# Assessment of impacts of coastal flooding on coastal urban areas and coastal erosion developed

## Final version

Deliverable number 5.5.4

| | |
|---|---|
| ***Deliverable title:*** | Assessment of impacts of coastal flooding on coastal urban areas and on coastal erosion |
| ***Deliverable number:*** | D.5.5.4. |
| ***Deliverable authors:*** | Ivan Federico (CMCC), Giovanni De Cillis (CMCC), Salvatore Causio (CMCC), Giovanni Scicchitano (UNIBA), Giovanni Scardino (UNIBA), Giusy Fedele (CMCC), Ilenia Manco (CMCC), Alessandro Pugliese (CMCC), Paola Mercogliano (CMCC), Valeria Intini (ASSET), Filomena Carbone (ASSET), Giuseppe Garofalo (ASSET) |
| ***Date release:*** | 2023-06-09 |
| ***Submitted by:*** | CMCC |

# TABLE OF CONTENTS

# 1 INTRODUCTION

The STREAM project deals with territorial challenges connected to flooding in the different Adriatic regions.

Areas of potentially significant risk of flooding and/or erosion can have very different characteristics, therefore requiring different modelling systems. The STREAM pilot sites involved in the modelling and forecasting activities greatly differ for morphological characteristics, potential flood and erosion hazards and vulnerabilities.

The Apulia Pilots have been grouped according to the following flood types:

- coastal flood: **Peschici and Manfredonia towns (Apulia Region, Italy).**
- fluvial flood: **Ofanto watersheds (Apulia Region, Italy);**
- coastal erosion: **Lecce and Torchiarolo wetlands (Apulia Region, Italy).**

In Deliverable D5.5.1 we have introduced the modelling settings for each component (ocean, atmospheric, flooding). Here we describe the main results and the validations.

# 2 RESULTS AND VALIDATION

## 2.1 Validation of the subregional and high-resolution coastal models

Here the main results of SHYFEM, WW3 and WRF models are reported for hydrodynamics, waves and atmospheric fields respectively.

### 2.1.1 Ocean and wave component

We have performed a simulation of 3 years (2018-2020) with SHYFEM model for the configuration AdriFs, while 2 years (2018-2019) of simulation for SOAP (Southern Adriatic Apulia forecasting system) system. Both model configurations are described in D5.5.1 §3.1.1. In Figure 1 we report the seasonal-average circulation fields at 30m produced by the AdriFs for the basin-scale, showing the main structures and dynamics of Adriatic Sea. The same has been reported for Figure 2 but for the SOAP configuration.

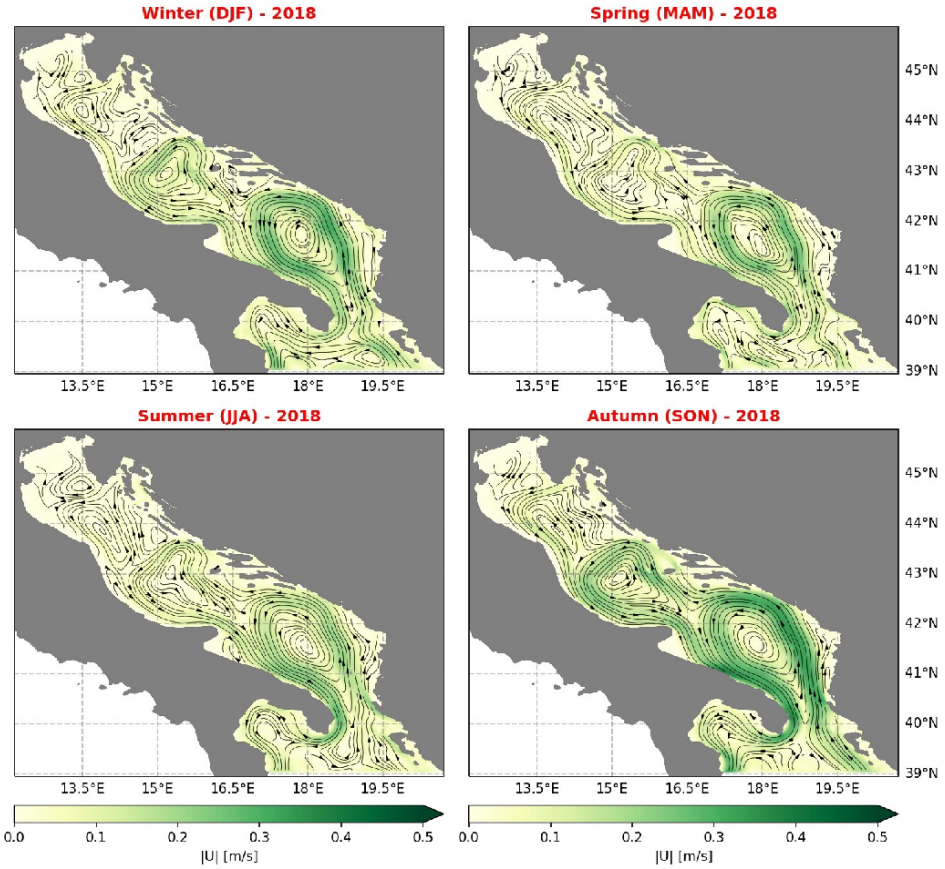


Figure 1: Seasonal-average basin-scale circulation at 30m of high-resolution coastal model AdriFs for Adriatic Sea

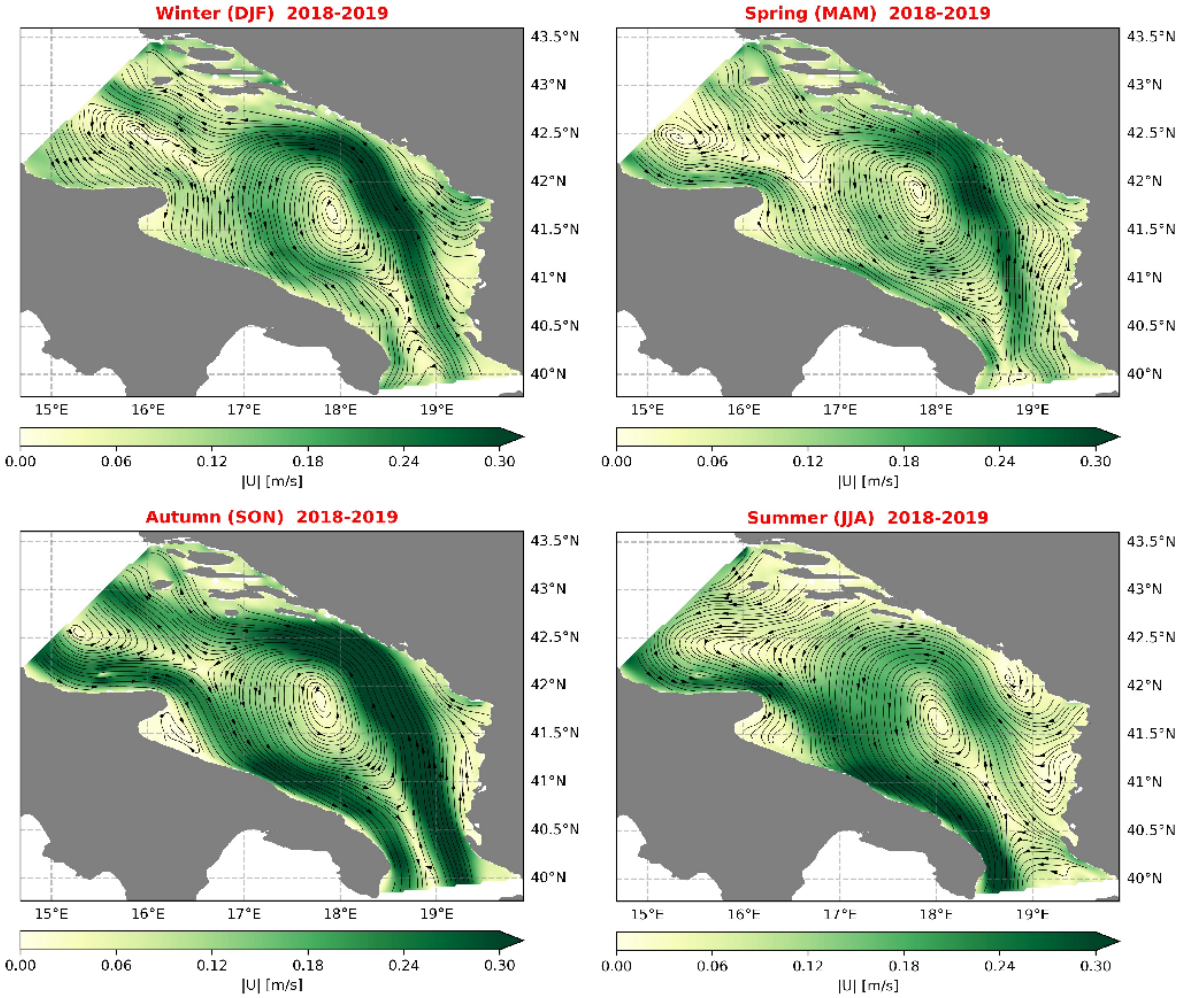


Figure 2: Seasonal-average basin-scale surface circulation in the Southern Adriatic Sea from very high-resolution SOAP coastal model system of Apulia Pilots

The AdriFs simulation carried out in active model for the period 2018-2020 (with one month - December 2017 - of spin-up) has been analyzed to assess the accuracy of model in comparison with satellite temperature. In particular, the temperature field at surface layer of model has been compared with CMEMS observation product: Mediterranean Sea Ultra High Resolution (0.01°) Sea Surface Temperature Analysis (SST_MED_SST_L4_NRT_OBSERVATIONS_010_004, Buongiorno

Nardelli et al., 2013). The comparison refers to the instantaneous value at 00:00 of the model surface temperature against the satellite foundation SST (~ SST at midnight). Figure 3 shows the time series (January 2018-December 2020) of daily SST, BIAS and RMSE between model and satellite observation, averaged over the whole Adriatic Sea domain. The model is in well agreement with the observation, describing the pattern of seasonal cycle of temperature (Figure 3a). BIAS and RMSE time series are displayed in Figure 3b and Figure 3c respectively, subdivided also for years. The mean (over the entire domain and entire timeseries 2018-2020) RMSE is 0.87C.

