

# **Alma Mater Studiorum - Università di Bologna**

**DOTTORATO DI RICERCA IN GEOFISICA**

**Ciclo XXXIII**

**Settore Concorsuale: 04/A4**

**Settore Scientifico Disciplinare: GEO/12**

## **Seamless Modelling of the Southern European Seas with an optimized MPI-based Unstructured-Grid Model**

**Presentata da: Ivano Barletta**

### **Coordinatore Dottorato**

Prof. Nadia Pinardi

### **Supervisore**

Dr. Ivan Federico

### **Co-Supervisor**

Prof. Nadia Pinardi

Dr. Giorgia Verri

**Esame finale anno 2021**



“... man had always assumed that he was more intelligent than dolphins because he had achieved so much —the wheel, New York, wars and so on— whilst all the dolphins had ever done was muck about in the water having a good time. But conversely, the dolphins had always believed that they were far more intelligent than man for precisely the same reasons.”

# Abstract

This interdisciplinary work combines computational science and physical oceanography.

The parallelization of the SHYFEM circulation model and its application for simulating the system of inter-connected basins in the Southern European Seas (SES) are the main areas of focus.

The SES system is made up of different water bodies that can generally be regarded as separate and having independent dynamics. Most modelling studies treat these water bodies as standalone. The recent studies that model the system as a whole typically focus only on a particular process or a short time-scale.

Representing the dynamics generated by narrow straits is difficult in regular grid models, and so those based on an unstructured grid represent a natural solution, as they are the most suitable for representing the dynamics of a geometrically complex system such as the SES.

Modelling aimed at a complete description of such complex physical systems incurs a considerable computational cost, which is directly affected by the size of the physical domain, the model resolution and the length of the simulation. Modern simulations involve numerical codes of circulation models that exploit the ongoing increases in computer power. The code of state-of-the-art models can be executed in parallel through high performance computing facilities, thus enabling the simulations to be completed quickly and efficiently.

In this context, the parallelization of the hydrostatic code of the SHYFEM unstructured grid model is the first task, which has been successfully applied in several coastal and regional configurations, and is a versatile code that includes several tools for environmental studies. The task of parallelization of SHYFEM includes an assessment of its reproducibility and parallel performances on a realistic test case.

In the second task we address the applicability of this parallel unstructured model for the more demanding SES system, demonstrating that this parallel code can effectively simulate the SES system from a physical point of view, seamlessly representing the physical domain in the long term. The SES system is simulated for one year with a fully baroclinic run.

The results obtained are assessed qualitatively and by comparison with observations where possible.

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	General Features of The SES System . . . . .	2
1.2	Dynamics Of The Mediterranean . . . . .	5
1.2.1	Sub-basin scale Circulation . . . . .	8
1.3	Dynamics of The Black Sea . . . . .	11
1.4	Dynamics of the Turkish Strait System . . . . .	14
1.5	State of the art of the SES modeling . . . . .	16
1.5.1	Modeling of the Mediterranean Sea . . . . .	17
1.5.2	Modeling of the Black Sea . . . . .	20
1.5.3	Modeling of TSS . . . . .	23
1.6	Unstructured Grid Models . . . . .	25
1.7	Parallel Ocean Numerical Models . . . . .	29
1.8	Aims and Structure of the Thesis . . . . .	34
<b>2</b>	<b>SHYFEM Description</b>	<b>36</b>
2.1	The SHYFEM modelling System . . . . .	36
2.1.1	The SHYFEM Computation Grid . . . . .	38
2.1.2	Model Equations . . . . .	39
2.1.3	Sub-grid scale Physics - Turbulence Closure . . . . .	41
2.1.4	Surface/Bottom boundary Conditions . . . . .	44
2.1.5	Lateral Open Boundary Conditions . . . . .	47
2.2	Numerics - SHYFEM Solution Algorithm . . . . .	48
2.2.1	Advancement of Transports . . . . .	50
2.2.2	Viscosity Operators . . . . .	51
2.2.3	The Barotropic Equation . . . . .	53

2.2.4	STEP 2 - Finalize Hydrodynamics . . . . .	54
2.2.5	Vertical Velocities . . . . .	54
2.2.6	Advection-Diffusion . . . . .	55
<b>3</b>	<b>Optimization and MPI Parallelization of SHYFEM</b>	<b>57</b>
3.1	The Need of SpeedUp Simulations . . . . .	58
3.2	Dependencies Between Physical Variables on the Numerical Grid .	60
3.3	Parallelization Approach . . . . .	62
3.3.1	Domain Decomposition . . . . .	63
3.3.2	Domain Decomposition Strategy . . . . .	63
3.3.3	The Domain Decomposition Algorithm . . . . .	65
3.3.4	MPI Directives . . . . .	66
3.3.5	Free Surface Solver . . . . .	69
3.3.6	I/O Management in MPI . . . . .	75
3.4	Test Case for Validation of MPI Code . . . . .	76
3.5	Results . . . . .	78
3.5.1	Performance Assessment . . . . .	80
3.6	Extension of Restart Capabilities . . . . .	86
3.7	Model Diagnostics . . . . .	86
<b>4</b>	<b>The Southern European Seas Modelling System</b>	<b>88</b>
4.1	The SES System and The Domain Configuration . . . . .	88
4.2	Experimental Setup . . . . .	91
4.3	Forcings . . . . .	92
4.3.1	Surface and Atlantic Boundaries . . . . .	92
4.3.2	Rivers . . . . .	95
4.4	Model Integration and Performances . . . . .	97
4.5	Validation of Results . . . . .	97
4.5.1	Strait Dynamics . . . . .	98
4.5.2	Sea Surface Temperature . . . . .	102
4.5.3	Circulation Patterns . . . . .	105
4.5.4	Thermohaline Structures . . . . .	108
<b>5</b>	<b>Conclusions and Future Work</b>	<b>119</b>

<i>CONTENTS</i>	1
<b>A River Dataset</b>	<b>124</b>
<b>B Lock Exchange Experiment</b>	<b>127</b>
<b>Acronyms</b>	<b>129</b>
<b>Bibliography</b>	<b>134</b>
<b>Acknowledgements</b>	<b>151</b>

# Chapter 1

## Introduction

### Contents

---

<b>1.1</b>	<b>General Features of The SES System . . . . .</b>	<b>2</b>
<b>1.2</b>	<b>Dynamics Of The Mediterranean . . . . .</b>	<b>5</b>
1.2.1	Sub-basin scale Circulation . . . . .	8
<b>1.3</b>	<b>Dynamics of The Black Sea . . . . .</b>	<b>11</b>
<b>1.4</b>	<b>Dynamics of the Turkish Strait System . . . . .</b>	<b>14</b>
<b>1.5</b>	<b>State of the art of the SES modeling . . . . .</b>	<b>16</b>
1.5.1	Modeling of the Mediterranean Sea . . . . .	17
1.5.2	Modeling of the Black Sea . . . . .	20
1.5.3	Modeling of TSS . . . . .	23
<b>1.6</b>	<b>Unstructured Grid Models . . . . .</b>	<b>25</b>
<b>1.7</b>	<b>Parallel Ocean Numerical Models . . . . .</b>	<b>29</b>
<b>1.8</b>	<b>Aims and Structure of the Thesis . . . . .</b>	<b>34</b>

---

### 1.1 General Features of The SES System

The system of interconnected basins that we call in this work Southern European Seas ( SES for brevity) covers less than 1% of the total ocean surface but has always raised scientific interest because of its unique characteristics.

The system encompasses the Mediterranean, being the largest concentration basin in the world, and the Black Sea, the largest brackish water basin.

Both the basins are semi-enclosed and connected with other marginal seas, the Azov and the Marmara sea, giving the system a connotation of high complexity, from both geometrical and dynamical point of view.

The SES system can be seen as a seamless cascade of basins that starting from the Azov Sea to the west is connected to the Atlantic Ocean. The Sea of Azov is an enclosed sea located in eastern Europe and is among the smallest and shallowest seas in the world having a surface of 39000 km<sup>2</sup> and an average depth of 8 meters with a maximum of 14. its Hydrological characteristics and the general circulation are governed by local winds, river runoff (95% of which are due to Don and Kuban rivers giving an annual average of nearly 40 km<sup>3</sup>/year), and water exchange with the Black Sea (Zavialov et al. 2020), the latter taking place in the Kerch Strait having a width of 5 km and an average depth of 55 meters. The Black Sea is a semi-enclosed basin located amid Eastern Europe and the Caucasus, having a surface of 436000 km<sup>2</sup> and a mean/maximum depth of 1250/2120 meters. The topography consists mainly of a wide shelf area in the North West of the basin, and an abyssal plain of nearly constant bathymetry in the center.

The main mechanisms driving the Black Sea circulation are wind stress, creating a cyclonic rim current, and the river runoff. The contribution of the precipitation over the Black Sea basin plus the riverine discharge is not balanced by the evaporation in the long period, with the result of a net outflow of brackish water towards the Marmara Sea through the Bosphorus Strait (Jarosz et al. 2011).

The Bosphorus Strait is a 30 km long narrow and jagged channel whose minimum width is around 700 meters and the deepest point of 110 meters. Estimates of the net flow of Black Sea waters to the Marmara Sea range from 200 km<sup>3</sup>/year to 300 km<sup>3</sup>/year (Ünlülata et al. 1990a) (Peneva et al. 2001) (Kara, Wallcraft, Hurlburt, and E.V. Stanev 2008).

The Sea of Marmara can be considered as a mixing zone between the Black Sea and the Mediterranean, connecting the two major water bodies through the Bosphorus Strait and the Dardanelles. The basin is relatively small if compared to the surrounding ones, covering a surface of 11500 km<sup>2</sup> and showing a rather complex topography. The Marmara has an east-west elongated shape with shallow bathymetry with the exception of the three topographic depressions in the

northern part that are seaward extensions of the well known North Anatolian Fault Zone spanning the Anatolian peninsula (Beşiktepe et al. 1994) with the deepest point at 1370 meters. Compared to the northern coastline, the southern is more complex with islands and peninsulas. The mean upper-layer circulation which is largely driven by the southward flowing Bosphorus jet is anticyclonic. The river runoff into this sea is relatively small so the net flows through the Bosphorus nearly equals the flow towards the Mediterranean through the Dardanelles Strait.

The Dardanelles Strait (also referred to as the Canakkale Strait) connects the Sea of Marmara to the Mediterranean. The Dardanelles Strait is about 61 km long with an averaged depth of 55 m. The strait is rather narrow with a width varying between 1.2 km and 7 km (Ünlülata et al. 1990a). The narrowest sections of the Dardanelles Strait are located in a sharp bend called the Nara Pass and close to the site of Canakkale. The two-layer flow, with brackish Black Sea waters moving westward in the upper layer and more saline Aegean waters flowing roughly eastward underneath, puts in communication the Marmara Sea with the Aegean Sea and, hence, with the Mediterranean. The Mediterranean is the largest of the semi-enclosed seas in the SES system, with a surface of 2.5 million km<sup>2</sup>. It is composed of two major sub-basins, the Western Mediterranean and the East Mediterranean, that communicate through the Strait of Sicily (140 km wide) and the narrower Strait of Messina (around 3 km wide), each one connected to a marginal sea, the Adriatic and the Aegean respectively.

With the exception of the Western Mediterranean, which shows relatively regular morphology and nearly constant bathymetry in its west part, the Mediterranean and its marginal seas are marked out by complex coastline features and bathymetry that play a role in its dynamics. The Mediterranean has several shelf areas, such as the Gulf of Gabes and the North of Adriatic, with average depths around 30 and 50 meters respectively. The abyssal plain of the Ionian is the deepest of the SES system, with average values of nearly 3000 meters and a maximum value of 5270 meters at Calypso Deep, in the Hellenic Trench.

Different from the Black Sea, The Mediterranean is an anti-estuarine or concentration basin, being the sum of rain and river runoff smaller than the evaporation. The Mediterranean region is marked out by a neat heat loss of 7 W/m<sup>2</sup> and net water flux of -0.64 m/yr as climatological mean in the estimate of (Pettenuzzo et al. 2010). The water lost by evaporation is compensated by the net inflow of

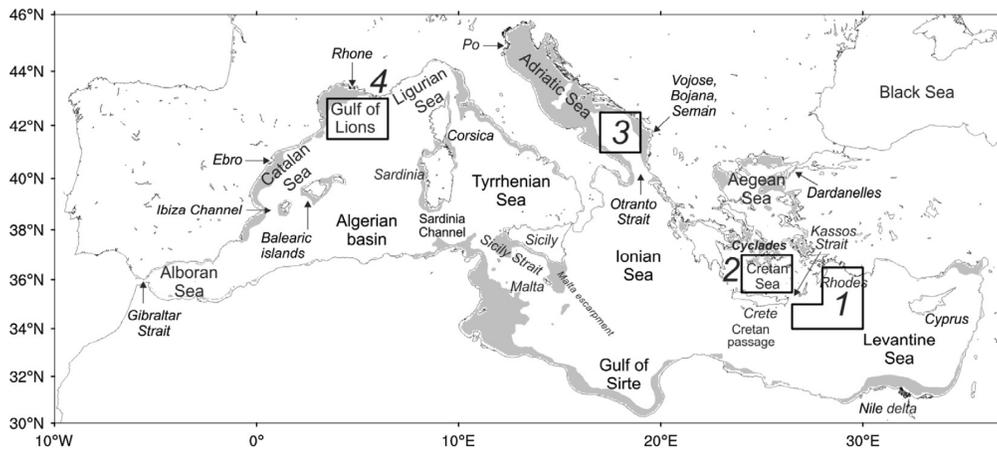


Figure 1.1: Geography of the Mediterranean Sea with labelling of relevant sub-areas. Grey shades indicate shelf areas. Numbered boxes refer to the deep water formation regions.

Atlantic water through the Strait of Gibraltar.

## 1.2 Dynamics Of The Mediterranean

Among the SES system sub-basins, the Mediterranean is the biggest and richer of features that characterize its dynamics. Many processes that are typical of the general circulation of the world ocean develop also in the Mediterranean in an identical or analogous manner, giving it a connotation of an isolated oceanic system (Robinson et al. 2001).

The geography of the entire Mediterranean is shown in figure (1.1). The Mediterranean Sea is composed of two nearly equal size basins, the Western and Eastern Mediterranean (WMED and EMED hereafter), connected by the Strait of Sicily.

