernández Jaramillo, J. M., Gertman, I., Mayerle, R., Goldman, R. (2023) An artificial neural network based system for wave height prediction, Coastal Engineering Journal, DOI: 10.1080/21664250.2023.2190002

Fernández Jaramillo, J. M., Dakar, E., Restrepo Álvarez, D., Goldman, R., Mayerle, R., Gertman, I. (2023). A wind model output compressor using a Deep Convolutional Neural Network. Manuscript in preparation.

Gertman, I., Murashkovsky, A., Levin, V., Kallos, G., & Rosen, D. (2006). Wave monitoring and wind input as key issues in operational wave forecasting systems. *European Operational Oceanography: Present and Future*, 743-749.

Gertman, I., Rosen, D. S., Kariel, S., & Raskin, L. (2000). Comparison of two years of wind and wave hindcasts via WAM based operational forecasting system versus field and other models data. *Preprints 6th International Workshop on Wave Hindcasting and Forecasting*, (pp. 91-98).

Jakobsson, M., Mayer, L., Coakley, B., Dowdeswell, J. A., Forbes, S., Fridman, B., Hodnesdal, H., Noormets, R., Pedersen, R., Rebesco, M., Schenke, H. W., Zarayskaya, Y., Accettella, D., Armstrong, A., Anderson, R. M., Bienhoff, P., Camerlenghi, A., Church, I., Edwards, M., Gardner, J. V., Hall, J. K., Hell, B., Hestvik, O., Kristoffersen, Y., Marcussen, C., Mohammad, R., Mosher, D., Nghiem, S. V., Pedrosa, M. T., Travaglini, P. G., Weatherall, P. (2012). The International Bathymetric Chart of the Arctic Ocean (IBCAO) Version 3.0. Earth and Space Science. Volume 39, Issue 12, June 2012. DOI: 10.1029/2012GL052219

Kallos, G., Nickovic, S., Papadopoulos, A., Jovic, D., Kakaliagou, N., Misirlis, L., Anadranistakis, E. (1997). The regional weather forecasting system SKIRON: An overview. *Proceedings of the symposium on regional weather prediction on parallel computer environments. 15*, p. 17. University of Athens Greece.

Restrepo Álvarez, D., Fernández Jaramillo, J. M., Goldman, R., Mayerle, R., Gertman, I. (2023). Implementation of an EnKF for wave prediction in the East Mediterranean Sea. Manuscript submitted for publication.

Weatherall, P., Marks, K. M., Jakobsson, M., Schmitt, T., Tani, S., Arndt, J. E., Rovere, M., Chayes, D., Ferrini, V., Wigley, R. (2015). A new digital bathymetric model of the world's oceans. Earth and Space Science. Volume 2, Issue 8, August 2015, (pp 331-345), DOI: 10.1002/2015EA000107.