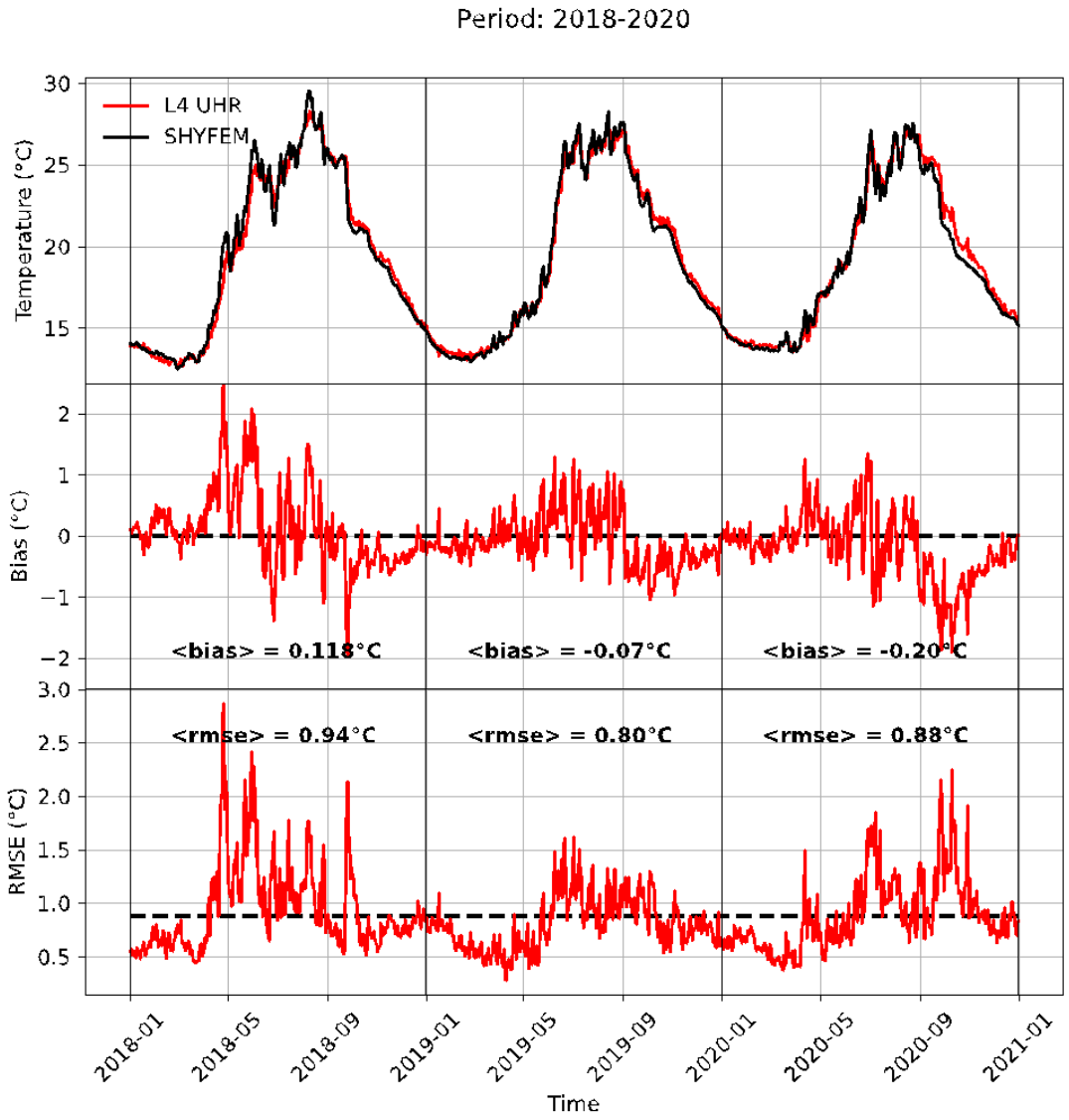

Figure 3: Time-series comparison between model surface temperature (instantaneous value at 00:00) and satellite foundation SST (~ SST at midnight), averaged over the Adriatic Sea basin (a); BIAS (b) and RMSE (c) between model and observation (CMEMS Observation products: SST_MED_SST_L4_NRT_OBSERVATIONS_010_004).

Same comparison as in Figure 3 has been performed for the SOAP system for the time period 2018-2019. The results are shown in Figure 4 and the mean RMSE is 0.77C.

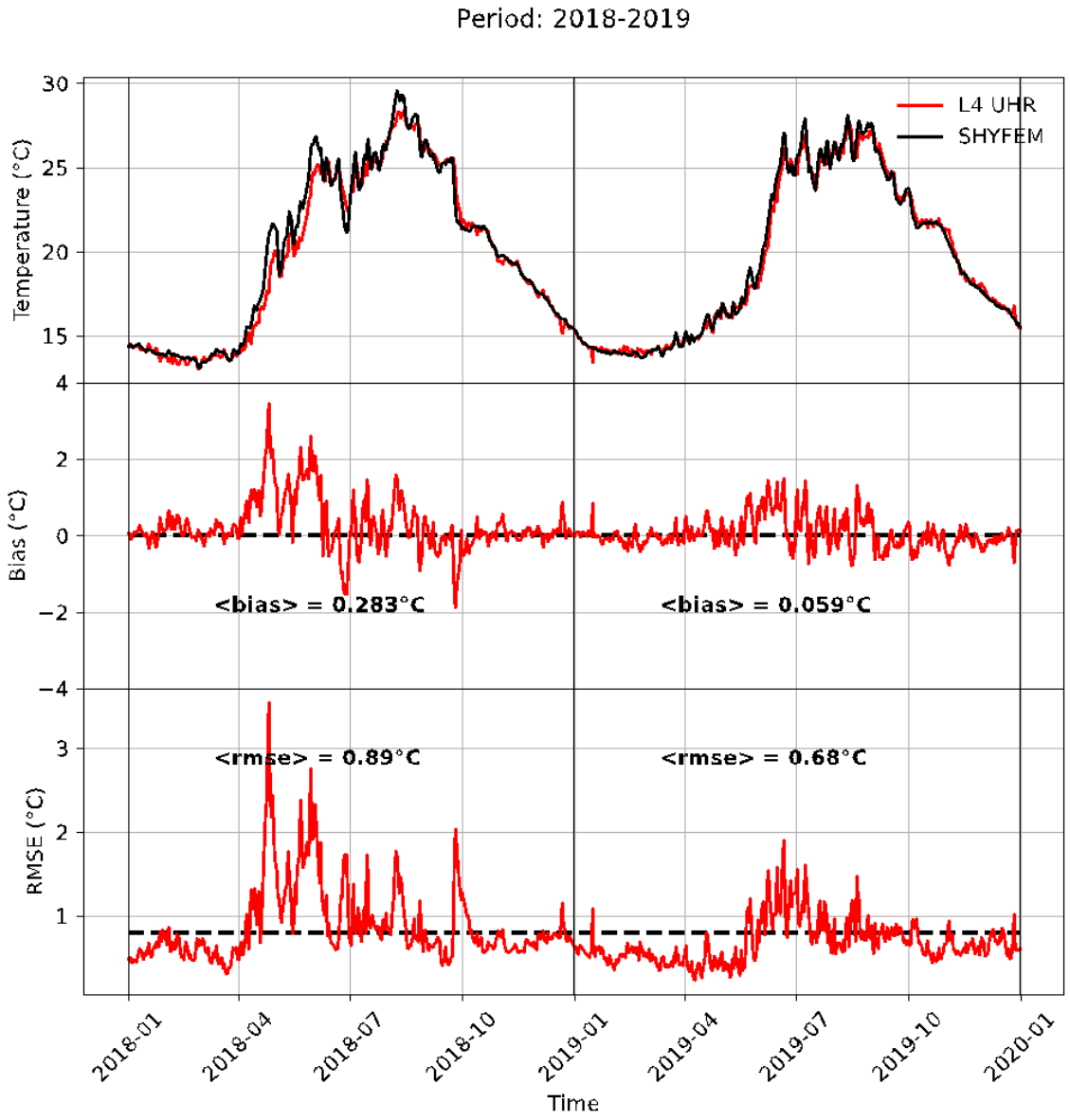


Figure 4: Time-series comparison between model surface temperature (instantaneous value at 00:00) and satellite foundation SST (~ SST at midnight), averaged over the Southern Adriatic basin for SOAP configuration (a); BIAS (b) and RMSE (c) between model and observation (CMEMS Observation products: SST_MED_SST_L4_NRT_OBSERVATIONS_010_004).

Comparisons with vertical profiles of temperature and salinity have performed for the SOAP system. The data used are ARGO floats provided by the Copernicus Marine Service. In Figure 5a we show the averaged profile of model temperature and salinity compared with the observed ones over the entire period (2018-2019) and domain (Southern Adriatic Sea). BIAS for temperature (Figure 5b) and salinity (Figure 5c) are also reported grouping the observed profiles for seasons (December-January-February DJF, March-April-May MAM, June-July-August JJA, September-October-November SON).