The WMED extends from Gibraltar to the west coast of Italian peninsula, encompassing the Alboran sea, the Algerian basin and the Thyrrhenian sea. The Algerian basin is separated from the Thyrrhenian by the Sardinia and Corsica islands. The EMED has a complex geographical distribution: it is composed of two marginal seas in its northern part, the Adriatic and Aegean Seas, connected to the rest of the basin by shallow straits. The Adriatic extends northward between Italy and the Balkans, communicating with the eastern Mediterranean

basin through the Strait of Otranto. The Aegean lies between Greece and Turkey, connected to the eastern basin through the several straits of the Grecian Island arc. The strait of Gibraltar provides connection between the WMED and the Atlantic ocean. It is a system of sills and narrows about 60 km long and 20 km wide, with a minimum width of less than 14 km in the Tarifa narrow section and a minimum depth of 290 m in the Camarinal sill, located west of Tarifa.

The anti-estuarine character of the Mediterranean circulation is driven by the baroclinic flow developing at the Gibraltar strait, where water flows are separated vertically, with warmer and fresher Atlantic Water enter the Mediterranean in the top layer, while colder and saltier waters stream beneath. The net flow through the strait has been estimated from both satellite and in-situ observations on the period October 2004 - January 2009 in 0.038 Sv showing high variability (Soto-Navarro et al. 2010). The Mediterranean Forecast System estimate for the transport is 0.057 Sv for the year 2017 (Clementi et al. 2019). The net inflow of Black Sea brackish waters through the Dardanelles ranges from 0.006 to 0.01 Sv as estimated by (Ünlülata et al. 1990a)(Peneva et al. 2001) (Kara, Wallcraft, Hurlburt, and E.V. Stanev 2008).

The forcing mechanisms for the circulation of the Mediterranean can be summarized into three main branches: 1) The net transport of Atlantic at Gibraltar; 2) The change of buoyancy properties of water masses due to heat and water fluxes and consequent formation of deep waters; 3) The action of wind stress.

The result of these driving forces is a time permanent pattern of cyclonic / anticyclonic gyres that are likely found at the border of boundary jets and mesoscale eddies along the horizontal, as well as vertical motion in the case of deep water formation, that is particularly intense in some specific areas of the SES system (see fig. 1.1). This gives rise to an intermediate water circulation. The overall picture of the Mediterranean circulation has been extensively illustrated in (Robinson et al. 2001) and (Pinardi, Zavatarelli, et al. 2015). In the Mediterranean system the large, sub-basin and mesoscale dynamical features coexist, having different importance / intensity depending on the timescale considered. As stated before, the Mediterranean presents aspects common with the ocean, including the thermohaline circulation, formation of water masses and their interaction with other water masses. The way atmospheric forcing acts on the basin scale circulation is the modification of density of relatively fresher waters of Atlantic

origin both because of salinity preconditioning (the evaporation is larger than precipitation) and by intense local cooling events, especially during the winter. These waters are referred to as Modified Atlantic Waters (MAW). The general picture of the thermohaline circulation is shown in figures (1.2) and (1.3), with the 3-dimensional and vertical structure respectively. Figure 1.2 shows the zonal and meridional components of the thermohaline overturning circulation. The zonal component is made up by the MAW (yellow path) that occupies the first 100-200 meters below the surface and on its route to the EMED gradually increases in salinity and density. In the Levantine basin the MAW eventually sinks forming the Levantine Intermediate Waters (LIW), that recirculate in the opposite direction between 200 and 700 meters (fig. 1.3). Further source of deep water masses are depicted in fig (1.2), forming mainly in the winter time:

- The Gulf of Lyon, where WMED Deep Waters (WMDW) are formed
- The Adriatic Sea, from where the Eastern Mediterranean Deep Water (EMDW) connects to the general pattern through the Strait of Otranto.
- The Cretan Deep Waters (CDW) that form in the Aegean Sea (not presented in fig. 1.2)

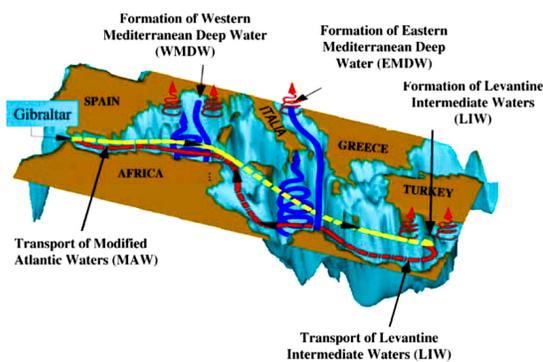


Figure 1.2: Schematic representation of the Mediterranean Overturning Circulation (MOC)

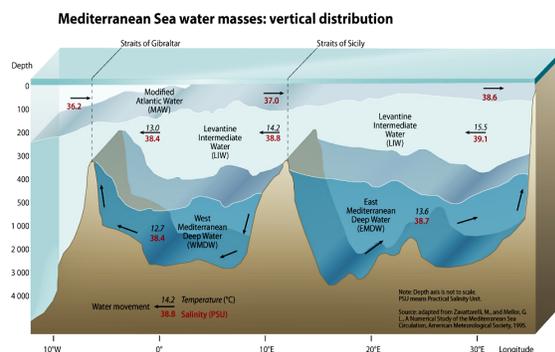


Figure 1.3: Zonal Section of MOC adapted from (Zavatarelli and Mellor 1995)

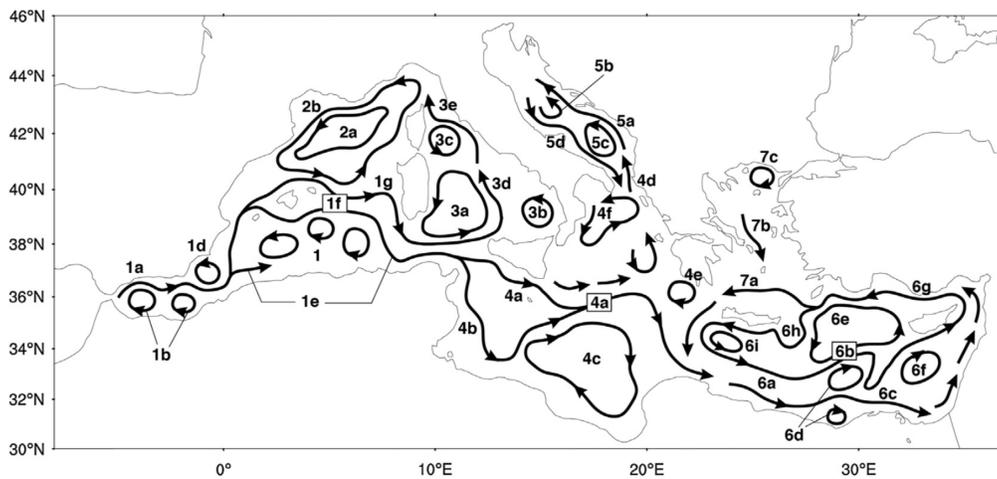


Figure 1.4: Surface Circulation Patterns in the Mediterranean Region from the 1987–2007 reanalysis mean flow field in (Adani et al. 2011) presented in (Pinardi, Zavatarelli, et al. 2015)

### 1.2.1 Sub-basin scale Circulation

The driving forces responsible for the basin scale overturning circulation play a role also in the sub-basin and the mesoscale circulation, that is rich of boundary currents, bifurcating jets, permanent and semi-permanent cyclonic and anticyclonic structures as well as mesoscale eddies.

The role of the wind stress is crucial in the motion throughout the basin, giving prevalence of cyclonic rotation in the gyres in the North, and opposite sense for those in the South. The cyclonic motion of Northern gyres is also corroborated by the processes of deep water formation.

Figures (1.4) and (1.5) give the resulting picture of a 23 long reanalysis of the Mediterranean ocean surface and intermediate depth circulation respectively carried by (Adani et al. 2011) and synthesized by (Pinardi, Zavatarelli, et al. 2015). The common structures visible in the circulation are clustered into systems made up of components.

Starting from west, first semi-permanent structures are found in the Alboran Sea that is heavily influenced by the intrusion into the Mediterranean Sea of surface fresh and relatively cold waters from the Atlantic Ocean: the Atlantic Water Current (AWC, Figure 1.4 1a) enters into the basin through the Gibraltar Strait and propagates towards the eastern part of basin, sinking while it travels

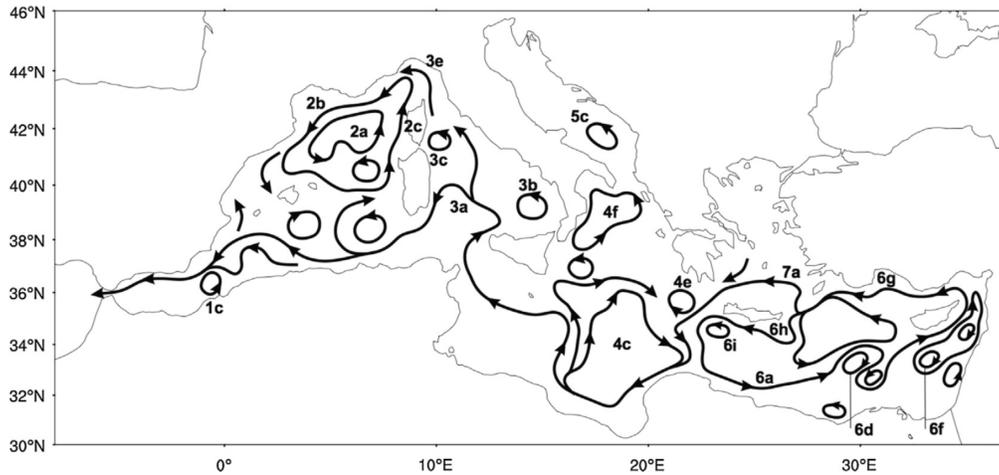


Figure 1.5: As in fig. 1.4 for the 200-300 m averaged currents

from West to East due to the modification of its thermohaline characteristics.

In the Alboran Sea the AWC flanks two semi-permanent anticyclonic gyres, West and East Alboran gyres respectively (fig. 1.4 1b), and the Almera-Oran cyclonic eddy on the opposite to the previous ones (fig. 1.4 1d).

East of the Almera-Oran eddy, the AWC bifurcates into two branches, the first proceeding to the east along the Algerian coast (Algerian Current segments, Figure 1.4 1e), while the second one flows northward and towards the Ibiza channel, to generating the Western Mid-Mediterranean Current (WMMC, Figure 1.4 1f) that runs along the balearic islands.

The Algerian current presents both anticyclonic and cyclonic structures. These short-lived structures are generated by the unstable character of the current (Robinson et al. 2001) and show low translational velocity and a highly variable direction of propagation (Pessini et al. 2018).

The WMMC constitutes the southern boundary of the Liguro-Provencal-Catalan Current (LPCC, fig. 1.4 2b) that flows around the Gulf of Lion Gyre (GLG, fig. 1.4 2a).

The Algerian Current splits into three branches at the Sardinia Channel: one flowing northward into the Tyrrhenian Sea, the Middle Tyrrhenian Current (MTC, fig. 1.4 3d ); the other two entering the Sicily Strait forming the Sicily Strait Tunisian Current (SSTC, Figure 1.4 4b) and the Atlantic Ionian Stream (AIS, fig. 1.4 4a).

The northern branch of the Algerian Current flows into the system of anticyclonic structures of the Tyrrhenian Sea circulation, which has three cyclonic gyres: the South-Western Tyrrhenian Gyre (SWTG), the South-Eastern Tyrrhenian Gyre (SET G) and the Northern Tyrrhenian Gyre (NTG), identified as 3a, 3b and 3c respectively in fig. 1.4. During summer, the dynamic picture of the gyres changes slightly with the formation a well-known dipole located to the east of the Bonifacio Strait, that separates Sardinia from Corsica (Iacono et al. 2013).

The Southern branch of the Algerian Current enters the Sicily Strait and splits into the Sicily Strait Tunisian Current (SSTC) flowing along the coast of Tunisia (fig. 1.4 4b) and the Atlantic Ionian Stream (AIS) that flows between the island of Malta and the southern coast of Sicily (fig. 1.4 4a). Entering the Ionian Sea, the AIS constitutes a jet that freely flows eastward and separates the anticyclonic structure of the so-called Sirte Gyre (SG, fig. 1.4, 4c) after (Nadia et al. 2006) and the Northern Ionian Cyclonic Gyre (fig. 1.4, 4f).

From the Northern Ionian Cyclonic departs what is called the East Ionian Current (EIC, fig. 1.4 4d) that enters the Adriatic Sea passing the Otranto Strait, contributing to the anticyclonic character of the sub-basin. The EIC proceeds northwards along the coast of Balkans, named as Eastern South-Adriatic Current (ESAC, Figure 1.4 5a), having a seasonal behavior, being identifiable all year round except for summer (Artegiani et al. 1997)(Zavatarelli and Pinardi 2003). The anticyclonic circulation is completed by the Western Adriatic Coastal Current (WACC, Figure 1.4 5d) flowing along the Italian coast back to the Otranto Strait and rejoining the Northern Ionian Gyre.

Two cyclonic structures are clearly identifiable in the Adriatic Sea, the Middle Adriatic Gyre and the South Adriatic Gyre (fi. 1.4 5b, 5c).

The AIS proceeds eastward crossing the Cretan Passage and entering the Levantine basin nesting inside its anticyclonic system. The gyre that flanks the Egyptian coast, the Western Asia coast and the Southern Turkish coast is made of Cretan Passage Southern Current (CPSC), Southern Levantine Current (SLC) and the Asia Minor current (fig. 1.4 6a, 6c, 6g).

Further cyclonic gyres are present in the Levantine basin. South and East of Crete the Western Cretan Cyclonic Gyre and the Rhodes Gyre (RG) are visible (fig. 1.4 6i and 6e). The Mid-Mediterranean Jet (MMJ, fig. 1.4 6b) flows between the southern borders of the aforementioned gyres and deviates to Cyprus Island to

merge into Asia Minor Current.

The Levantine basin is characterized also by significant anticyclonic mesoscale systems such as the Mersa Matruh Gyre System (MMGS, fig. 1.4 6d), the Shikmona Gyre System (SGS, fig. 1.4 6f), the Ierapetra Gyre (IPG, fig. 1.4 6h )

The Aegean Sea presents a cyclonic-like circulation, caused by intrusion of Asia Minor Current from the South East (Nittis and Perivoliotis 2002) and from the Black Sea Waters coming from the Dardanelles. This inflow produces an anticyclone in the North of the sub-basin (North Aegean Anticyclone, NAA fig. 1.4 7c) and a southward current flowing along the Greek coastline (Southward Cyclades Current, SCC fig. 1.4 7b) that eventually reaches the Cretan Sea joining the Cretan Sea Westward Current (CSWC, fig. 1.4 7a)

The intermediate depth circulation shown in figure 1.5, representing the 200-300 meters average of the (Adani et al. 2011) reanalysis. The circulation pattern is in great part similar to the surface case, having similar structures. The anticyclonic patterns in the Levantine are almost equal, as well as the SG, the Northern Ionian Gyre and the GLG. Gyres that were present at surface in the shallower seas like Adriatic and Aegean, such as the ESAC and WACC and the North Aegean Anticyclone are not visible anymore at intermediate depth.