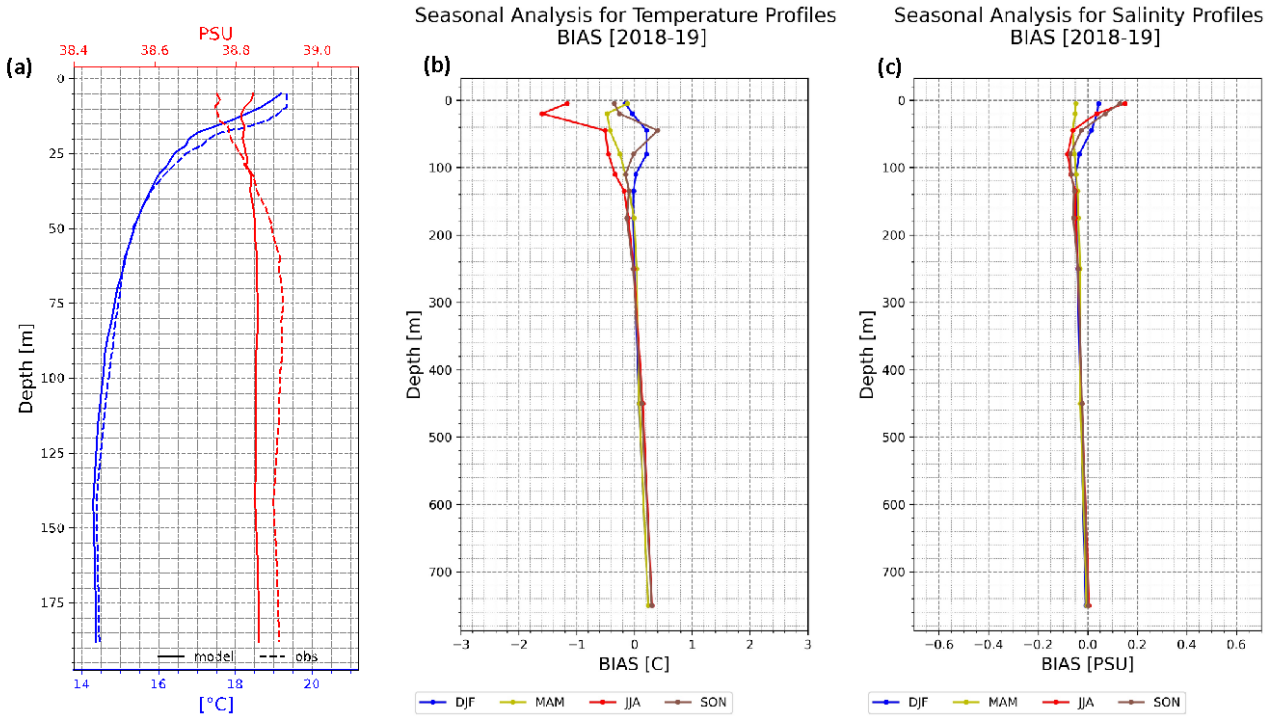


Figure 5. Model (SOAP) average profiles (temperature and salinity) compared with the observed ones taken from Copernicus Marine Service. Temperature and salinity BIASes are also reported grouping the observed profiles for seasons.

A further analysis has been performed on significant wave height over the entire Adriatic Sea, considering the satellite track for the period 2018-2020. The comparison is described in Figure 6 in terms of scatter plot. The satellite used are Criosat-2, AltKa, HY-2B, Jason-3, Sentinel-3A, Sentinel-3B.

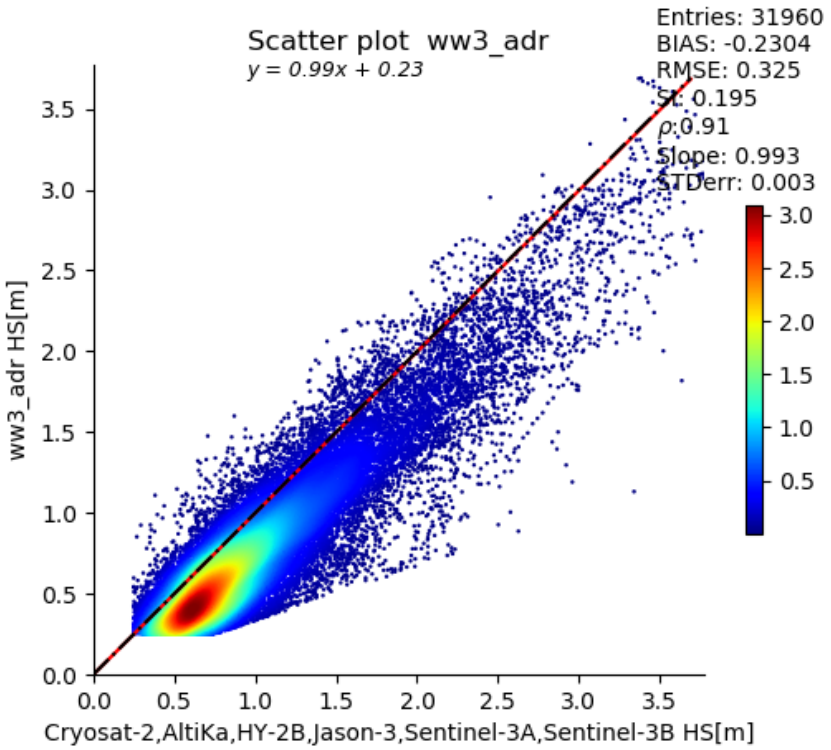


Figure 6. Scatter plot of modelled and satellite observed Significant Wave Height over the entire Adriatic Sea for the period 2018-2020

Further investigations and comparisons have been done on sea level tailored the storm surge event. Here we report for the area of Manfredonia the time series of the detited model sea level and we qualitatively compare with SLA data satellite-derived (Figure 7). Then we perform a short-term simulation for the event of November 2019 and we extract timeseries of sea level and significant wave height (Figure 8) for the areas close to Manfredonia-Peschici and Ofanto to force the downstream flooding models (§2.2).

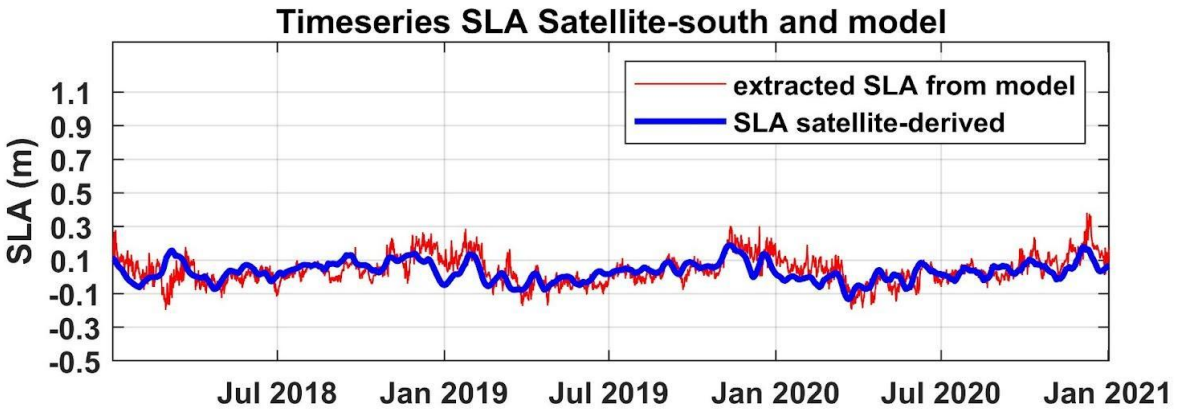


Figure 7. Time series of de-tided modelled and satellite-derived sea level in Mafredonia for the period 2018-2020

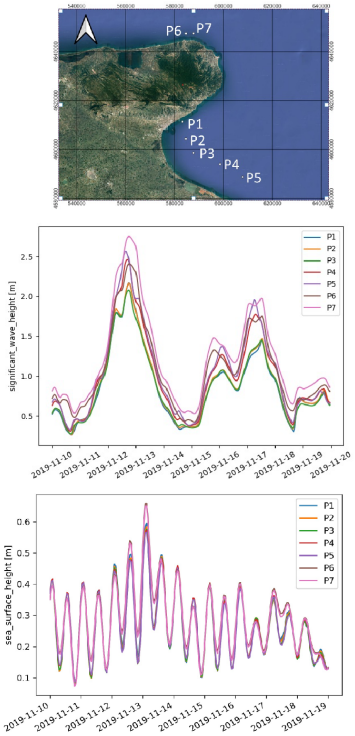


Figure 8. Time series of modelled significant wave height and sea level in Peschici_manfredonia e Ofanto for the event of November 2019

## 2.1.2 Atmospheric component

In this section the main results of the simulations performed over the Apulia Region with the Weather Research and Forecasting Model (hereafter WRF; Skamarock et al., 2019) in the context of the STREAM project are presented. In addition to the two WRF configurations presented in D.5.5.1, at high (WRF-2km) and very high (WRF-566m) horizontal resolution respectively, an additional configuration at medium resolution (WRF-6km), previously developed at CMCC over the Italian Peninsula, with the same modelling configuration of WRF-2km and WRF-566m, has been integrated, in order to investigate the impact of the horizontal resolution at different spatial scales over the Apulia Region (Figure 9). The modelling configuration of these three simulations is shown in Table 1.

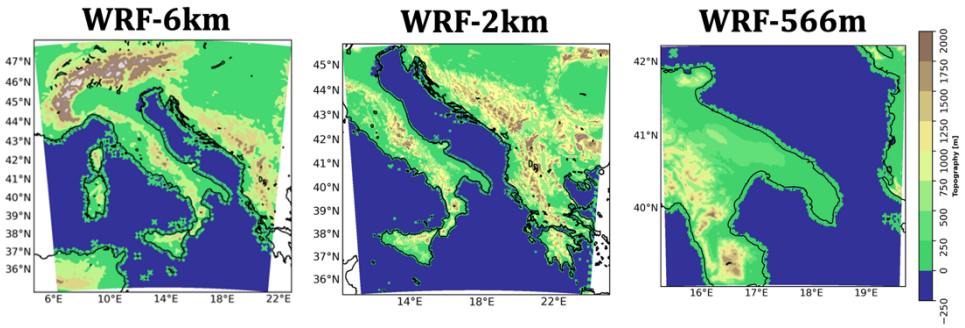


*Figure 9. Simulated domain in (a) WRF-6km, (b) WRF-2km and (c) WRF-566m*

| Table 1. Modelling Configuration | WRF-6km/ WRF-2km / WRF-566m |
|---|---|
| Model | WRF |
| Forcing | IFS (ECMWF) 0,075° |
| Grid | Lat-Lon regular (Lambert) |
| Horizontal Resolution | ~6000m/1700m/566m |
| Horizontal Discretization | Arakawa-C Grid |
| Timestep | 36 s / 12 s / 3 s |
| Vertical Coordinates | Sigma-pressure (60 vertical levels) |
| Temporal Integration Scheme | 3rd order scheme Runge-Kutta |
| Spatial Integration Scheme | 6th order centered difference |

By construction, since the IFS analyses are used as initial and boundary conditions for the parent domains, the models' performance can only be associated to their modelling configuration, and therefore any information can be provided in relation to the forecast error.

In order to assess the model skills in simulating the short-term weather variability, three different datasets have been analyzed: (i) the UERRA regional reanalysis (Ridal et al, 2017), (ii) the VHR-REA_IT dataset (Raffa et al., 2021) and (iii) meteorological data acquired from networks of weather stations covering the Apulia Region provided by ARIF-Puglia (*Agenzia Regionale Attività Irrigue e Forestali*; http://www.agrometeopuglia.it/osservazioni/mappa-stazioni-meteo). These three datasets have been chosen for their enhanced horizontal resolution over the region of interest, in order to evaluate the model ability to capture the weather features at finer spatial scales.

(i) The UERRA datasets contain regional reanalysis data for Europe from 1961 to 2019 (Ridal et al., 2017). The system is made by two components: the three-dimensional model version called UERRA-HARMONIE at 11 km resolution and the two-dimensional MESCAN-SURFEX surface analysis at 5.5 km resolution. Within the Stream Projects, the MESCAN-SURFEX dataset is investigated, since it provides atmospheric variables above the model topography with an enhanced horizontal resolution every 6-hours. Hereafter the MESCAN-SURFEX dataset is referred to as UERRA. Temperature and wind data are analyzed, while the precipitation is not considered since here it is cumulated from 06:00AM to 06:00PM.

(ii) The VHR-REA_IT Dataset (ERA5-2km) is a Very High-Resolution Dynamical Downscaling of ERA5 over Italy (about 2 km) conducted with the Regional Climate Model COSMO-CLM to provide a detailed dataset of climatological data for the last 30 years (Raffa et al., 2021). ERA5 is the climate reanalysis produced by ECMWF (European Center for Medium-Range Weather Forecasts) and provide hourly data of different atmospheric, land-surface and sea-state parameters at spatial resolution of 31 km (Hersbach et al., 2020).

(iii) In situ measurements acquired from 89 out of 94 weather stations around the Apulia Region in the reference period are used. The five stations not used in this work don't cover the period of interest. Data are kindly provided by the *ARIF Puglia* (www.agrometeopuglia.it) for different meteorological variables: the 2m temperature, the cumulated precipitation and the wind speed with a raw time-sampling that is 1 hour, 10 minutes and 10 minutes respectively.

Different atmospheric fields are here analyzed to estimate the WRF ability to capture the observed weather variability for the period 01/01/2019-05/01/2019: the 2-meter temperature, the cumulated precipitation and the wind speed.

The selected time period has been chosen to characterize the model ability to capture the weather variability in instable conditions. In fact, Italy was affected by humid south-western winds during the reference period, which caused conditions of instability, even intense: the Euro-Atlantic synoptic configuration is characterized by an extensive polar-origin trough extended from Russia to the North African coasts, with two large and robust headlands on the sides. First there is the isolation of a cut-off on Italy in gradual translation to the south and then a deepening and intensification of the trough (Figure 10).

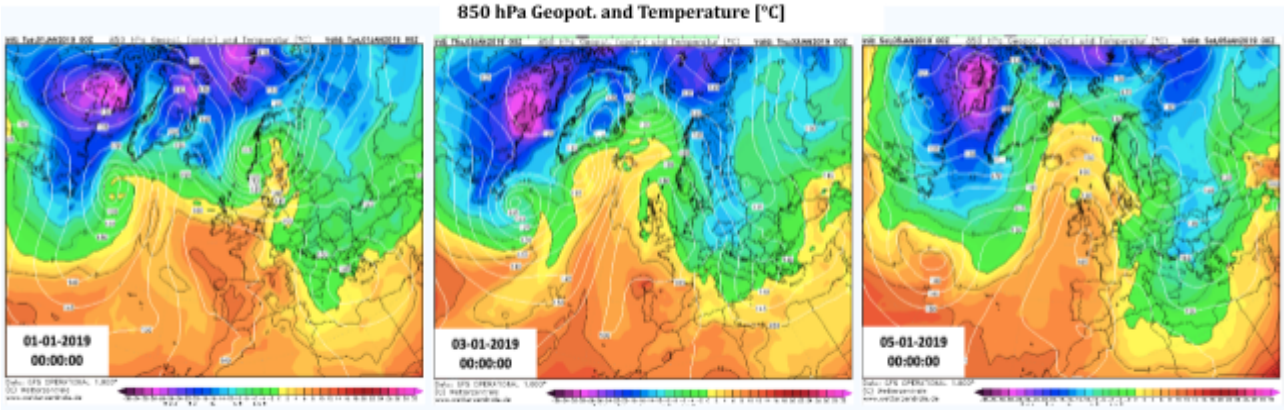


*Figure 10. Synoptic Configuration: 850 hPa Geopotential and Temperature [°C] for 01-01-2019 (left), 03-01-2019 (center) and 05-01-2019 (right); adapted by https://www.wetterzentrale.de/*

The analysis is divided in two parts: (i) first the models are evaluated against ERA5-2km and UERRA, focusing separately on the "land" and "sea" domains (since worse results can be expected over the sea from the two atmospheric model configurations), and then (ii) the model performance is assessed against the weather stations data, taking advantage of the "Nearest Neighbor Search" method (NNS; Vaidya, 1989), which consists in an optimization problem that aims to select the nearest point in a dataset in respect with a specific point. Applying this method to each dataset, the validation has been done comparing the weather stations with three new set of data composed by the respective nearest points in the models and ERA5-2km at hourly frequency. To do so, the 10-minutes wind speed and precipitation data provided by the weather stations have been postprocessed to obtain hourly fields. In this respect, UERRA is not used since it has a 6-hourly time sampling by construction.

To evaluate how the model reproduce the short-term weather variability over land and over sea, a preliminary postprocessing has been done computing the daily data of maximum, mean and minimum temperature ($T_{max}$, $T_{mean}$, $T_{min}$ respectively) on each dataset and the cumulated daily precipitation ($Tp$) in the models and ERA5-2km. The UERRA dataset is not included in the analysis of the precipitation since it is cumulated between 06:00AM and 06:00PM.

Many diagnostics have been used to characterize the model performance: time-average fields, the Probability Density Function (*PDF*), mean values (*mean*) and the Standard Deviation (σ, Eq. 1), where $N$ is the total number of the samples, $X_i$ is the value in the data distribution and μ is the population mean.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(X_i - \mu\right)^2} \qquad \text{(Eq. 1)}$$

The models' performance is preliminarily evaluated investigating the simulated 2m temperature, 10m wind speed and the precipitation against ERA5-2km and UERRA, focusing separately on the "land" and "sea" domains since, by construction, worse results can be expected over the sea from the two atmospheric model configurations.

The 2m temperature, 10m wind speed and the total precipitation fields averaged in the period of interest are shown in Figure 11. The main temperature patterns over the domain are in agreement among the datasets, with stronger differences over the sea. In this respect, weaker amplitudes are found in the ERA5-2km dataset above the sea. Stronger temperatures are found over the Tyrrhenian Sea in UERRA, which are not present in the others. A wide pattern of lower temperatures interests the Campania and Calabria Regions, spanning over the Southern Apennines. Lower temperature spots are also localized over Manfredonia and Bari, surrounded by higher temperatures, with stronger values in Foggia, Brindisi and Lecce. The hint of the different horizontal resolutions is evident in Figure 11 as expected, with finer details in WRF-566m compared to the other datasets, which are exacerbated over the Southern Apennines.

Moreover, from a visual inspection, the simulated wind speed patterns look consistent with those found in UERRA and ERA5-2km, with minima wind speed peaks over the Southern Apennines and Calabria Region, and higher values over the Gargano region, Brindisi and Bari Provinces. More localized spotted wind peaks are found over the Calabria e Campania regions, which are less evident in the other datasets. The mean wind field over the Ionian Sea is similar between

WRF-566m and WRF-2km, both in terms of shape, values and wind speed gradient along the coasts, while WRF-6km, ERA5-2km and UERRA have similar wind patterns over the Ionian Sea, with one-only warmer branch confined in the Taranto Gulf. Similar wind patterns are also found in the Adriatic Sea among the datasets, while weaker wind amplitudes are captured by the models in the Tyrrhenian Sea compared to UERRA and ERA5-2km. Nevertheless, there are differences among the wind mean patterns, the datasets are consistent between each other.