The recirculation of the LIW is visible in the intermediate depth portrait. The LIW forms mainly in the North of Levantine for meteorological reason (Robinson et al. 2001), starting to flow westward in correspondence of the CSWC. It comes out of Cretan Sea and meanders the south border of SG before pointing northwards to cross the Sicily Strait and meander cyclonically around the SWTG. Here the LIW bifurcates at east of the Sardinia coast, where one branch flows northward leaving the Italian coast to the right and then reaching the GLG, while the other branch flows towards the Gibraltar Strait, crossing an area of significant mesoscale activity, and finally exiting from the Mediterranean Sea basin.

### **1.3 Dynamics of The Black Sea**

The main driving forces of the Black Sea dynamics are the wind stress, with the result of the anticyclonic behaviour of its general circulation, and the riverine discharge, that acts as a reinforcement. The general circulation pattern is commonly

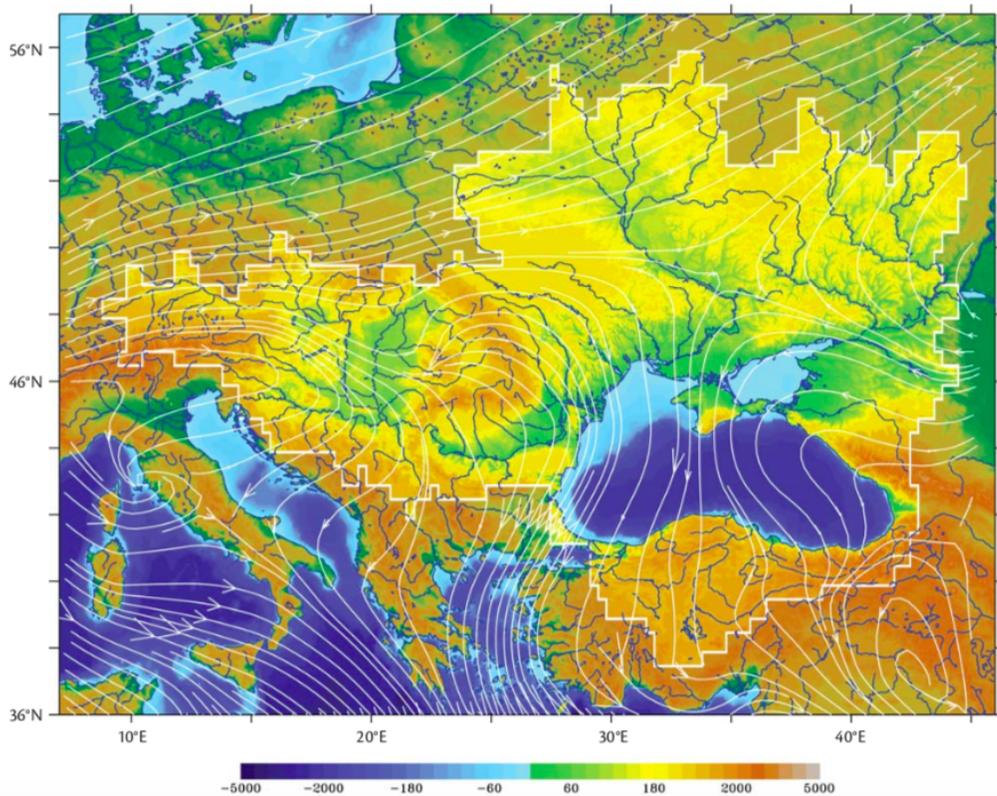


Figure 1.6: The Black Sea nested in its catchment (white polyline) from (E.V. Stanev 2005)

referred to as the Rim Current. The considerable amount of freshwater provided by its catchment related to the size of the basin makes the Black Sea a typical estuary water body. Figure (1.6) shows the Black Sea inside its catchment area that encompasses Eastern Europe and part of Asia Minor.

Most of this river runoff is inserted in the northwestern part of the basin, where Danube and Dniepr rivers alone provide 70% of the total runoff in the Black Sea (Kara, Wallcraft, Hurlburt, and E.V. Stanev 2008). The river runoff acts as a reinforce to the anticyclonic Rim Current flowing in the coastal area, associated to a difference of sea level that is around 20 cm above the sea level in open waters (see fig. 1.7).

The rim current envelops two cyclonic gyres, generally referred as Western and Eastern gyre. Several mesoscale eddies are visible also bordering the rim current. This happens in areas of steep slope of the bathymetry where the current

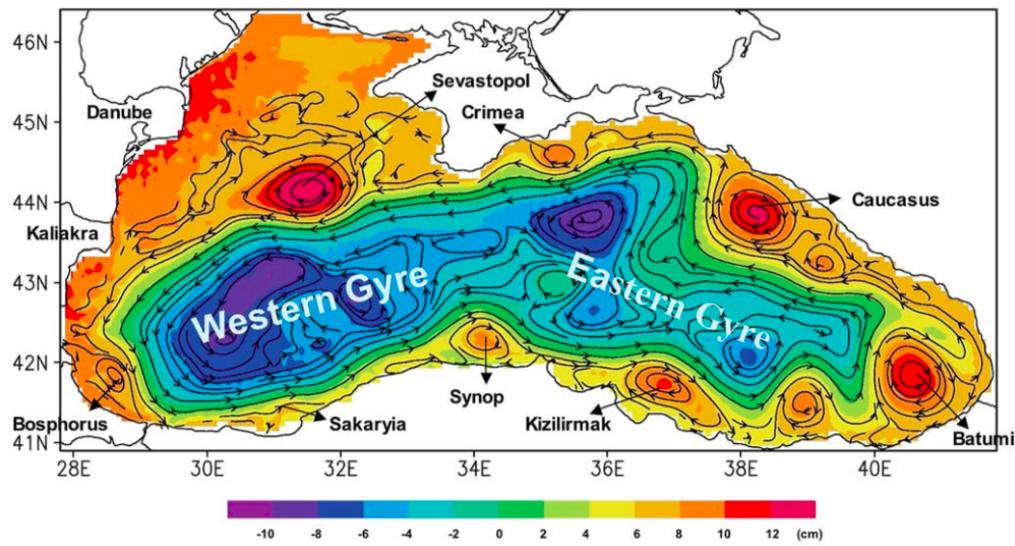


Figure 1.7: Sea level and simulated currents by the DieCAST model in (J. Staneva et al. 2001)

interacts with the bottom slope zone approximately as a vertical wall, and the eddy field becomes highly unstable and meandering (E.V. Stanev 2005).

The dynamical scheme shown in fig. (1.7)) is consistent with the summer more intense mesoscale activity, with more pronounced eddies. The eastern eddy, called Batumi Anticyclonic Eddy, can increase and fill the entire corner of the basin (Korotenko 2015).

During the winter season, the rim current intensifies and the circulation is organized in one gyre system encompassing the entire basin (the Batumi sub-basin gyre is absent) (E.V. Stanev 2005).

An important feature of the Black Sea thermohaline circulation is the formation of the so-called Cold Intermediate Layer (CIL), that generally can be found between 25 and 150 m depth (Miladinova et al. 2018) with temperatures less than 8 C, colder than the surface water, that in summer can reach 25 C and the deep water, having values around 9 C. The strong stratification deriving from the presence of the CIL in subsurface waters hampers the vertical mixing and results in lack of oxygenation of the deep waters.

The formation and evolution of the CIL are not entirely understood but there is large agreement about the presence of two main driving mechanisms (Ivanov et al. 1997)(E.V. Stanev, Bowman, et al. 2003) : 1) the cooling of the surface waters

in the basin interior during the winter and 2) transportation of the cold water masses, formed in winter on the North West Shelf, by the Rim Current.

## 1.4 Dynamics of the Turkish Strait System

Exchange flows between The Black Sea and the Mediterranean take place in the Turkish Strait System ( TSS ), and are driven by the density differences. This difference of density results, also, in a permanent gap in the BS sea level and the Aegean sea level, for which estimates range from 30 (+/- 10) cm (Beşiktepe et al. 1994) to 55 cm (Alpar and Yüce 1998).

The climate of the Marmara region has moderated wind, because of the conformation of land topography with prevalence on northeasterly winds (Beşiktepe et al. 1994), that weakly influence the general circulation.

Figure (1.8) shows the budget of water masses flowing across the components of TSS of fluxes and salinity computations by (Ünlülata et al. 1990a) based on the Knudsen (Knudsen 1900) relations expressing the steady-state mass budgets. Arrows leaving/entering the surface represent evaporation/precipitation respectively and tilted arrow in the Black Sea compartment represent the estimate of the annual average of the total river runoff. Numbers in brackets represent the values of salinity.

The resulting steady state flow is a made of an upper layer of fresher BS water that progressively become more saline in the path to the Aegean. A counter flow of saltier and denser water from the Aegean to the BS closes the water budget.

The mass budget shows clearly that the 2 layer flow of Marmara is almost totally driven by the BS excess of water ( $352 \text{ km}^3/\text{y}$  by runoff +  $300 \text{ km}^3/\text{y}$  by precipitation -  $353$  by evaporation) being the corresponding values for the Marmara much smaller ( $-4 \text{ km}^3/\text{y}$  from the evaporation/precipitation budget). The net flow of water through the Bosphorus resulting from this steady state mass budget is around  $300 \text{ km}^3/\text{y}$ , with similar value in the Dardanelles.

The upper layer waters, then, largely reflect the seasonal characteristics of Black Sea water modified in transit through the Bosphorus, and by local heating/cooling and mixing in the Sea of Marmara.

Seasonal variations are essentially confined to the well aerated upper layer

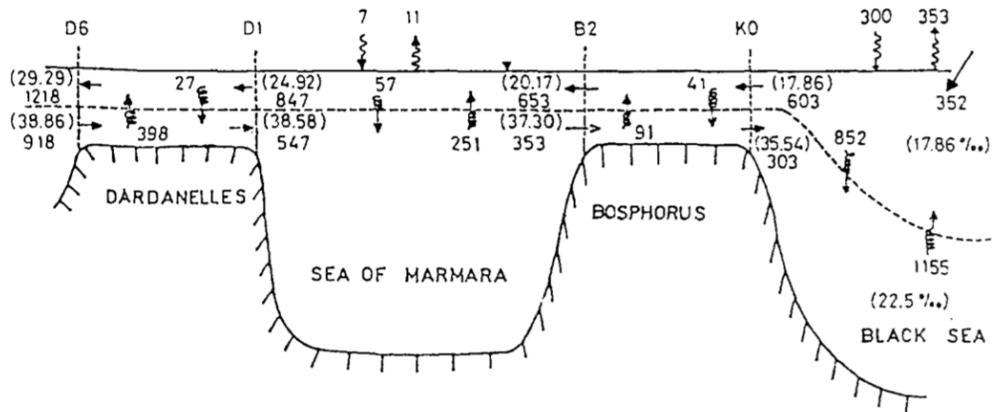


Figure 1.8: Volume fluxes in the components of the TSS (Beşiktepe et al. 1994)

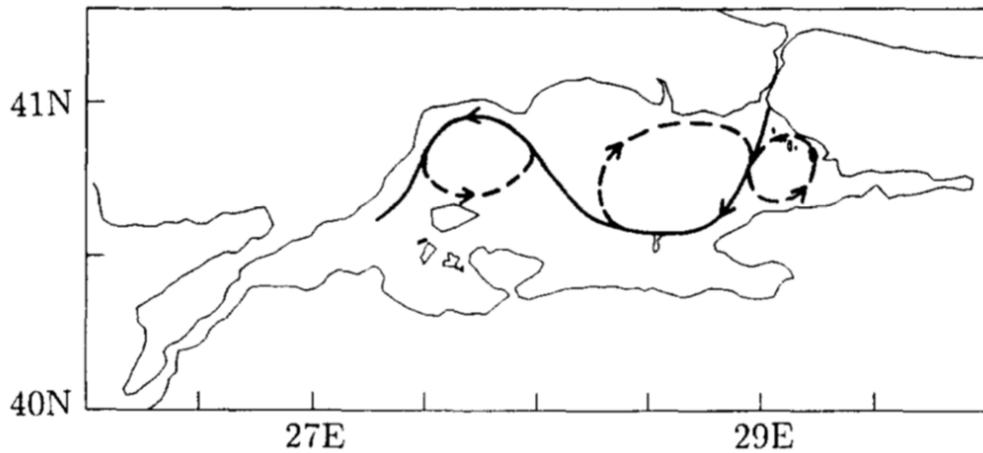


Figure 1.9: Surface Circulation in Marmara Sea. From (Beşiktepe et al. 1994)

(first 25m), but extreme wind mixing events are able to erode the halocline down to a depth of about 40m (Beşiktepe et al. 1994). The upper layer salinity is in the range of  $23 \pm 2$  PSU, reaching a maximum in winter, as a result of increased wind mixing in the basin and reduction in the influx from the Black Sea.

Figures (1.9) shows typical circulation patterns in the Marmara Sea for the surface.

The surface circulation shows the S-shape current with anticyclonic character triggered by the Bosphorus Jet. The jet initially flows to the south; upon reaching the coast of the Bozburun Peninsula, it is deflected west and then northwest, circulating around a closed anticyclonic circulation attached to the Thracian coast

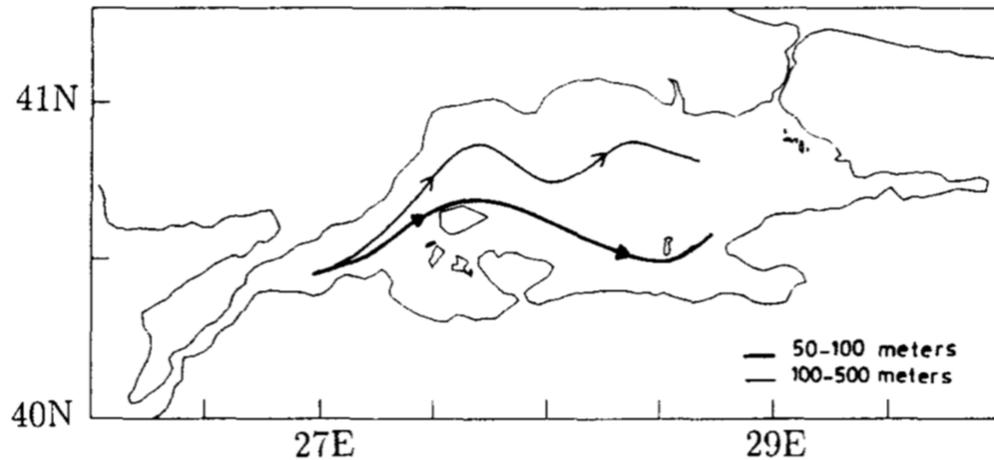


Figure 1.10: Circulation in Marmara Sea in deep waters. From (Beşiktepe et al. 1994)

before flowing towards the Dardanelles. This pattern, regulated by the Bosphorus inflow and the basin geometry, is typical of summer period, and it is weakly influenced by the wind.