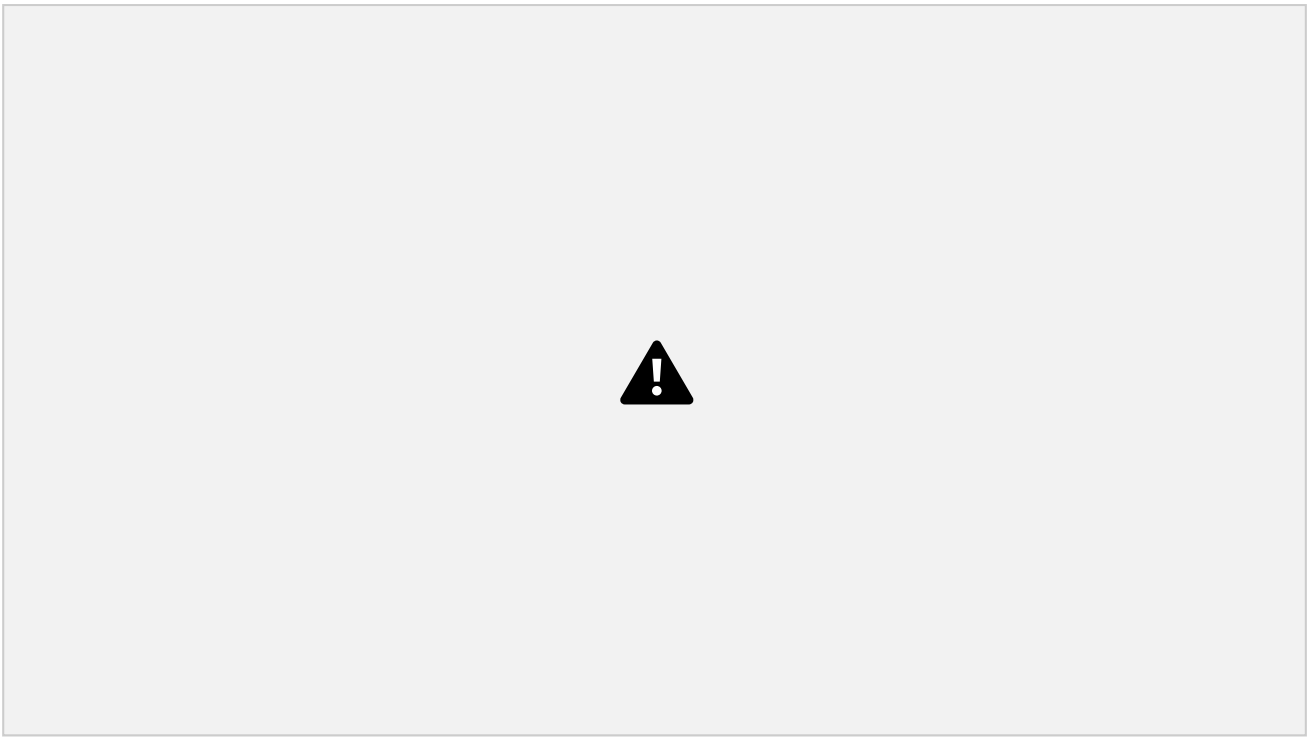

*Figure 11. Upper panel: Maps of time-mean hourly 2m-temperature for WRF-6km, WRF-2km, WRF-566m, ERA5-2km and UERRA-5.5km (from left to right). Center panel: Maps of time-mean hourly 10m-wind speed for WRF-6km, WRF-2km, WRF-566m, ERA5-2km and UERRA-5.5km (from left to right). Bottom panel: Maps of cumulated 5-days precipitation for WRF-6km, WRF-2km, WRF-566m and ERA5-2km (from left to right).*

In contrast, bigger differences are found in the precipitation pattern comparing ERA5-2km and the models. UERRA is not included in the precipitation analysis as previously mentioned. WRF-566m and WRF-2km present similar features, while WRF-6km captures a much higher cumulated precipitation in 5-days with respect to the finer resolution. Nevertheless, even if the total precipitation over the 5days is overestimated by each simulation, the precipitation cores are better geolocalized by WRF-566m and WRF-2km.

In order to better capture the differences and similarities among the datasets, the PDFs of each variable and dataset are investigated below, separating the analyses in two categories: (i) Land-only and (ii) Sea-only.

The PDFs of the simulated $T_{max}$, $T_{mean}$, $T_{min}$ are consistent with the respective PDFs of the reference datasets over land and sea (Figure 12). However, a translation of the Sea-only PDFs of UERRA, WRF-6km, WRF-2km and WRF-566m is observed with respect to ERA5-2km probably due to the absence of data in the easternmost sector of the domain in the latter. WRF-566m and WRF-2km have similar mean and σ values for $T_{max}$, $T_{mean}$, $T_{min}$ over land, while WRF-6km maintains a σ comparable to the other model configurations for each variable, but its mean value is shifted toward lower magnitude, in agreement with the reanalyses' mean value. While, over the sea, WRF-566m and WRF-2km overlap each other and WRF-6km maintain their same range of value, with higher probabilities associated to higher magnitudes. Over the sea, the models are coherent with the UERRA distribution, while ERA5-2km is shifted toward lower values. It is important to note, that the ERA5-2km dataset does not cover the entire region of interest (Figure 11), and then the results over the entire domain would suffer for this gap over the sea.

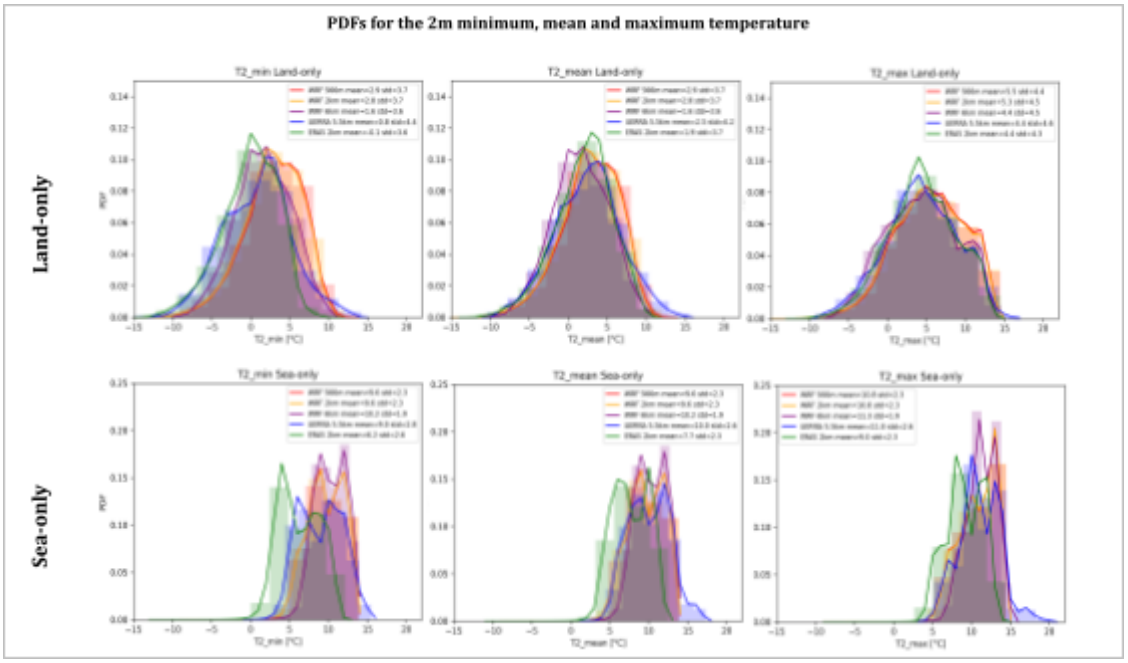


*Figure 12. PDFs and related histograms for the minimum (left), mean (center) and maximum (right) hourly 2m-temperature [°C] over land (top panel) and over sea (bottom panel) in WRF-566m (red), WRF-2km (yellow), WRF-6km (purple), UERRA (blue) and ERA5-2km (green). Inside the legend, the mean and standard deviation of each distribution are shown.*

Analyzing the wind PDFs (Figure 13), by visual inspection each model configuration is in good agreement with the respective PDFs in ERA5-2km and UERRA over land, while over the sea only WRF-2km and WRF-566m are consistent with the wind data of the reference datasets. WRF-566m and WRF-2km, both over land and sea, have mean values and σ in good agreement with those found in the reanalyses, while WRF-6km over land is more shifted toward higher values and has higher probabilities centered around the mean (~9.6 m/s).

Investigating the PDFs for the precipitation, although there are substantial differences in the simulated precipitation pattern over the entire domain with respect to ERA5-2km (Figure 11), the PDFs are consistent between each other. However, the intensity of the precipitation pattern tends to be overestimated by the models but is better geolocated by WRF-2km and WRF-566m than by WRF-6km. (Figure 14).

Nevertheless, in order to assess the models' performance, in situ measurements acquired from 89 weather stations around the Apulia Region in the reference period are used. Data are kindly provided by ARIF Puglia for different meteorological variables: the 2m temperature, the cumulated precipitation and the wind speed with a raw time-sampling that is 1 hour, 10 minutes and 10 minutes respectively. The wind speed has been averaged each hour and the precipitation has been cumulated hourly, in order to have the same time-sampling.

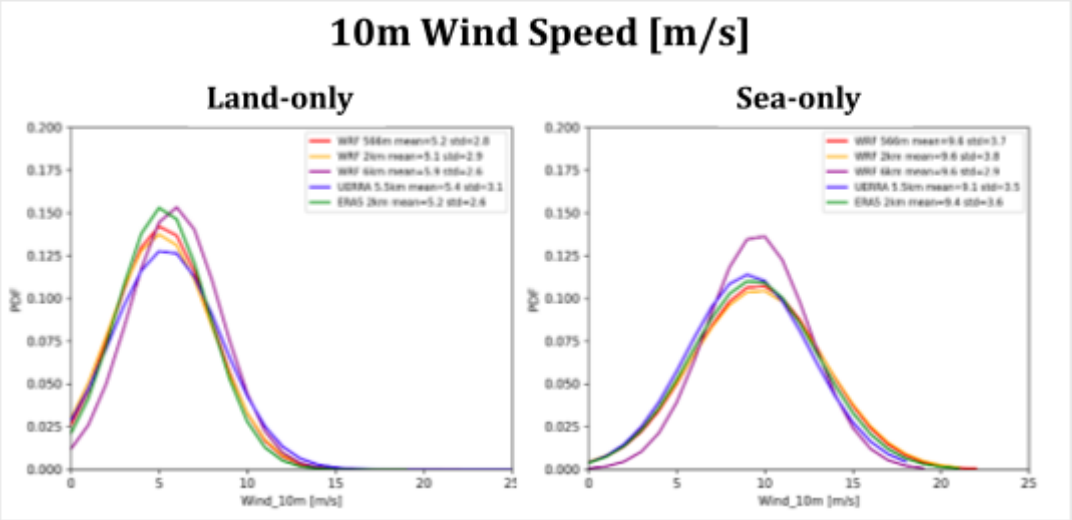


*Figure 13. PDFs for the hourly 10m-wind speed [m/s] over land (left) and over sea (right) in WRF-566m (red line), WRF-2km (yellow line), WRF-6km (purple line), UERRA (blue line) and ERA5-2km (green line). Inside the legend, the mean and standard deviation of each distribution are shown.*

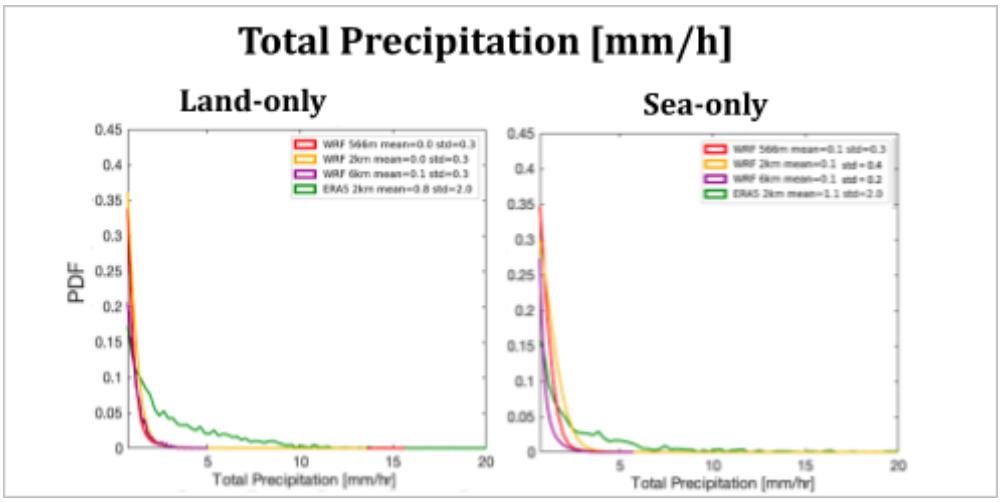


Figure 14. PDFs for the hourly total precipitation [mm/h] over land (left) and over sea (right) in WRF-566m (red line), WRF-2km (yellow line), WRF-6km (purple line) and ERA5-2km (green line). *Inside the legend, the mean and standard deviation of each distribution are shown.*

To make a fair comparison between the WRF 566m, WRF 2km, WRF-6km and ERA5 2km with the weather stations, we take advantage of a method which is widely used in literature known as "Nearest neighbor search" (NNS; Vaidya, 1989). The nearest grid point (which minimizes the distance) for each station is therefore selected in the gridded datasets. The hourly data from WRF-6km, WRF-2km, WRF-566m and ERA5-2km are evaluated against the weather-stations. The UERRA dataset is not included since it has a 6-hourly time-sampling.

Figure 15 shows the PDFs for in situ data, ERA5-2km and the three model configurations for the hourly 2m temperature, wind intensity and total precipitation.

Analyzing the 2m temperature, it is less sensitive to the variation of the model horizontal resolution. Similar PDFs for the mean temperature at hourly frequency are found for WRF-2km and WRF-566m, with higher probabilities shifted toward higher values compared to the other datasets. Nevertheless, the range of values, the mean and σ of the distributions are consistent with each other. While, investigating the 10m wind field, progressive improvements in wind reproducibility are achieved by switching from a resolution of 6km to a resolution of 2km and finally 566 m. Finally, the resolution increase from ~6km to 2km greatly improves the precipitation simulation. On the contrary, there is no noticeable variation in performance from ~2km resolution to ~566m.

Good performance is achieved by WRF-2km and WRF-566m, whose PDFs tend to overlap with the station data.

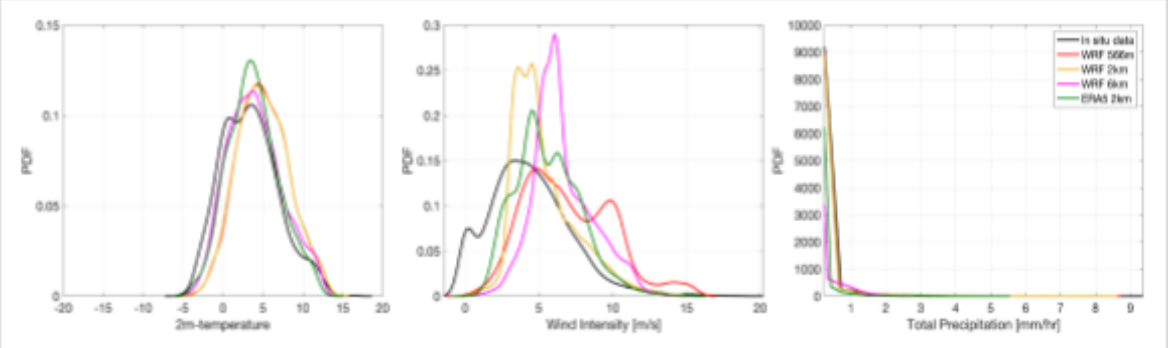


*Figure 15. PDFs for the hourly 2m temperature [°C] (left), 10m wind intensity [m/s] (center) and the cumulated hourly precipitation [mm/h] (right) for the in-situ data (black line), and the respective datasets defined through the NNS method in ERA5-2km (green line), WRF-6km (purple line), WRF-2km (yellow line) and WRF-566m (red line).*

In conclusion, improvements in the model performance are evident switching from 6km to 2km (without perturbing the rest of the model setup), while changing the horizontal resolution from 2km to 566m, the WRF model better reproduces the wind speed, while the temperature and precipitation are less sensitive to the enhanced resolution. Therefore, both WRF-2km and WRF-566m configurations are considered suitable for the project purposes.

## 2.2 Coastal and inland flooding for Peschici-Manfredonia ond Ofanto river pilots

The area of Peschici is the less exposed to the flooding of those considered in this project. This is probably due to its geomorphological setting, mainly represented by and alternance of high promontories and small pocket beach, and to the absence of coastal lagoon. In this area, as reported in the Fig. 16, the medium flooding, referred to the event of November 2019, could be of about 28-30m, with a maximum value of 35m estimated along the pocket beaches. The rocky promontories should be not interested by flooding.

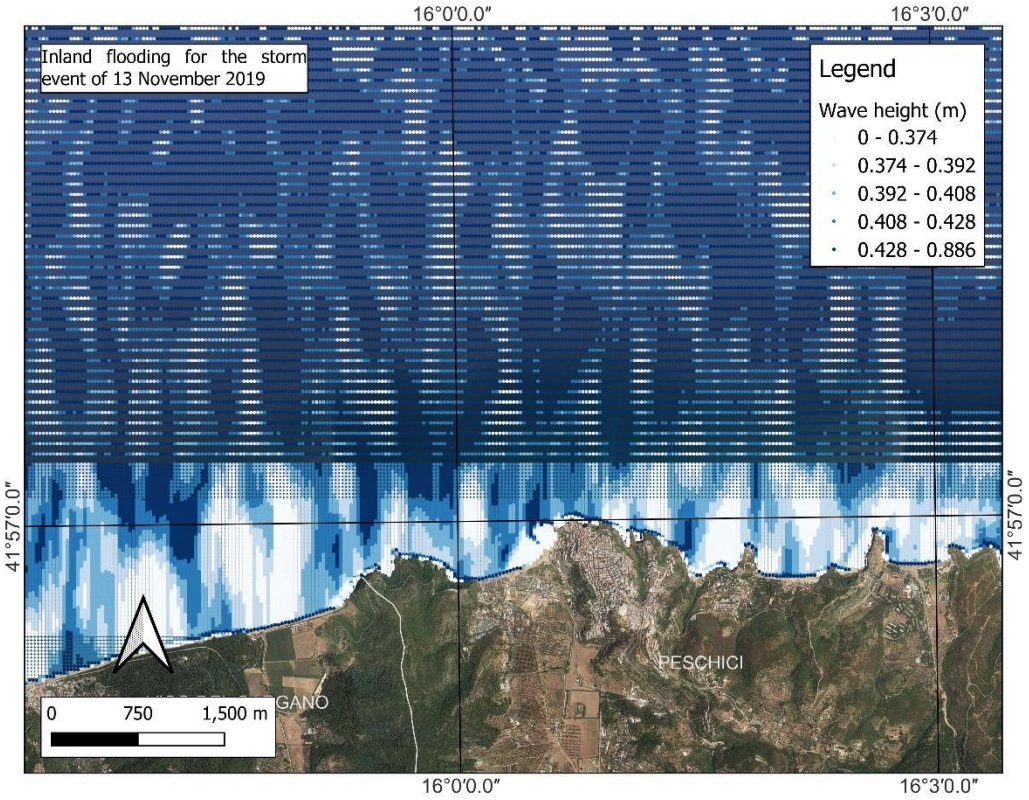


*Figure 16 – Flooding in Peschici for the event of November 2019*

The area of Manfredonia is the most exposed to the flooding. As shown in Fig. 17, this area would be heavily impacted by the effects of storms, probably cause of the presence of large coastal lagoons and also for its geomorphological setting characterized by low lying beaches. In this sector, flooding related to the event of November 2019, could reach a maximum value of 1,5 km in proximity of the coastal lagoon located close to Margherita di Savoia.

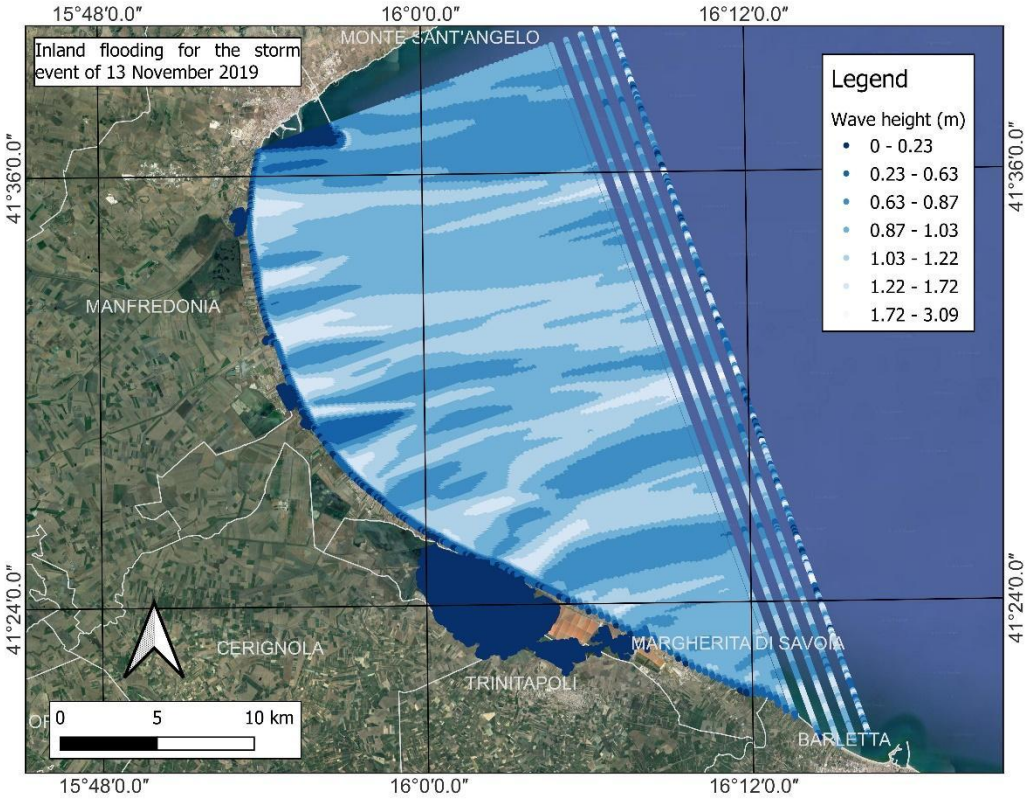


*Figure 17 - Flooding in Manfredonia – Margherita di Savoia for the event of November 2019*