During the winter period, with stronger wind regime and moderated BS river runoff and consequent smaller inflow through the Bosphorus, the circulation pattern can change radically from the summer, with reversal of the circulation. The combination of small throughflow and strong wind can cause the reversal of the Dardanelles flow, with upper water flowing from Aegean to Marmara.

The schematic representation of the deep circulation of the Marmara is in fig. (1.10). The circulation in the sub-halocline of the Marmara Sea is very slow, the net motion being a slow drift from west to east following the bathymetry.

## 1.5 State of the art of the SES modeling

Literature has plenty of examples of numerical modeling in the Southern European Seas, most of them focused on a specific area, like the coastal models, or at higher scales, covering an entire region of the system.

To the author's knowledge, only 2 recent studies have explored the possibility of modelling the SES system entirely.

The study of (C. Ferrarin, Bellafiore, et al. 2018) considers the full system of

the inter-connected Mediterranean, Marmara, Black and Azov (briefly named MMBA) modelled seamlessly with the unstructured grid model SHYFEM that does not consider baroclinic contributions in the computations. The main topic of this study is the role of tides in the dynamics of the system, for which a rather resolution of 12 km in open water is deemed as sufficient, with 2 km grid spacing at coastline and few hundred of meters in the straits. The simulation is forced with internal equilibrium tidal potential and Atlantic tides. The validation is made against a set of 77 sea-level observation from tide gauges.

The tides are also studied in the work of (Palma et al. 2020), in a similar work carried with MITgcm model (Marshall et al. 1997), that has a nearly constant resolution of 2 km in the domain, increased to 380 m in Gibraltar and 120 in the Turkish Straits. The model is initialized and forced with tracers, sea level and velocities at Atlantic boundary from CMEMS fields with the addition of tidal potential and main lunar, solar components (M2, O1, S2, K1) prescribed at the open boundary. Differently from the previous study, the study includes the baroclinic terms in the computations, focusing on the role of tides in modulating the transports through the straits and the intensity of the circulation patterns in the short period (few weeks). The main conclusion of this study is that the tides affect the circulation significantly and that their contribution generates non-linear effects that do not simply overlap the general circulation.

This study, as well as the previous, does not consider the contribution of the atmospheric forcing nor the effects of such non-linear contribution of tides in the medium-long period.

This section unfolds the advancements of modelling studies from the first implementations to presents days, organised for each one of the main basins of the SES system.

Given the abundance of examples, we provide here a selection that suits more with the scope of this work.

### **1.5.1 Modeling of the Mediterranean Sea**

A 10-year long simulation of the Mediterranean has been carried out by (Zavatarelli and Mellor 1995) with the Princeton Ocean Model (Blumberg and Mellor 1987). The horizontal mesh is orthogonal-curvilinear with maximum stretch in

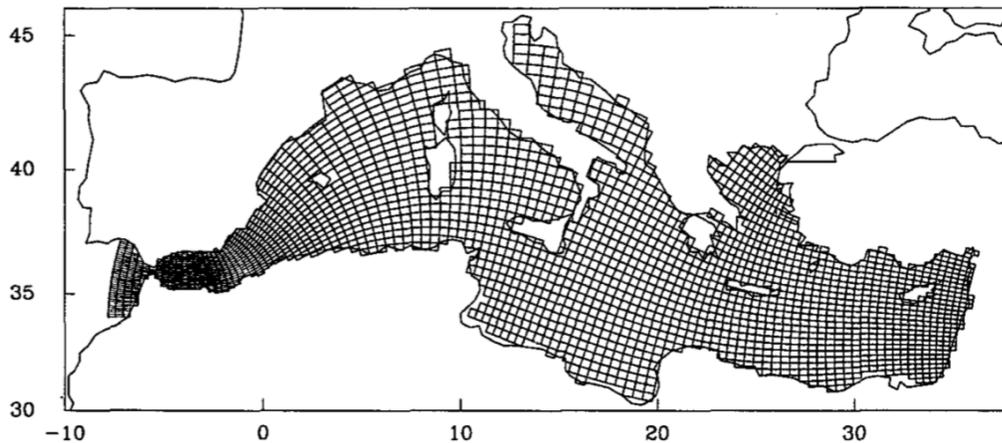


Figure 1.11: Curvilinear grid used by (Zavatarelli and Mellor 1995) to model the Mediterranean with POM

correspondence of the Gibraltar Strait with 8 km resolution and a minimum resolution in the North Adriatic, with a resolution of nearly 60 km (fig. 1.11).

The model was forced at the surface by monthly climatologies of wind stress, heat, and salinity fluxes. The lateral open boundary west of Gibraltar was forced with a climatology of Temperature and Salinity with a Sommerfeld radiation condition.

The model was able to catch the large scale features such as the formation of deep water, the 2 layer inflow/outflow at Gibraltar, and the propagation of the MAW across the Strait of Sicily and into the EMED.

The Project of MFS (Pinardi, Allen, et al. 2003) has laid the foundation for the advancements both in the observing system of the Mediterranean and in the modeling. The project has demonstrated the feasibility of running short term forecast at the basin scale and the possibility of downscaling the problem via nesting different resolution models.

The model used in this project was the Modular Ocean Model ( MOM ) (Pacanowski et al. 1990) with a resolution of 1/8. The model had a rigid lid formulation with computation of sea level anomaly in a diagnostic manner (Pinardi, Rosati, et al. 1995)

Further advancements in the framework of MFS have been made by (Tonani, Pinardi, Dobricic, et al. 2008) and (Oddo et al. 2009) . The horizontal resolution of the previous implementations has been doubled (1/16, approximately 6.5 km)

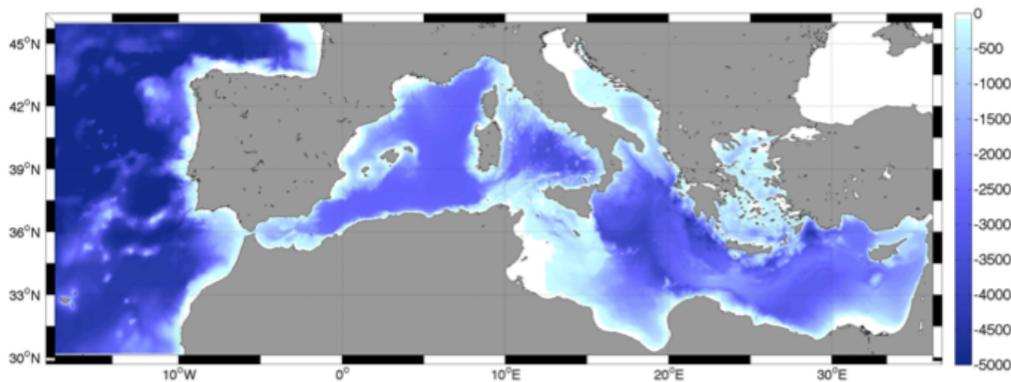


Figure 1.12: Domain and Bathymetry of Mediterranean Forecast System

as well as the vertical, which has been upgraded to 72 levels with a resolution of nearly 3 meters near the surface and around 300 meters in the deepest levels. The numerical model used in the simulations was Ocean PArallelise (OPA) (Madec et al. 1998). The resolution of 6.5 km renders the simulation mesoscale resolving since the first baroclinic Rossby radius of deformation for the Mediterranean region is around 15 km.

The remarkable advancement of this implementation is the free surface, that allows for net transport through Gibraltar strait to compensate for the water loss by evaporation at the basin scale.

In order to simulate Atlantic - Mediterranean water exchanges properly, the model setup considers a large Atlantic box, shown in figure (1.12)

The model has relaxation to a climatology of active tracers at the lateral boundaries of the Atlantic box and has enhanced diffusivity in the area of Gibraltar to get have realistic values of transports. The results obtained in two 6 years-long experiments are in good agreement with observations with the exception of the water flux.

The experiments carried by (Oddo et al. 2009) make use of the same model configuration described above in a 4 years-long simulation, with the addition of nesting the Mediterranean model in a global model, from which total velocities are taken, to test the sensitivity of the Mediterranean circulation in response to the influence of the Atlantic. Results are compared to a simulation that has closed boundaries and relaxation of tracers to climatology in the same fashion as in (Tonani, Pinardi, Dobricic, et al. 2008).

Major benefits derived from the nesting the model into the global solution are found in the representation of Atlantic waters and the variation of the Mediterranean water volume. The comparison of the closed simulation with the open one highlights the difficulties of implementing workarounds for the conservation of total volume. To preserve the total volume in the closed simulation, the water lost by evaporation in the Mediterranean region is compensated with a positive factor in the Atlantic box, with the cost of excessive dilution of Atlantic water and too freshwater flowing through Gibraltar. The open boundary simulation shows a more realistic salinity for inflowing Atlantic water.

A similar configuration is used currently in the framework of MFS to provide analysis and short-term forecast of the main physical parameters in the Mediterranean Sea (Clementi et al. 2019) as a product of the CMEMS catalog.

The analysis and forecast Mediterranean currents system are provided by means of a coupled hydrodynamic-wave model implemented over the whole Mediterranean basin and extended into the Atlantic Sea in order to better resolve the exchanges with the Atlantic Ocean at the Strait of Gibraltar. The model horizontal grid resolution is 1/24 (ca. 4.5 km) and has 141 unevenly spaced vertical levels.

The hydrodynamics are supplied by the NEMO (v3.6 (G. and Team n.d.)) while the wave component is provided by WaveWatch-III (Tolman 2008). The model solutions are corrected by the variational assimilation (based on a 3DVAR scheme) of temperature and salinity vertical profiles and along track satellite Sea Level Anomaly observations.

The Dardanelles Strait is implemented as a lateral open boundary condition by using CMEMS Global daily Analysis and Forecast product and daily climatology derived from a Marmara Sea box model (Maderich, Ilyin, et al. 2015).

## 1.5.2 Modeling of the Black Sea

Most of the modeling studies for the Black Sea have been carried since late nineties of XX century and early 2000's using regular grid models.

The DieCast z coordinate rigid lid model (Dietrich 1998) was used by (J. Staneva et al. 2001) with a high horizontal resolution configuration (1/12). Both the high resolution and the fourth-order accurate approximation for advection terms and

pressure gradient led to represent important mesoscale structures of the Black Sea realistically (1.7). Previous studies (J. V. Staneva and E. V. Stanev 1997) (Emil Stanev and J. Staneva 2000) used the MOM model with the same resolution could replicate quite accurately the formation of the wintertime intermediate water but the Rim Current was more diffuse and the coastal anticyclones produced were larger than found in nature.

The study of (J. Beckers et al. 2002) aimed at modeling the Black Sea with particular focus on the North West Shelf under the influence of relevant rivers such as Danube, Dniepr and Rioni. The model implemented for the study was GHER-3D (J.-M. Beckers 1991), initialized with an horizontally homogeneous stratification spun-up for several years under the forcing of climatological averages of monthly mean wind-stress fields, sea surface temperature and sea surface salinity. Two twin experiment with different resolution (15 km and 5 km) we carried, both showing a general circulation pattern compatible with classical views of circulation in the Black Sea. Of these two experiment, the high resolution one has more marked fronts and greater internal variability. For the Danube region, both experiment simulate a reversal of the current during the summer season.

Since 2016, the Black Sea is modelled in the framework of Black Sea Observing System (Palazov et al. 2019)(Ciliberti et al. 2020) with state-of-the-art models for the physics (NEMO, (G. and Team n.d.)), biogeochemistry (BAMHBI, (Grégoire et al. 2008)) and waves (WAM, (Group 1988)). The framework provides analysis and forecasts released as CMEMS product.

The physical component of the Black Sea near real time forecasting system has horizontal grid resolution of 1/36 and 1/27 in the zonal resolution and in meridional resolution respectively (nearly 3 km) and has 31 unevenly spaced vertical levels. The model solutions are corrected by the variational assimilation (based on a 3DVAR scheme), originally developed for the Mediterranean Sea and later extended for the global ocean. The observations assimilated in the BS-Currents includes in-situ profiles, along-track sea level anomalies (SLA) and gridded sea surface temperature (SST) from ECMWF. The river sources are modelled as surface freshwater inputs.

A more recent implementation of NEMO in the Black Sea has been undertaken by (Gunduz et al. 2020). The model implemented has an horizontal resolution of 2.5 km and 60 z levels. The advancement done with respect to previous works

is the inclusion of the Azov Sea in the domain and a box in the Marmara Sea to simulate the response of the Black Sea with the Bosphorus, though the full representation of dynamics that can occur in the strait are beyond the model setup capabilities. The boundary condition imposed at the Marmara box boundaries include relaxation to climatological values of Temperature and Salinity. The 10 years long simulation is forced at surface with ERA5 atmospheric reanalysis from ECMWF and 10 rivers, being the biggest the Danube, Dnieper and Don.

Overall, 10 years of simulation of temperature and salinity do not show any significant drift, which is an important feature that would allow the model to be used for long-term climate change simulations. The main features of the Black Sea oceanography such as CIL water mass, rim current and upwelling along the southern coast are all well detected by the newly developed model. The model was also able to reproduce the SSH variability in comparison to the tide gauge observations.

Last years have seen also the application of unstructured grid models to the Black Sea. The study of (Bajo et al. 2014) uses the SHYFEM finite element hydrodynamical model (Umgiesser, Canu, et al. 2004) with resolution varying throughout the domain from a maximum of 100 meters along the Romanian coast, where the study is focused to a minimum of 1.5 km in the open waters. The Azov Sea is not considered in the model and the Bosphorus open boundary is forced with sea level observation. The atmospheric forcing is from NCEP/NCAR reanalysis. The aim of the study is the influence of the rivers in the mesoscale circulation near the Romanian coast. The analysis is enriched including passive tracers at river mouths to determine the path of river plumes across the seasons.

A more comprehensive application is the one implemented by (E.V. Stanev, Grashorn, et al. 2017), using SCHISM (Y.J. Zhang et al. 2016) to model the system of interconnected basins from the Azov to the Mediterranean through the Kerch, Bosphorus and Dardanelles strait with a seamless grid.

The model simulation is 4 year long, including the 2 years of spin up.

The lateral open boundary in the Aegean Sea is forced with fields from CMEMS Copernicus products, while atmospheric forcing includes 6-hourly wind, atmospheric pressure, air temperature, and dew point temperature from the 0.2 ECMWF product.

The model shows good skills in representing both stratification and the

mesoscale structures of the Black Sea circulation. The seamless fashion of the modelling system allowed for deeper insights in the sea level difference between the Black Sea and the Aegean.

As a result of the numerical simulations, the sea level in the interior part of the Black Sea can be lower than that in the Marmara Sea and even in some parts of the Aegean Sea. The correlation analysis of the numerical results demonstrated that it is not only the sea-level difference at the both sides of the straits that controls the transport between the Black Sea and Marmara Sea but also the intensity of circulation.

### 1.5.3 Modeling of TSS

Literature concerning the modeling of Turkish Strait System is not ample. Modeling efforts have regarded either the components of TSS as standalone or de-coupled systems. Among the relevant literature we mention (YS 1993), who introduced a 2D reduced gravity ocean model of the Dardanelles inflow into the Marmara Sea. Other simplified 2D models aimed at reproduce the dynamics of the Bosphorus exchange flows ((Maderich and Konstantinov 2002); (Ilicak, Özgökmen, et al. 2009); (Maderich, Ilyin, et al. 2015)).

A first attempt to model the circulation of Marmara in response to the atmospheric forcing and the interaction with the two straits was carried by (Chiggiato et al. 2012). The model employed for this simulation is ROMS (Haidvogel et al. 2008), a hydrostatic finite difference model that discretizes the horizontal space onto a curvilinear-orthogonal mesh, and the vertical space with  $s$ -coordinates. In this simulation the horizontal resolution would vary from a minimum of 1.5 km in the center of basin to a maximum of 500 m in the Bosphorus. None of the straits is properly resolved. The model is nudged at the lateral boundaries to match the data of T,S and currents measured during NR/V Alliance campaign coordinated by NATO Undersea Research Centre during September 2008 - February 2009. The model, also, was forced at the surface with data produced by Italian Air Force National Meteorological Center, and limited in duration by the length of the observations used to force the boundaries

The model solutions produce a general circulation in the Sea of Marmara consistent with previous studies. The simulations show that both the wind curl

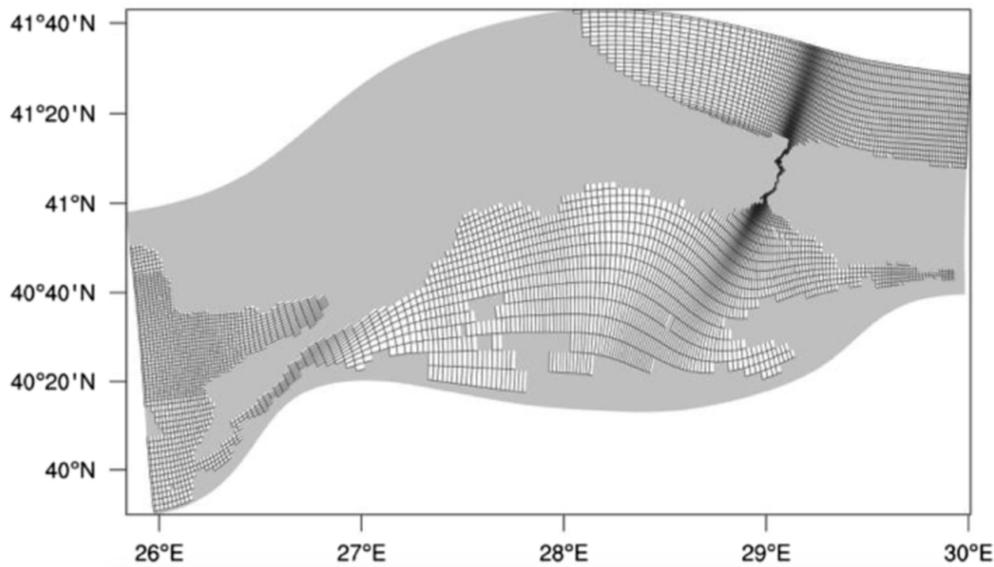


Figure 1.13: Curvilinear grid used by (Sannino et al. 2017) to model TSS (grid is undersampled)

and the strait flows are found to influence the strength and location of the main semi-permanent features of the circulation.

A 18-year long simulation was carried, using a regular grid baroclinic hydrostatic model with constant resolution (nearly 1 km along both directions), by (Demyshev et al. 2012), without contribution of wind. The same S-shape jet as described in (Beşiktepe et al. 1994) was detected as semi-permanent structure surrounding the main anticyclonic structure occupying the center of Marmara.

Similarly to (Demyshev et al. 2012), also (Sannino et al. 2017) does not consider the wind effect in his model for TSS. The hydrostatic MITgcm model, with implicit treatment of free surface and bottom representation with partial steps was used for this study, covering the Marmara and the areas beyond the straits with a curvilinear grid as in fig. (1.13). Average resolution is 1km and straits are resolved with 50 m grid spacing. The vertical column is discretized with 100 z levels.

The response of the TSS system was studied under the effect of different barotropic signals applied at the boundaries. The circulation changes from a large anticyclonic circulation at the centre of the basin at low flux values, to different gyres wrapped around an S-shaped jet as the net flux is increased.

The last relevant work is the one carried by (Aydoğdu et al. 2018), modeling the TSS under the response to complete atmospheric forcing and strait dynamics using the unstructured grid model FESOM (Danilov, Kivman, et al. 2004) in a 6 years-long simulation.

The model covers the TSS system with resolution varying from 65m and 150m in the straits and 1.6km in open water, almost entirely the Aegean Sea and the South West corner of Black Sea, where the resolution is around 5 km. The model has closed open boundaries in the Aegean and BS, hence volume and salinity conservation are imposed in a buffer zone in the Black Sea to prevent drifts in the tracer fields.

The model results for the transports at Bosphorus compare well against the observations. Results also confirm the role of the Bosphorus jet in the anticyclonic circulation structure in the eastern part of Marmara and make account of the wind role in the in the creation of a cyclone in case of intensified wind stress, with small scale vortices are also formed in various parts of the basin.

## 1.6 Unstructured Grid Models

Recent decades have seen an increase of application of numerical models to scales smaller than the basins scales mentioned in the previous sections, directing the attention to coastal, strait or even harbour scales.

The application of finite difference models has highlighted the difficulty of representing complex features of the coastline, such as in the case of (Chiggiato et al. 2012) and (Gunduz et al. 2020) to model the Turkish Straits.

Coastal modelling aims at providing connection between the general circulation and the local dynamics of the area interested by the study, that is closely connected to human activities.

The common approach adopted is the dynamical downscaling, that differently from the case where the coastal model describes both large scale and local scale dynamics, the coastal model creates the transition between the large scale, described by a parent model, and the area of interest.

Global or large scale ocean forecast systems have accomplished large advancements but many practical applications (e.g., calculating the probability of extreme

events, forecasting oil spill trajectories, supporting marine search and rescue, interpreting the movement of tagged marine animals) require information on spatial scales that are too small to be resolved. The need for higher resolution can be particularly acute on continental shelves where variability on scales of several km and less can be of practical importance. However, regional ocean conditions are often controlled by processes operating on large scales, e.g., western intensification of boundary currents, propagating Rossby waves. This leads to the need for downscaling to help estimate local and regional features from coarser scale patterns (Katavouta and Thompson 2016).

The downscaling, moreover, allows to ease the cost of simulation, since the nested model covers only the part of the domain of interest. The implementation of a spatially confined simulation under the influence of a coarse/larger simulation (the so-called parent model) can present the non-trivial problem of the open boundary specification, for which several possible solutions are present in the literature (Marchesiello et al. 2001).

Literature of coastal application has plenty of examples and a complete dissertation would be too long. Here we just report those examples of applications related to human activities such as lagoons, river deltas, fisheries.

With a 2D unstructured grid barotropic model, based on finite element discretization, (Umgiesser, Canu, et al. 2004) has implemented a system of the prediction of tides in the lagoon of Venice (fig. 1.14). The model is forced at the 3 inlets of the lagoon by a prescribed water level and at the surface by wind, rain and solar radiation. The hydrodynamics is de-coupled from the thermodynamics, since in the particular environment of the Venice lagoon the pressure gradient due to density differences is deemed as negligible with respect to the surface pressure gradient. The results show the model capabilities to cope with drying of the shallow tidal flats of the lagoon and the agreement with the tide gauges in a 1 year-long simulation.

The tides are also studied in the work of (C. Ferrarin, Bellafiore, et al. 2018) that considers all the sub-basin of the SES system (called Mediterranean-Marmara-Black Sea-Azov system modelled in (C. Ferrarin, Bellafiore, et al. 2018)) in a 1 year long 3D SHYFEM simulation without baroclinic contribution nor atmospheric forcing. The model was evaluated against observed sea level in the Mediterranean, Marmara and Black seas, illustrating the capability of the model to simulate the

internal tidal generation and the exchange of water and energy between adjacent seas.

The same hydrodynamical model has been used by (C. Ferrarin, Davolio, et al. 2019) in a domain that covers the Adriatic Sea and includes the Venice and Marano Lagoons and the Delta of Po river. The model shows the capability of representing the extent of salt intrusion inside the Po Delta during flood tide phases and storm surge events in the Lagoons. The model, additionally, is equipped with a lagrangian particle tracking module (Cucco et al. 2012) that turns useful to simulate the dispersion of pollutants, as it happens in the Gulf of Trieste when discharge of Isonzo river increases significantly during flood events (Covelli et al. 2007).

Another example of interaction of downscaling to asses the interaction of the local dynamics with the ecological factors is given in (Huret et al. 2007), where the FVCOM model (Chen et al. 2003), has been used to model the transport of cod larvae in the Gulf of Maine. Further applications of FVCOM are related to large scale, with the study of response of Ocean Mixed Layer under the Hurricane Sandy (Li et al. 2020) or in a fully coupled sea-ice model to simulate the seasonal and inter-annual variability of the Arctic (Yu Zhang et al. 2016).

The unstructured grid models make commonly use of semi-implicit algorithm to avoid numerical instabilities, especially in the case of high resolution simulations. Some models can run in unconditionally stable conditions, using Eulerian-Lagrangian schemes for the momentum advection. This feature is a legacy from the atmospheric models where this kind of scheme is widely used. The SCHISM (Semi-implicit Cross-scale Hydroscience Integrated System Model) (Y.J. Zhang et al. 2016) model uses such scheme and is devised to simulate river-estuary-shelf scales. Recent applications of SCHISM consider larger domains, as in the case of (E.V. Stanev, Grashorn, et al. 2017) described in section 1.5.2 or even larger application as in the case of (Fernandez Montblanc et al. 2019) for the 40 year long hindcast of sea level trend and variability along the coast of Southern and Northern Europe.

The advancements of the computational power has allowed numerical simulation that cover large spatial scales and temporal scales, retaining the horizontal resolution necessary to resolve the main processes of interest.

In the recent years unstructured grid models such as FESOM2 FESOM2

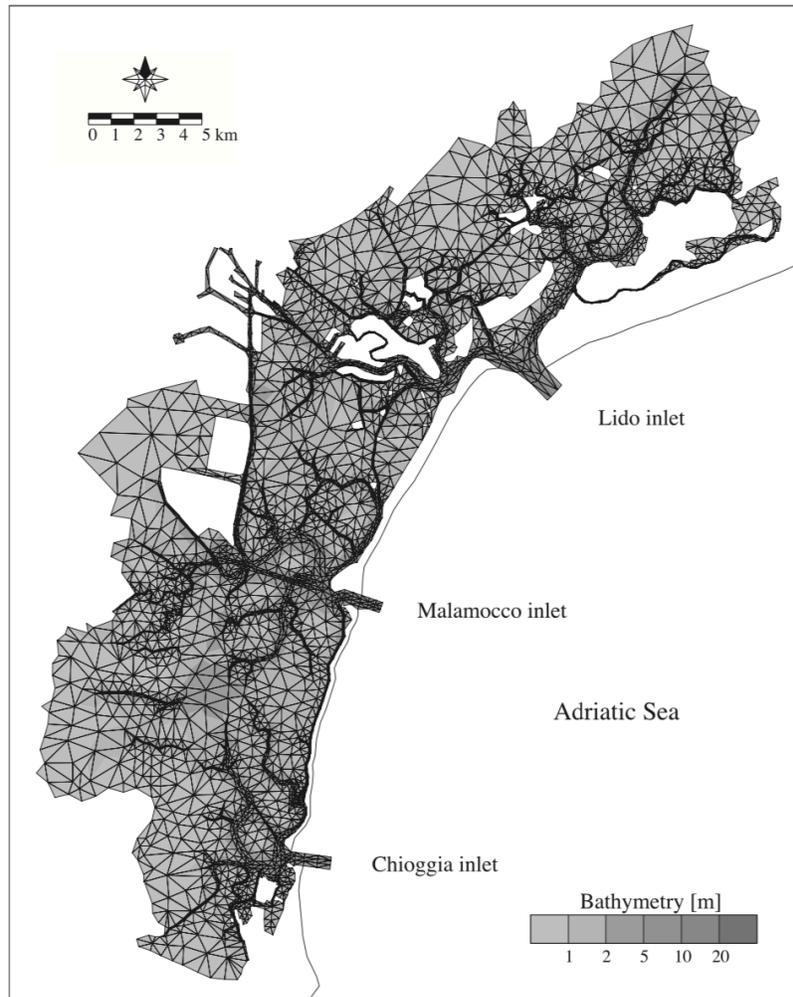


Figure 1.14: SHYFEM Grid and bathymetry used by (Umgiesser, Canu, et al. 2004) to predict tides into Venice Lagoon. The grid has 7842 triangular elements and 4359 nodes.