As discussed in the previous chapter, flooding along the coastal sector close to the river mouth of Ofanto will be the result of river inundation and coastal flooding related to the storm impact. In this case, as shown in Fig. 18, the increased water level, induced by storm impact along the river mouth, generate a river inundation up to 500m landward. The combined river (with HEC-RAS) and ocean (with XBEACH) flooding is shown in Figure

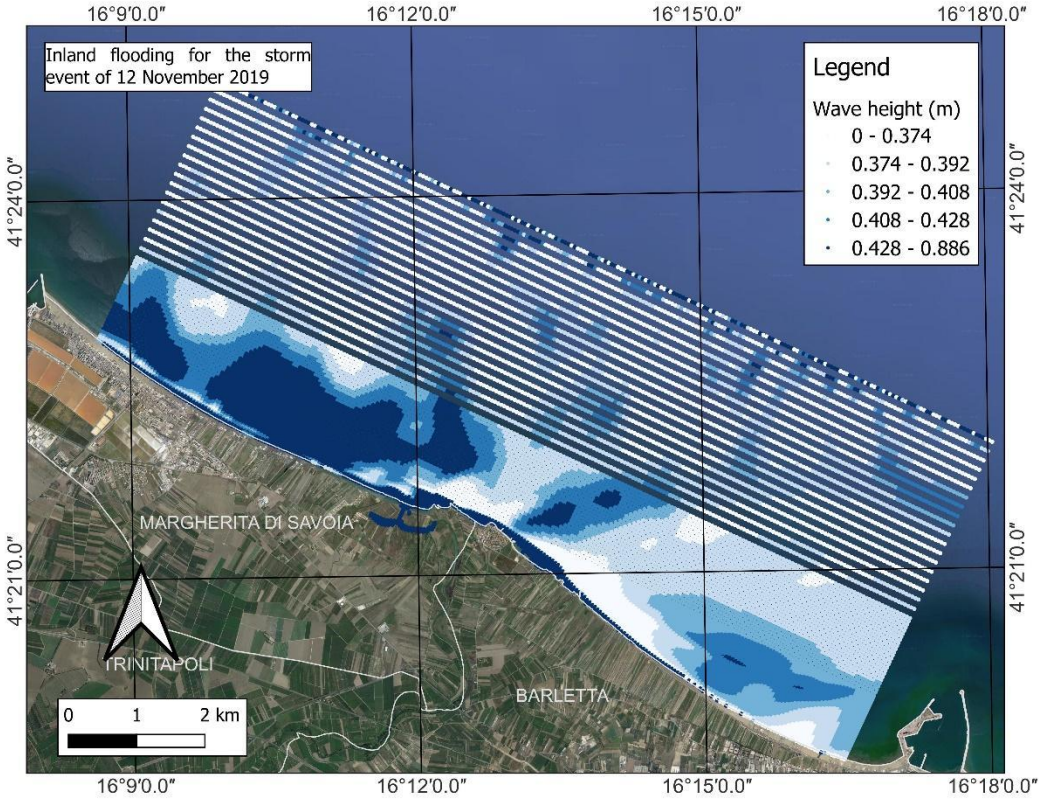


*Figure 18 – Ocean flooding in Ofanto for the event of November 2019*

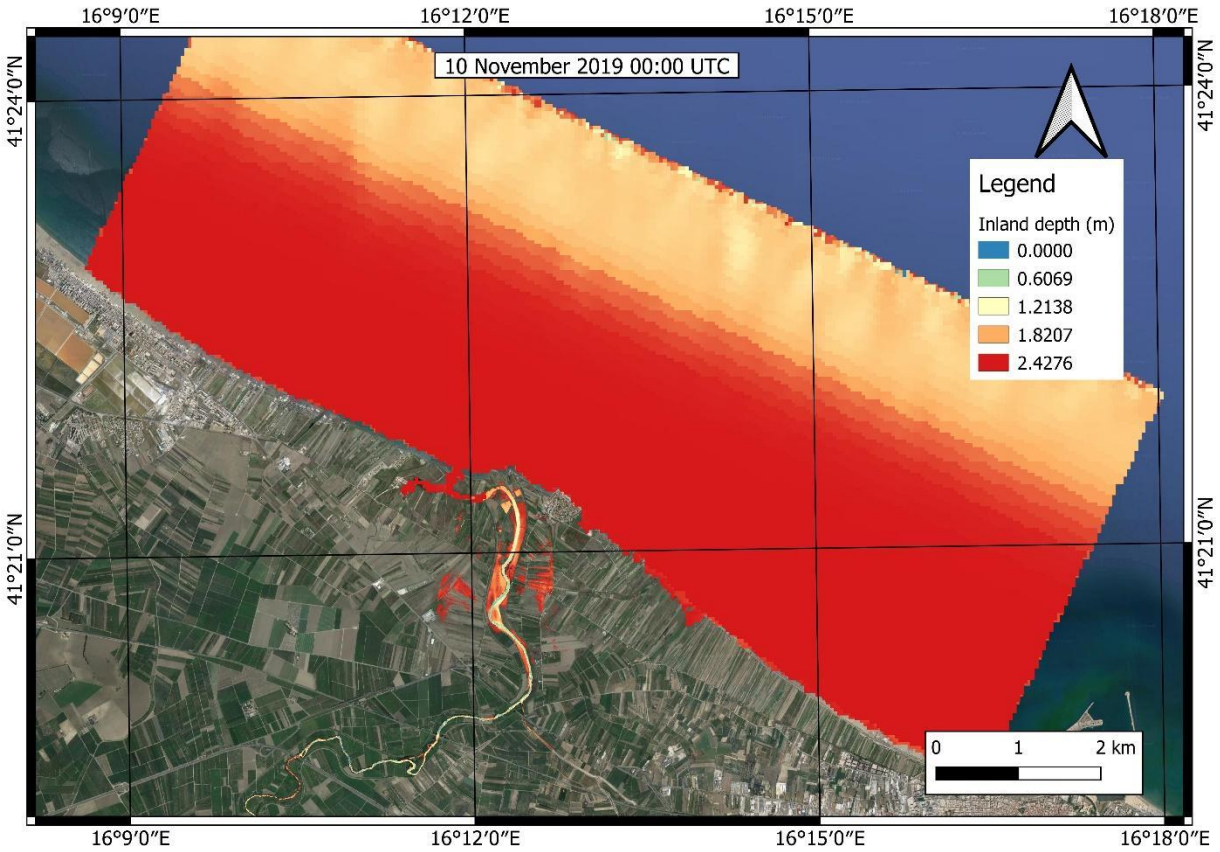


*Figure 19 – Combined river and ocean flooding in Ofanto for the event of November 2019*

## 2.3 Coastal erosion for the Lecce-Torchiarolo pilot

Diachronic analysis of coastline between Torchiarolo and Lecce highlights a general erosional trend. As shown in Fig. 20, maximum rates of shoreline regression are present in the south areas close to Lecce, with values of about -1,3m/y. The central part is characterized by an almost stable coastline that has not shown linear displacements in the last decade, while the northern sector is marked by erosional rates up to -0,9m/y.

As consequence the area most exposed to submersion in 2050 and 2100 are the southern one, as shown in Fig. 21.