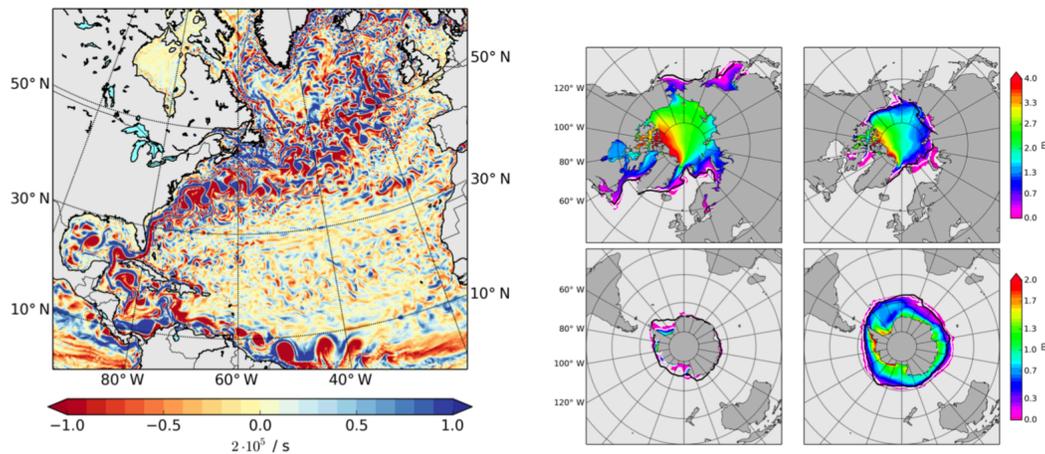


Figure 1.15: [Left Panel] Snapshot of Sub-Surface Relative Vorticity eddy permitting simulation [Right Panel] Modelled Sea Ice thickness in March (left) and September (right). From (Danilov, Sidorenko, et al. 2017)

(Danilov, Sidorenko, et al. 2017) and MPAS (M. R. Petersen, Asay-Davis, et al. 2019), formerly devised specifically for small scale application for their ability of representing complex features of coastlines, have been employed to simulate the global ocean. The simulations carried in the two cases have an eddy permitting and eddy resolving horizontal resolution respectively. The length of the simulation and the coupling with a sea ice model in both cases shows the capability of the unstructured models as valuable tools in the study of climate.

## 1.7 Parallel Ocean Numerical Models

The aforementioned applications are marked out by high horizontal / vertical resolution and long periods of simulation and are very demanding of computational resources. Climate simulation have computational domains of nearly  $10^8$  grid points and simulate tens of year.

Numerical weather prediction (NWP) is one of the first applications of scientific computing since mid 40s of XX century (CHARNEY et al. 1950) and remains an insatiable consumer of high-performance computing today.

The advancements made by both computer power and by scientific applications based on numerical methods are mutual, in the sense that the former have accomplished improvements in response of ever higher demanding of computa-

tional resources by the scientific community while the latter, especially for what concerns modelling earth sciences, has made remarkable advancements upheld by the increased capability of HPC facilities.

In order to carry these kind of simulation in a reasonable time, it is necessary to adapt the numerical codes to be executed on state-of-the-art machines like HPC facilities.

Parallelization refers to the process of taking a serial code that runs on a serial CPU and spreading the work across multiple CPUs.

The process of modifying a code for scientific purposes, as it can be a code for solution of PDE, to be run on HPC supercomputers is referred to as "parallelization", and it can be more or less intrusive depending on the type of directives / approach is meant to be adopted.

The three main methods of code parallelization and associated directives are:

- Distributed memory approach: MPI
- Shared Memory approach: OpenMP
- GPU: CUDA, OpenACC

Each on the aforementioned methods entail advantages and disadvantages.

The distributed memory approach implies the distribution of the physical domain into sub-domains that can be seen as sealed compartments that do not know anything about the others and to attain information need to communicate through an interface. Sub-domains in a distributed memory framework communicate using the Message Passing Interface ( MPI ) protocol. This kind of approach is potentially unlimited in the level of domain decomposition allowing to ease both the computation and the memory workload of each sub-domain. The reduction of workload is also referred to as "scalability". The distributed memory approach is not straightforward to implement and can imply adversities.

The shared memory approach implies the concept of "multithreading", in which a primary thread divides a task among a certain number of sub-threads performing the task concurrently. The shared memory approach is, in principle, relatively easy to implement by means of the OpenMP directives, but its main drawback is the limitation on the number of threads, that is circumscribed by the number of CPUs on a machine, generally few tens.

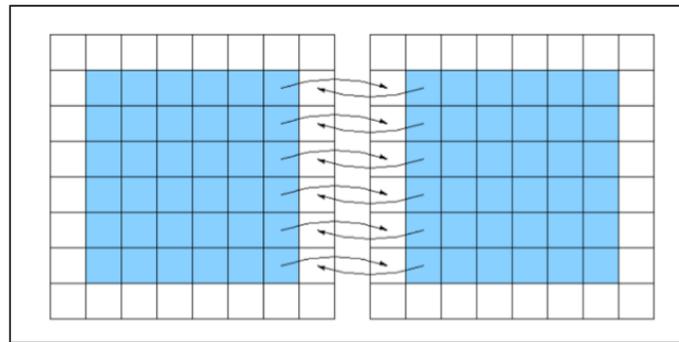


Figure 1.16: Example of domain decomposition of a regular grid for the solution of PDE into 2 MPI Processes (process 0 [left], process 1 [right]). Calculations are made on inner domain (cyan cells). White cells represent the ghost cells or "halo" where lateral boundary conditions for each sub-domain are updated by means of "messages" (black arrows)

The GPU were formerly devised to handle computations for computer graphics and are marked out by higher performances if compared to conventional CPUs. The employment of GPUs in general circulation models is at early stages.

The main point of the parallelization of codes is to allow to complete numerical simulations in a reasonable time. The parallelization, though, addresses 2 further desired requirements.

Key concepts of parallelization are:

1. Time scalability. The parallel simulations is faster and is carried in a time that is considerably shorter that the serial case.
2. Memory scalability. Highly computationally and memory intensive simulations are allowed.
3. Code reproducibility. The parallelization does not change in a sensible way the representation of the physical problem with respect to the serial code

The MPI paradigm satisfies 1),2) and 3) points but the implementation is not straightforward, since MPI processes do know nothing about the others and need to communicate through "messages".

As an example of solution of PDE exploiting a distributed memory approach we take the 2D heat equation

$$\frac{\partial T}{\partial t} = k \left( \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right) \quad (1.1)$$

defined in the space  $0 \leq x \leq L_x, 0 \leq y \leq L_y$  with  $k$  the diffusion coefficient and its finite differences discretization with a FTCS scheme

$$T_{i,j}^{n+1} = T_{i,j}^n + \frac{k\Delta t}{\Delta^2} \left[ T_{i-1,j}^n + T_{i,j+1}^n - 4T_{i,j}^n + T_{i+1,j}^n + T_{i,j-1}^n \right] \quad (1.2)$$

being  $\Delta t$  the time step of integration and  $\Delta$  the spacing of the mesh, taken constant along  $x$  and  $y$  for the sake of simplicity. Both space and time are organised in discrete fashion

$$\begin{cases} x_i = i\Delta & i = 0, 1, \dots, N_x \\ y_j = j\Delta & j = 0, 1, \dots, N_y \\ t_n = n\Delta t & n = 1, 2, \dots \end{cases} \quad (1.3)$$

The equation 1.2 is explicit in time, being the new value of temperature  $T_{i,j}^{n+1}$  in position  $i, j$  depending on its value at time level  $n$  plus the diffusional part that depends on the 4 surrounding points  $i-1, j, i, j+1, i+1, j, i, j-1$ . This access pattern of points needed to calculate a value at the next time step is commonly referred to as a **stencil**. In the MPI case with 2 MPI processes, as in fig. 1.16, each of the 2 processes performs the stencil calculation in a reduced space  $(N_x/2, N_y)$ . The reduced workload of the 2 MPI process determines a speed up with respect to the single process case.

Both MPI processes perform the stencil calculation in their inner domain (Cyan region in fig. 1.16). The white cells between the 2 processes represent an extended memory space made of grid cells where computations are not made and that overlap the computational cells of the neighbour MPI process. These cells are referred to as "halo" or "ghost cells" and are essential for the stencil calculation at inter-process edgcuts.

Figure 1.17 shows the temperature values needed to calculate the stencil in the case of right edge for process 0 to advance to new value of temperature  $T_{N_x/2,j}^{n+1}$ . For each of the iteration in time, the value in point  $N_x/2 + 1, j$  is kept updated with its correspondent value inside neighbour process by means of a

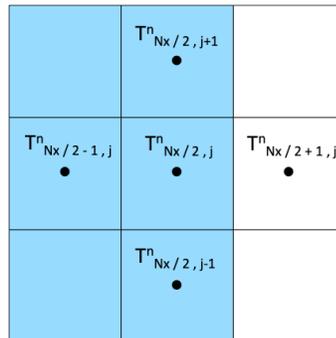


Figure 1.17: Stencil calculation for MPI process 0 in correspondence of right edge. Value in grid point  $N_x + 1, j$  is kept updated by correspondent point of inner domain of process 1

communication. The neighbour process, in a similar fashion, needs the updated value of Temperature in point  $N_x/2, j$  to calculate the stencil in its left edge. The communication pattern is represented graphically by the arrows in fig. 1.16.

The time required for these communications is not zero and the overhead depends on the amount of data to exchange and on the communication frequency.

In the simple example given here, the time to carry the calculation is, in the ideal case of instantaneous communications, halved because each MPI process solves the equation in half of the domain. In the practical case each iteration in time is burdened by a cost due to communication causing the execution time slightly (hopefully) longer than exactly half of the time necessary for one single MPI process. These issues are addressed to as "scalability" and are described more in detail in section 3.5.1

The example given above is elementary but represents the principle that has been applied to parallelize several of the ocean (or atmospheric) circulation models based on regular grids. Examples for the ocean community are Nucleus for European Modelling of the Ocean (NEMO) (G. and Team n.d.), Modular Ocean Model 5 (MOM) (Griffies 2012), MITgcm (Marshall et al. 1997).

Ocean models of the unstructured community, such as Finite Element Sea-Ice Ocean Model 2 (FESOM2) (Danilov, Sidorenko, et al. 2017), Finite Volume Community Ocean Model (FVCOM) (Chen et al. 2003)(Cowles 2008), Semi-implicit Cross-scale Hydroscience Integrated System Model (SCHISM) (Y.J. Zhang et al.

2016) are able to be run on HPC machines, often exploiting both distributed and shared paradigms. In the case of unstructured grid, though, the distributed memory parallelization addresses the problems of domain decomposition, that is not as trivial as in the case of regular grid, and of the semi-implicit methods, common among such models.

Chapter 3 gives deeper insights on both these issues.

## 1.8 Aims and Structure of the Thesis

The simulations outlined above, especially those carried in the last decade, are marked out by a high computational cost. This cost is motivated by several factors, such as the resolution, and hence the number of grid points in the mesh, the time step used and the length of the simulation. The 4 years simulated for cascading basins system in (E.V. Stanev, Grashorn, et al. 2017) has required around 4 weeks of machine time using 144 MPI processes with a time step of 90 seconds.

The global applications carried by (Danilov, Sidorenko, et al. 2017) and (M. R. Petersen, Asay-Davis, et al. 2019) need to be run on modern High Performance Computing (HPC) facilities in order to be accomplished in a reasonable time, since they need a massive parallelism with deployment of thousands of MPI processes. In these case the metric to measure the throughput <sup>1</sup> is Simulated Years per Day (SYPD) and it is 17 in the case of (Danilov, Sidorenko, et al. 2017) using 1728 processes.

State of the art numerical simulations, then, need to be endorsed by plentiful computational infrastructures.

The aim of this work is to set up a comprehensive simulation of the Southern European Sea as a unique model such as in study of (C. Ferrarin, Bellafiore, et al. 2018) for the MMBA system but including all the physics representable.

None of the modelling studies that have been outlined in the previous sections provides a comprehensive description of the SES system, with 3 main key points:

- A complete description of the problem, including the various processes such as water mass formation, representation of stratification, role of tides

---

<sup>1</sup>The throughput indicates the efficiency of a code in the context of a climatic simulation, measuring the rate of production of output

- Modelling the system in a seamless fashion, including the exchange of thermohaline properties among the various sub-basins of the SES system
- A study of the physical aspects of the SES system in a inter-annual or, decadal scale, aiming to a description of the system in a climatic perspective

For this purpose, in this work we explore the possibility of implementing a modelling system for the Southern European Seas based on the SHYFEM unstructured hydrodynamical model, including baroclinic contribution to the representation of dynamics and accounting for coastal-open water interaction with a proper resolution of coastal features. We acknowledge the necessity of massive computational power to achieve this and, notably, the necessity of adapting the code of SHYFEM to be run on state-of-the-art HPC facilities.

The thesis is organised as follows:

- In the Chapter 2 there is a complete description of the SHYFEM model, with some hints about the discretization of primitive equation and a description of the solution algorithm derived by the code tracking.
- The description of all the steps made to parallelize the SHYFEM code is in chapter 3, that shows the parallelization strategy for distributed memory architectures, the validation of the parallel outputs against the serial version using a realistic test case and some further optimizations.
- The chapter 4 describes the SES system modeling framework and shows the validation of 1 year of simulation against observational data.

# Chapter 2

## SHYFEM Description

### Contents

---

<b>2.1 The SHYFEM modelling System</b>	<b>36</b>
2.1.1 The SHYFEM Computation Grid	38
2.1.2 Model Equations	39
2.1.3 Sub-grid scale Physics - Turbulence Closure	41
2.1.4 Surface/Bottom boundary Conditions	44
2.1.5 Lateral Open Boundary Conditions	47
<b>2.2 Numerics - SHYFEM Solution Algorithm</b>	<b>48</b>
2.2.1 Advancement of Transports	50
2.2.2 Viscosity Operators	51
2.2.3 The Barotropic Equation	53
2.2.4 STEP 2 - Finalize Hydrodynamics	54
2.2.5 Vertical Velocities	54
2.2.6 Advection-Diffusion	55

---

### 2.1 The SHYFEM modelling System

SHYFEM is a modeling system built around its baroclinic thermo-dynamical core based on Discontinuous Galerkin (DG) finite element discretization of primitive

equations. The modeling system (<https://github.com/SHYFEM-model/shyfem>) is endowed with several features that renders it suitable for the study of coastal processes.

As stated in the previous section, SHYFEM has been formerly devised for tidal prediction and applied both in coastal/lagoon areas (Umgiesser, C. Ferrarin, et al. 2014) and in larger scale applications (C. Ferrarin, Bellafiore, et al. 2018).

The coupling of SHYFEM with the and the 3rd generation spectral wave model WWM (Wind Wave Model, (Roland 2008)) has been applied in the Gulf of Mexico and in the Adriatic Sea.

SHYFEM can simulate the erosion and sedimentation rates under either steady currents or the combined and time-dependent influence of waves and currents is external module SEDTRANS (Neumeier et al. 2008) is used. The transport of sediments from the Venice lagoon to the Adriatic has been modelled in (Christian Ferrarin et al. 2010).