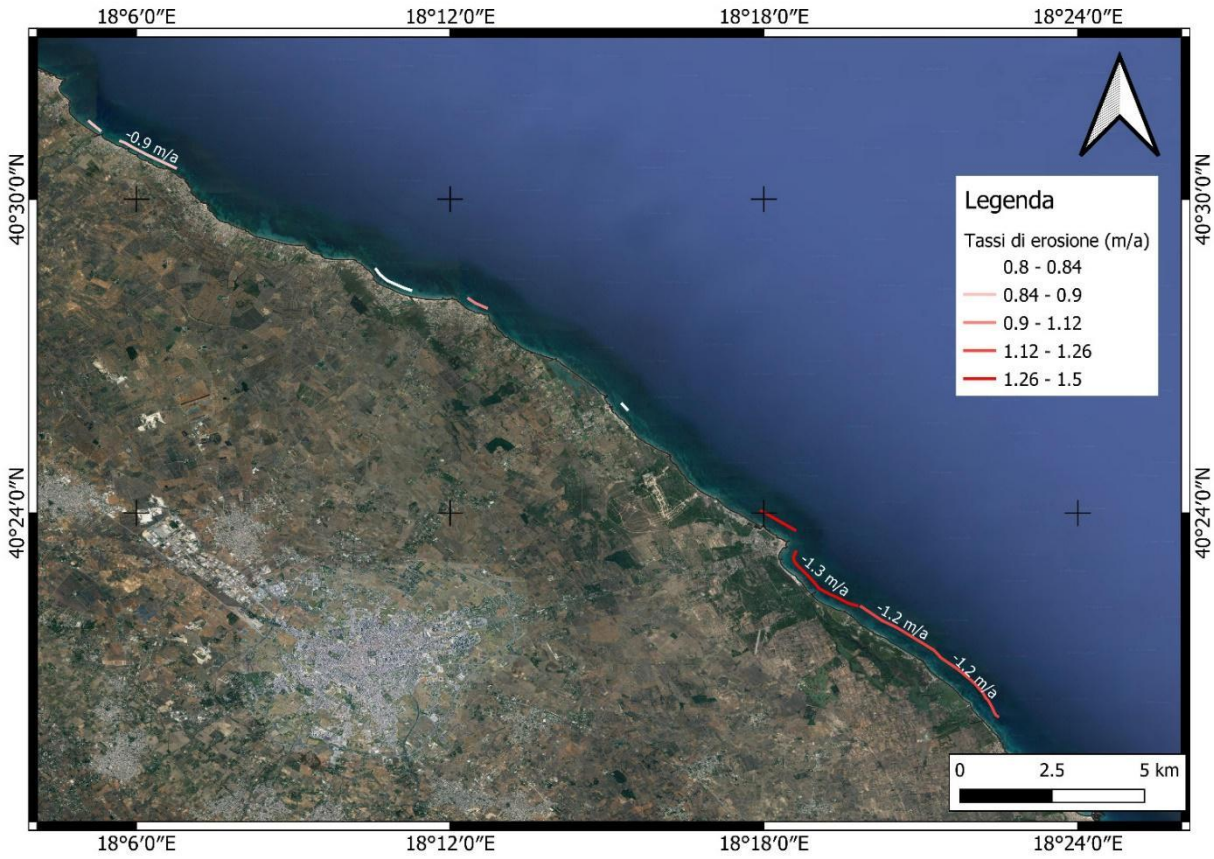


*Figure 20 - Diachronic analysis of coastline between Torchiarolo and Lecce*


*Figure 21 - Submersion scenario in 2100*

# 3 The Operational Forecasting Chains

The operational chain built for the ocean models in the STREAM framework is capable every day to provide 3-days of forecasts with hourly frequency. The outputs have been delivered both in unstructured (native) format and regridded format to serve downstream services (e.g. tool for visualization) and applications. The ocean fields released on the daily basis are 3D currents, temperature and salinity, and sea level.

The operational forecasting methodology is based on the high-resolution model re-initialization every day, similar to the short-term limited-area atmospheric modelling practice (Mesinger et al., 1988). The re-initialization strategy allows exploiting the high-quality systematic fields of parent model Med CMEMS (provided by data assimilation), which supplies operational forecasting products in the framework of CMEMS service. This type of approach has been adopted by other forecasting systems downscaled from Med-CMEMS, as reported in Napolitano et al. (2016).

Here we describe the daily forecast cycle for the model configuration. With j as the current day, the initializing fields (taken from the Med-CMS simulation products) of forecast procedure are imposed at 12:00 of 3 days backward (j − 3) as the instantaneous fields. The forecasting run exploits the Med-CMS simulations (for j − 2 and j − 1) and the Med-CMS forecasts (for j +1, j +2 and j +3) at the lateral open boundary, while the surface boundary conditions run over the ECMWF analysis (j − 3, j − 2, j − 1) and forecasts (j + 1, j + 2, j + 3). The forecast is prepared and runs automatically. The operational chain is activated as soon as the atmospheric forcings are available. The technical procedures through scripts and codes for computing the forecast fields can be summarized in the pre- processing of input data, model run and post-processing of the output model.

The main informatic procedures are based on workflow manager and python-based scripts. Since the operational chain was designed with a scalability and reusability approach, the modules used there were relocated and further developed to manage the tasks required by the full operational cycle. The main informatic aspects about the operational chains are:

- Managed workflow (python Luigi based, https://github.com/spotify/luigi)
  - No direct usage of cluster's scheduler
  - Already adopted in operational chains at CMCC
- No bash-based automations
  - Easier to maintain
  - Uncoupled function
  - Optimized I/O
- Consolidated production framework (e.g. Spotify)
- Pre-and-post processing intrinsically parallelized
  - Faster execution time for large dataset
  - Easy recover in case of failures
- Front-end included

The managed workflow vs. the flat automation design, allowed us to design and develop a more structured and modular system able to manage every aspect of the operational routines (unexpected situations included) and to serve as foundation for more complex and higher-level operations. The indirect usage of the cluster's scheduler makes it possible to choose amongst different running models, each fitting different cluster conditions and unexpected situations.

The bash automation was kept to a minimum, preferring the use of a general-purpose language to describe the operational tasks. This led to a more structured codebase, easier to maintain with a clear separation of concerns between modules.

The use of command-line based data manipulation tools has been eliminated in favor of the direct use of programmatic libraries that are at the foundation of the aforementioned commands. This allowed a more rationale usage of manipulation routines and a considerable reduction of disk usage for intermediate results storing.

The workflow manager used to run operatively was Spotify's Luigi, a mature piece of software, already used in production by a large number of big IT industries. Luigi is a Free Open Source software too, which allow us to inspect its internals, make changes freely and contribute to the project.

The operational chain is constituted mainly by three large steps, each of which dealing with Preprocessing, model running/monitoring, Post processing. Being the pre and post processing are composed almost entirely by a large number of interpolations, there is a lot of room for parallelization. The way the tasks are designed, allowed us to exploit this intrinsic parallelization almost for free by splitting the large interpolations in blocks and by choosing a running model with different concurrent workers.

The first concern of the operational staff is to understand the state of the system in case of malfunctioning.

The Luigi's scheduler used in the operational chains integrates a web based front end showing in real time the status of every workflow running. In the schematic representation, each dot indicates a single task need to be accomplished in order to advance the workflow. The color of that dot denotes visually the state of the task, being: yellow when pending, green when done, blue when running, red when failed.

When hovering, each task shows the relevant details useful to precisely identify it and, in case of failure, details about the error are shown too.

The representation gives a consistent picture about the system state and simplifies the identification of problems.

The forecasting data for the AdriFs and SOAP systems are displayed at the website https://adri.cmcc.it/ and https://soap.oceanity.eu/. Some examples of visualization are reported in Figure 22.

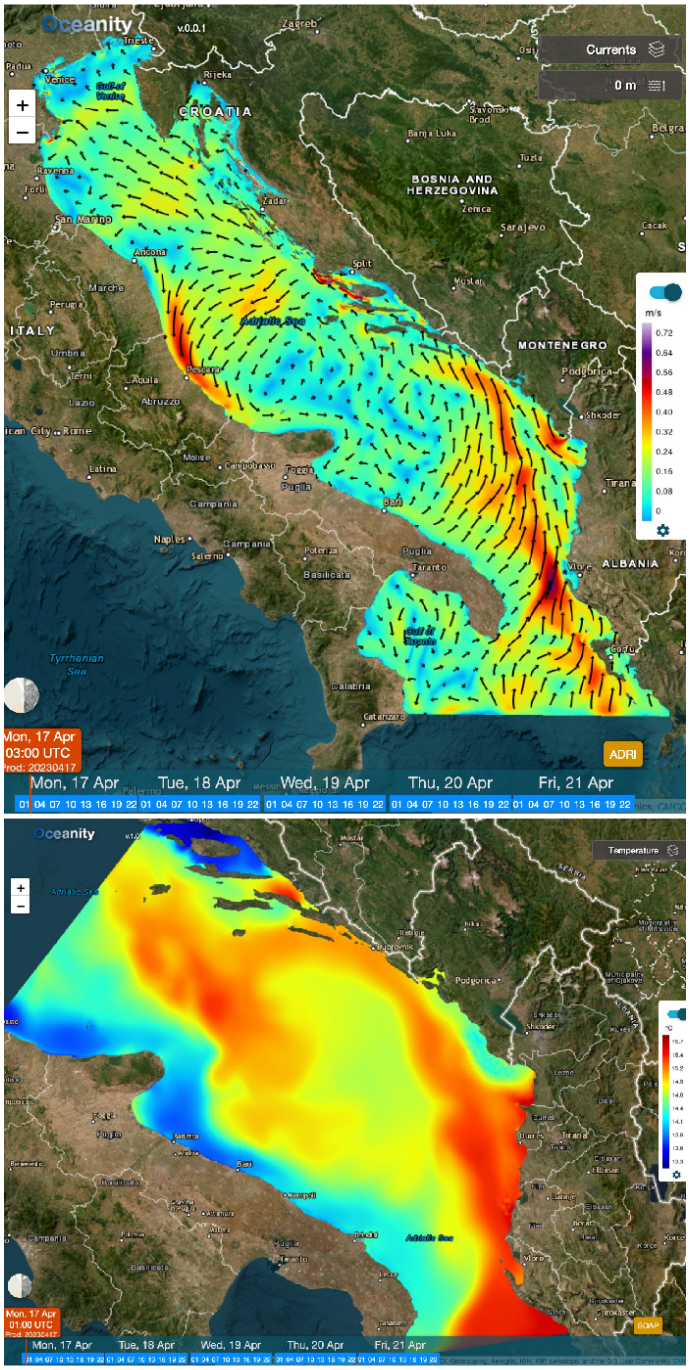


Figure 22. Examples of visualization of circulation for AdriFs and temperature for SOAP systems

# 4 REFERENCES

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J-N. (2023): ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), DOI: 10.24381/cds.adbb2d47

Raffa, M.; Reder, A.; Adinolfi, M.; Mercogliano, P. (2021) A Comparison between One-Step and Two-Step Nesting Strategy in the Dynamical Downscaling of Regional Climate Model COSMO-CLM at 2.2 km Driven by ERA5 Reanalysis. Atmosphere 2021, 12, 260. https://doi.org/10.3390/atmos12020260

Ridal M., E. Olsson, P. Unden, K. Zimmermann and A. Ohlsson (2017) HARMONIE reanalysis report of results and dataset. UERRA deliverable D2.7.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., … Huang, X. -yu. (2019). A Description of the Advanced Research WRF Model Version 4.1 (No. NCAR/TN-556+STR). doi:10.5065/1dfh-6p97

Vaidya, P. M. (1989). "An O(n log n) Algorithm for the All-Nearest-Neighbors Problem". Discrete and Computational Geometry. 4 (1): 101–115. doi:10.1007/BF02187718