Simulation that aim at study the quality of water can be performed by means of the modified version of EUTRO module of WASP (released by the U.S. Environmental Protection Agency (EPA) (Ambrose et al. 1993). It simulates the evolution of nine state variables in the water column and sediment bed, including dissolved oxygen (DO), carbonaceous biochemical oxygen demand (CBOD), phytoplankton carbon and chlorophyll a (PHY), ammonia (NH<sub>3</sub>), nitrate (NO<sub>3</sub>), organic nitrogen (ON), organic phosphorus (OP), orthophosphate (OPO<sub>4</sub>) and zooplankton (ZOO). An example of application of the EUTRO module in the Venice Lagoon is given in (Umgiesser, Melaku Canu, et al. 2003).

SHYFEM can also be run without considering the hydrostatic assumption. In this case, phenomena like dense water formation and inertia gravity waves are better captured by the model as stated in (Bellafiore, McKiver, et al. 2018).

The richness of external modules renders the SHYFEM modelling system versatile and suitable for studies of environmental that processes that develop in the open ocean - coastal zones interaction.

In this perspective, we want to extend the potential of this modelling system, with the aim of rendering it capable of running computationally demanding simulations.

In this work we focus on the hydrostatic hydrodynamical core of SHYFEM , updated for 3D computations by (Bellafiore and Umgiesser 2010) with the aim of

adapting the code to be run on distributed memory architectures. In this essay we refer to SHYFEM or its parallel version based on distributed memory approach, SHYMPI, indifferently.

In this chapter we describe the spatial discretization of model variables, we introduce the model equations and boundary conditions and, by means of a trough code tracking, we unfold the solution algorithm.

### 2.1.1 The SHYFEM Computation Grid

The horizontal space is discretized into a tassellation of irregular triangles joined by a connectivity list. The shared segments/grid points between elements are called edged / nodes respectively (see table (2.2) for the notation adopted in this report). Figure (2.1a) shows how model variables are located onto the unstructured grid. The horizontal transports and the sea level are staggered and located at element centroids and nodes respectively to have conservation of mass [reference]. Also active tracers, vertical velocities and turbulent quantities are node referenced. The dashed around the node in the center delimit the finite colume associated with the node.

The vertical grid is created from extrusion of the horizontal into  $nlv$  layers separated by  $nlv+1$  interfaces (see fig. 2.1b)

$$- \max(H(x, y)) = z_{nlv} < z_{nlv-1} < \dots < z_1 < z_0(x, y, t) = \eta(x, y, t) \quad (2.1)$$

where  $H(x, y)$  represents the depth of the ocean and the top most interface,  $z_0(x, y, t)$ , is the only allowed to move and corresponds to the free surface  $\eta$ . The thickness of these layers are defined as

$$h_l(e) = z_{l-1} - z_l \quad (2.2)$$

where we retain the index ( $e$ ) of the element because, in this z-model, the last active layer of each element is variable to represent the bottom topography (gray shade in fig. 2.1b). Variables are staggered also in the vertical grid. Horizontal transports and tracers are located at the centers of the cells, vertical velocities and turbulent quantities at the layer interfaces.

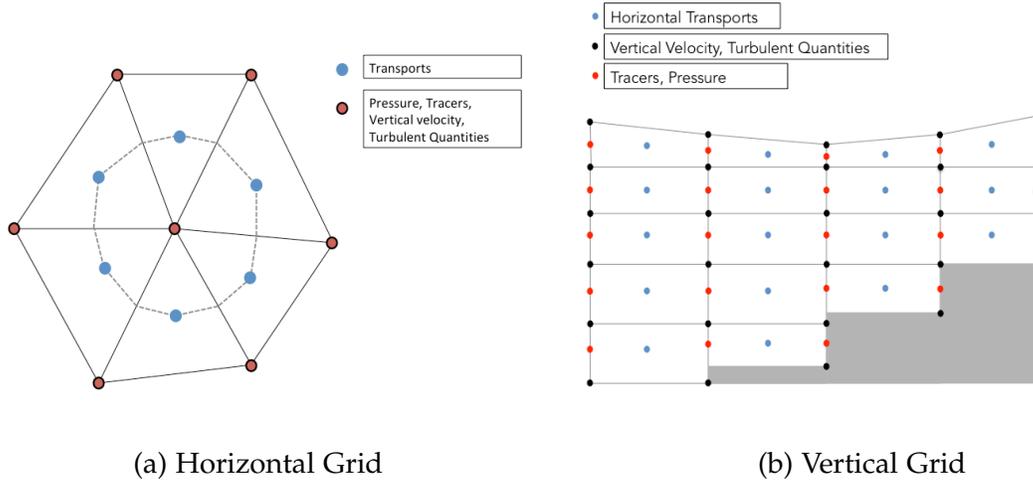


Figure 2.1: Disposition of Model Variables on the grid

## 2.1.2 Model Equations

SHYMPI equations are layers integrated, having transports ( $U, V$ ) as prognostic variables of the momentum. The prognostic equations for a generic layer  $l$  are for momentum (2.3),(2.4)

$$\begin{aligned} \frac{\partial U_l}{\partial t} + u_l \frac{\partial U_l}{\partial x} + v_l \frac{\partial U_l}{\partial y} + w_{l-1} u_{l-1} - w_l u_l = & -\frac{h_l}{\rho_0} \frac{\partial p_{atm}}{\partial x} - g h_l \frac{\partial \eta}{\partial x} - g \frac{h_l}{\rho_0} \sum_{k=l}^1 \frac{\partial \rho'_k}{\partial x} h_k + \\ & + (\tau_{xz}^{l-1} - \tau_{xz}^l) + A_H \nabla^2 U_l + f V_l \end{aligned} \quad (2.3)$$

$$\begin{aligned} \frac{\partial V_l}{\partial t} + u_l \frac{\partial V_l}{\partial x} + v_l \frac{\partial V_l}{\partial y} + w_{l-1} v_{l-1} - w_l v_l = & -\frac{h_l}{\rho_0} \frac{\partial p_{atm}}{\partial y} - g h_l \frac{\partial \eta}{\partial y} - g \frac{h_l}{\rho_0} \sum_{k=l}^1 \frac{\partial \rho'_k}{\partial y} h_k + \\ & + (\tau_{yz}^{l-1} - \tau_{yz}^l) + A_H \nabla^2 V_l - f U_l \end{aligned} \quad (2.4)$$

where the meaning of symbols and variables is listed in table (2.2). Stresses at tob/bottom interfaces of layer are denoted as  $\tau^{l-1}$  and  $\tau^l$ .

Equations (2.5),(2.6) represent the vertically integrated continuity equation and

the layer integrated continuity. The first is used prognostically to advance the solution of free surface  $\eta$  accounting for the contribution of mass fluxes at surface, the latter to diagnose vertical velocities.

$$\frac{\partial \eta}{\partial t} + \frac{\partial \bar{U}}{\partial x} + \frac{\partial \bar{V}}{\partial y} = P - E \quad (2.5)$$

$$\frac{\partial U_l}{\partial x} + \frac{\partial V_l}{\partial y} = w_l - w_{l-1} \quad (2.6)$$

with  $\bar{U}, \bar{V}$  the components of barotropic velocity.

The advancement of active tracers is computed in (2.7),(2.8). The equations are analogous, with the equation for the Temperature having a source term due to solar radiation penetration. Further details on the heat fluxes are provided in section (2.1.4)

$$\left. \frac{\partial(h_l S_l)}{\partial t} + \frac{\partial(U_l S_l)}{\partial x} + \frac{\partial(V_l S_l)}{\partial y} + (w_l S_l) \right|_{z_l}^{z_{l-1}} = \nabla \cdot (K_H \nabla(h_l S_l)) + \gamma_H \nabla^2(h_l S_l) + (K_V + \gamma_V) \left. \frac{\partial S_l}{\partial z} \right|_{z_l}^{z_{l-1}} \quad (2.7)$$

$$\left. \frac{\partial(h_l \theta_l)}{\partial t} + \frac{\partial(U_l \theta_l)}{\partial x} + \frac{\partial(V_l \theta_l)}{\partial y} + (w_l \theta_l) \right|_{z_l}^{z_{l-1}} = \nabla \cdot (K_H \nabla(h_l \theta_l)) + \gamma_H \nabla_h^2(h_l \theta_l) + (K_V + \gamma_V) \left. \frac{\partial \theta_l}{\partial z} \right|_{z_l}^{z_{l-1}} + \frac{h_l}{\rho_0 C_p} \left. \frac{\partial I}{\partial z} \right|_{z_l}^{z_{l-1}} \quad (2.8)$$

Pressure and density under hydrostatic assumption are computed in equations (2.9),(2.10).

$$\frac{\partial p_l}{\partial z} = -\rho_l g \rightarrow p_l = \rho_0 g (\eta - 1/2(z_{l-1} + z_l)) + \sum_{k=1}^{l-1} \rho'_k g h_k + 1/2 g \rho'_l h_l \quad (2.9)$$

equation term	weights of n+1/n time level	value
barotropic pressure in momentum	$a_M/\tilde{a}_M$	0.6/0.4
transport in continuity equation	$a_Z/\tilde{a}_Z$	0.6/0.4
vertical viscosity in momentum	$a_T/\tilde{a}_T$	1/0
vertical advection in AD equation	$a_A/\tilde{a}_A$	0/1
vertical diffusion in AD equation	$a_D/\tilde{a}_D$	1/0
Coriolis term	$a_F/\tilde{a}_F$	0.5/0.5

Table 2.1: Weights of time levels for Semi-implicit method with typical values used in SHYFEM simulation

$$\rho_l(x, y, 1/2(z_{l-1} + z_l), t) = \rho_l(S_l, \theta_l, p_l) \quad (2.10)$$

Where water density is computed with the UNESCO equation of state (Fofonoff and Millard 1983).

The equations of turbulence closure are described in (2.1.3)

The flow chart in fig (2.2) describes the solution method of SHYMPI. The step **1)** is the prediction of transports, where the transports at time level \* are computed, treating Coriolis semi-implicitly and vertical viscosity fully implicitly (eq. 2.35 and 2.36). In the step **2)** the equation for prediction of free surface is solved (eq. 2.52). The free surface at n+1 is used to correct the transports in the step **3)** (eq. 2.53). Details on the method of solution are provided in section 3.3.5. The vertical velocities are diagnosed in the step **4)** using the conservation of mass (eq. 2.54). The advection-diffusion equation for Temperature and Salinity is solved in step **5)** (2.55). Afterwards, the hydrostatic pressure and Temperature and Salinity at time level n+1 are used to calculate the new density field in step **6)** (eq. 2.9 and 2.10). Lastly, the vertical mixing coefficient are calculated in step **7)** (eq. 2.12). An extensive description of the solution method is given in (2.2).

### 2.1.3 Sub-grid scale Physics - Turbulence Closure

The effects of unresolved sub-grid scale fluid motions in the equation for the momentum are modelled according to (Smagorinsky 1963) (**subroutine smagorinsky**)

name	description	unit	index range
nel,nkn,nlv	total number of elements/nodes/layers	NA	NA
l	vertical layer index	NA	1:nlv
e	element index	NA	1:nel
k	node index	NA	1:nkn
$u_l(e),v_l(e)$	zonal/meridional velocity in element e at layer l	m/s	l=1:nlv,e=1:nel
$U_l(e),V_l(e)$	zonal/meridional transport in element e at layer l	$m^2/s$	l=1:nlv,e=1:nel
$w_l(k)$	vertical velocity at node k at interface l	m/s	l=0:nlv,k=1:nkn
$h_l(e)$	thickness of element e at layer l	m	l=1:nlv,e=1:nel
$\theta_l(k),S_l(k)$	Temperature/Salinity at node k at layer l	C/PSU	l=1:nlv,k=1:nkn
$\nu_l(k),K_V^l(k)$	vertical turbulent viscosity/diffusivity at node k at interface l	$m^2/s$	l=0:nlv,k=1:nkn
$\eta(k)$	free surface elevation	m	k=1:nkn
$\rho'_l(k)$	deviation of density from reference value	$Kg/m^3$	l=1:nlv,k=1:nkn
$\rho_0$	reference density	$Kg/m^3$	NA
$\nabla$	2D nabla operator	1/m	NA
$\Delta t$	model time step	s	NA
$f(e)$	Coriolis parameter in element e	1/s	e=1:nel
$A_H^l(e)$	Horizontal viscosity coefficient at element e at layer l	$m^2/s$	l=1:nlv,e=1:nel
$C_p$	Specific Heat at constant Pressure	J/Kg/K	NA
$g$	Gravity Acceleration	$m/s^2$	NA
$I$	Intensity of Solar radiation	$J/m^2/s$	*

Table 2.2: Notations adopted in this work

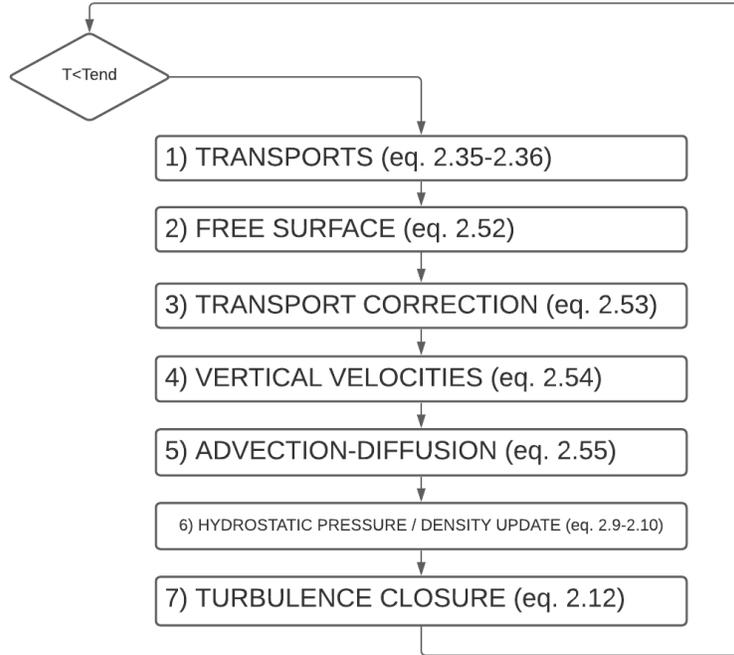


Figure 2.2: Simplified Diagram of Hydrostatic SHYFEM solution method

$$A_H = A \sqrt{2 \left( \frac{\partial u}{\partial x} \right)^2 + 2 \left( \frac{\partial v}{\partial y} \right)^2 + 2 \left( \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right)^2} \quad (2.11)$$

where  $A$  is the area of the grid cell. The Smagorinsky's eddy viscosity coefficient  $A_H$ , that appears in the equations (2.3) and (2.4), can be seen as the product of a characteristic grid size  $\sqrt{A}$ , bigger than the scales that are considered unresolved, times a characteristic velocity  $\sqrt{A} \sqrt{(\cdot)}$ .

Vertical mixing coefficients are calculated with the  $k - \epsilon$  model implemented in GOTM 1D model (Burchard and O. Petersen 1999) (**subroutine gotm\_shell**) following (Rodi 1987), where prognostic equations for dissipation rate of turbulent kinetic energy ( $\epsilon$ ) and turbulent kinetic energy ( $k$ ) are solved

$$\begin{aligned} \frac{\epsilon^{n+1} - \epsilon^n}{\Delta t} &= \frac{\partial}{\partial z} \left( \frac{\nu^n}{\sigma_\epsilon} \frac{\partial \epsilon^{n+1}}{\partial z} \right) + \frac{\epsilon^n}{k^n} \left( c_{\epsilon 1} P + c_{\epsilon 3} G - c_{\epsilon 2} \epsilon^n \right) \\ \frac{k^{n+1} - k^n}{\Delta t} &= \frac{\partial}{\partial z} \left( \frac{\nu^n}{\sigma_k} \frac{\partial k^{n+1}}{\partial z} \right) + P + G - \epsilon^n \end{aligned} \quad (2.12)$$

treating diffusional part fully implicitly.

$\sigma_k$	$\sigma_\epsilon$	$c_{\epsilon 1}$	$c_{\epsilon 2}$	$c_{\epsilon 3}$
1	1.3	1.44	1.92	1/-0.74

Table 2.3: Constants for the k- $\epsilon$  model according to (Rodi 1987)

Coefficients are set according to (Rodi 1987) in table (2.3). Subscripts  $n$  and  $n + 1$  indicate time levels and corresponds to the same time levels of SHYMPI model. The equations account for the production terms from velocity shear ( $P$ ) and buoyancy ( $G$ ) defined as

$$P = \nu \left[ \left( \frac{\partial u^{n+1}}{\partial z} \right)^2 + \left( \frac{\partial v^{n+1}}{\partial z} \right)^2 \right] \quad (2.13)$$

$$G = K_V \frac{g}{\rho_0} \frac{\partial \rho'}{\partial z} \quad (2.14)$$

where velocity at time level  $n+1$  are used and  $\nu, K_V$  denote the turbulent viscosity and diffusivity respectively.

Equations (2.12) are solved by simplified Gaussian elimination and vertical mixing at  $n+1$  is calculated from Kolmogorov-Prandtl relation

$$\nu^{n+1} = c_\mu \frac{(k^{n+1})^2}{\epsilon^{n+1}} \quad (2.15)$$

where the stability functions  $c_\mu$  are calculated with a quasi-equilibrium model following (Schumann and Gerz 1995).

## 2.1.4 Surface/Bottom boundary Conditions

### Momentum Fluxes

To perform a realistic simulation, SHYMPI needs to model air-sea interaction using information from an atmospheric model. The required fields are wind velocity in x/y [m/s], atmospheric pressure [Pa], air temperature and dew point temperature [C] and cloud cover [0-1]. SHYFEM offers several possibilities to compute air-sea fluxes, including COARE formulation (Fairall et al. 2003) and the same formulation used in POM model (Blumberg and Mellor 1987). In this section we point our attention on the bulk formulation used in the Mediterranean

Forecast System (MFS) (Pinardi, Allen, et al. 2003) (Tonani, Pinardi, Fratianni, et al. 2009), extensively described in (Pettenuzzo et al. 2010).

The boundary condition for the stress at the top of the ocean is

$$\vec{\tau}_{z=\eta} = C_D \rho_{air} |\vec{u}_{surf}| \vec{u}_{surf} \quad (2.16)$$

where  $\vec{u}_{surf}$  is the wind at air-sea interface provided by atmospheric model,  $\rho_{air}$  the density of air at ocean-atmosphere interface,  $C_D$  is the dimensionless drag coefficient calculated according to (Hellerman and Rosenstein 1983), with its polynomial form

$$\begin{aligned} C_D = & \alpha_1 + \alpha_2 |\vec{u}_{surf}| \\ & + \alpha_3 \Delta T + \alpha_4 |\vec{u}_{surf}|^2 \\ & + \alpha_5 \Delta T^2 + \alpha_6 |\vec{u}_{surf}| \Delta T \end{aligned} \quad (2.17)$$

where  $\Delta T = T_a - T_s$  the air-sea temperature difference and  $\alpha_1=0.934 \times 10^{-3}$ ,  $\alpha_2=0.788 \times 10^{-7}$ ,  $\alpha_3=0.868 \times 10^{-4}$ ,  $\alpha_4=-0.616 \times 10^{-6}$ ,  $\alpha_5=-0.120 \times 10^{-5}$  and  $\alpha_6=-0.214 \times 10^{-5}$ .

SHYFEM provides several options for the bottom stress formulation. The most relevant in large scale applications are the bottom stress with quadratic formulation

$$\vec{\tau}_{z=-H} = C_D |\vec{u}_{l_{max}}| \vec{u}_{l_{max}} \quad (2.18)$$

where  $-H$  is the position of the ocean bottom,  $\vec{u}_{l_{max}}$  is the norm of the current vector of last layer. The drag coefficient  $C_D$  has values of the order of  $10^{-2}$  or the logarithmic formulation

$$\vec{\tau}_{z=-H} = \left[ \frac{k}{\log\left(\frac{\lambda+0.5h_{l_{max}}}{\lambda}\right)} \right]^2 |\vec{u}_{l_{max}}| \vec{u}_{l_{max}} \quad (2.19)$$

where  $k$  is the Von Karman constant (0.4) and  $\lambda$  is the bottom roughness.

**Heat Fluxes (subroutine qflux3d)**

The dominant terms that control the heat exchange between the ocean and the atmosphere contribute to the net downward (positive for the ocean) heat flux  $Q_T$

$$Q_T = Q_S + Q_L + Q_E + Q_H \quad (2.20)$$

where  $Q_S$  is the net shortwave radiation flux,  $Q_L$  is the net longwave radiation flux,  $Q_E$  is the latent heat flux of evaporation and  $Q_H$  is the sensible heat flux.

$$Q_S = q_{surf}(1 - 0.62C + 0.0019\beta)(1 - \alpha) \quad (2.21)$$

with  $q_{surf}$  the clear sky radiation reaching the sea surface,  $C[0-1]$  the cloud cover,  $\beta$  the solar altitude in degrees at noon and  $\alpha$  the albedo.

$$Q_L = \epsilon\sigma T_S^4 - [\sigma T_A^4(0.653 + 0.00535e_A)](1 + 0.1762C^2) \quad (2.22)$$

with  $T_A$  the air temperature and  $e_A$  pressure of water vapor.

Fluxes of sensible and latent heat depend on sea-air difference in temperature ( $T_S - T_A$ )

$$Q_H = -\rho_A C_P C_H |\vec{u}_{surf}| (T_S - T_A) \quad (2.23)$$

and on the difference between the specific humidity of air ( $q_A$ ) and specific humidity saturated at  $T_S$  ( $q_S$ )

$$Q_E = -\rho_A L_E C_E |\vec{u}_{surf}| (q_S - q_A) = L_E E \quad (2.24)$$

where  $\rho_A$  is the density of moist air,  $C_P$  the specific heat capacity,  $L_E$  the latent heat of vaporization. The values of  $C_E$  and  $C_H$  are calculated according to (Kondo 1975).

The shortwave radiation,  $Q_S$ , consists of energy distributed across a wide spectral range. The ocean is strongly absorbing for wavelengths longer than 700 nm and these wavelengths contribute to heating the upper few tens of centimetres. The fraction of  $Q_S$  that resides in these almost non-penetrative wavebands,  $R$ , is  $\sim 58\%$ . It is assumed to penetrate the ocean with a decreasing exponential profile, with an e-folding depth scale,  $\lambda_1$ , of a few tens of centimetres (typically  $\lambda_1 = 0.35$

m). For shorter wavelengths (400-700 nm), the ocean is more transparent, and solar energy propagates to larger depths where it contributes to local heating. The way this second part of the solar energy penetrates into the ocean depends on which formulation is chosen. A decay law for  $Q_S$  that makes account for the shorter wavelengths is as follows

$$I(z) = Q_S [R e^{-z/\lambda_1} + (1 - R) e^{-z/\lambda_2}] \quad (2.25)$$

where  $\lambda_2$  is the second extinction length scale associated with the shorter wavelengths and it is usually chosen to be 23 m . The set of default values ( $\lambda_1 = 0.35\text{m}$ ,  $\lambda_2 = 23\text{ m}$ ,  $R = 0.58$ ) corresponds to a Type I water in Jerlov's (1968) classification (oligotrophic waters) (Jerlov 1968).

Diffusive fluxes at air-sea interface for Temperature and Salinity make account of the freshwater inputs from precipitation / evaporation budgets. The surface boundary conditions for the active tracers read

$$K_v \frac{\partial T}{\partial z} \Big|_{z=\eta} = T_{z=\eta} (E - P) - \frac{Q_S - Q_U}{\rho_0 C_p} \quad (2.26)$$

where  $Q_U = Q_L + Q_E + Q_H$  and

$$K_v \frac{\partial S}{\partial z} \Big|_{z=\eta} = S_{z=\eta} (E - P) \quad (2.27)$$

## 2.1.5 Lateral Open Boundary Conditions

In the case of simulation with open lateral boundaries, the model needs to be forced with fields of sea level, total velocities, temperature and salinity from an external (or parent) model.

The type of lateral boundary condition depends on the prognostic variable.

The water level is set at open boundaries nodes with a Dirichlet condition at the nodes  $k$  that belong to the boundary  $\partial$

$$\eta_{k \in \partial} = \eta^{\text{ext}} \quad (2.28)$$

while total velocities are relaxed in the outmost layer of grid cells  $e$  that belong to  $\partial$  with a correction factor in the momentum equation

$$-\frac{1}{\tau_{\text{nudge}}}(\vec{u}_m - \vec{u}_{\text{ext}})_{e \in \partial} \quad (2.29)$$

where  $\vec{u}_m$ ,  $\vec{u}_{\text{ext}}$  are the model and the external velocity respectively and  $\tau_{\text{nudge}}[\text{s}]$  the nudging parameter that regulates the entity of the relaxation of internal solution to the external. A big value of  $\tau_{\text{nudge}}$  brings small contribution from external velocity while a small (and  $> 0$ ) value relaxes the solution closer to  $\vec{u}_{\text{ext}}$ .

A Dirichlet boundary condition is set to a generic tracers  $T$  if the flow enters the domain, while a zero-gradient condition (based on the weighted average of closest internal values  $\langle T_{\text{int}} \rangle$ ) is imposed at the boundary nodes in case of outflow

$$\text{LBC on tracer } T = \begin{cases} T_{k \in \partial} = T_{\text{ext}} & \text{if } \vec{n} \cdot \vec{u}_{\text{ext}} < 0 \\ T_{k \in \partial} = \langle T_{\text{int}} \rangle & \text{if } \vec{n} \cdot \vec{u}_{\text{ext}} > 0 \end{cases} \quad (2.30)$$

where  $\vec{n}$  represents the normal direction to the boundary segment pointing outwards.

## 2.2 Numerics - SHYFEM Solution Algorithm

Considering a forward time differencing with two time levels  $n$  and  $n+1$ , SHYFEM treats some terms in the prognostic semi-implicitly, giving weights to the 2 time levels. For a generic variable  $A$ , naming  $\theta \in [0, 1]$  the weight of the  $n+1$  time levels, the semi-implicit treatment implies

$$A = \theta A^{n+1} + (1 - \theta) A^n \quad (2.31)$$

As proposed by (Dukowicz and R. Smith 1994), a fraction  $\gamma \in [0, 1]$  of the surface pressure gradient in the momentum equation is evaluated at time level  $n+1$

$$\frac{\partial \mathbf{u}}{\partial t} = -g \nabla (\gamma \eta^{n+1} + (1 - \gamma) \eta^n) + \frac{\partial}{\partial z} \left( \nu \frac{\partial \mathbf{u}^{n+1}}{\partial z} \right) + \mathbf{F} \quad (2.32)$$

and a fraction  $\beta \in [0, 1]$  of the divergence of barotropic flow in both continuity

equation and vertically integrated continuity is evaluated also at time level  $n+1$

$$\beta \nabla \mathbf{u}^{n+1} + (1 - \beta) \nabla \mathbf{u}^n + \frac{\partial \mathbf{w}^{n+1}}{\partial z} = 0 \quad (2.33)$$

$$\frac{\partial \eta}{\partial t} = -\nabla \left( \beta \int \mathbf{u}^{n+1} dz + (1 - \beta) \int \mathbf{u}^n dz \right) + P \quad (2.34)$$

Where the term  $\mathbf{F}$  contains other contributions to momentum (Coriolis, horizontal viscosity, baroclinic pressure). The vertical viscosity in the momentum is commonly treated implicitly in numerical ocean models as it is in eq. (2.32) since the inversion of the tri-diagonal matrix associated to the system is not computationally demanding.

Interesting choices of  $(\beta, \gamma)$  are

- $(\beta, \gamma) = (1, 0)$  or  $(0, 1)$ ; the scheme is called explicit and big restrictions conditions the model time step to have stability with respect to external gravity waves
- $(\beta, \gamma) = (1, 1)$ ; fully implicit scheme, unconditionally stable for external gravity waves by damping their energy
- $(\beta, \gamma) = (0.5, 0.5)$ ; Crank-Nicholson scheme. Stable against external gravity waves and energy conserving in the case linear Free Surface without momentum advection (Campin et al. 2004)

The advancement of momentum and free surface from time level  $n$  to  $n+1$  is described in (Campin et al. 2004), justifying the prediction / correction step for the momentum equations.

Additionally, SHYMPI considers the semi-implicit treatment of Coriolis term, since it can produce instability in cases of too low friction. The weights assigned to time level  $n$  and  $n+1$  of this term in the ocean models that do semi-implicit treatment is commonly  $0.5/0.5$ , being the most accurate scheme in the representation of inertial waves (Wang and Ikeda 1995).

This section provides details on the semi-implicit method of SHYMPI.

## 2.2.1 Advancement of Transports

In the STEP I, the prognostic equations for transport ( $U, V$ ) [ $m^2/s$ ] are advanced from time level  $n$  to a temporary time level  $*$ . The main routine for the advancement of transports is the **hydro\_transports**. The routine contains other subcalls to calculate the bottom friction (**subroutine bottom\_friction**) and the terms treated explicitly (horizontal viscosity, baroclinic PG, advection) (**subroutine set\_explicit**).

The advancement is done by each element of the mesh along the entire water column (**subroutine sp256v\_intern**).

The real numbers ( $\in [0, 1]$ )  $a_F, a_T$  weight the new time level (denoted as  $*$ ) Coriolis force and vertical viscosity respectively. For each element and for a fixed depth layer  $l$ , the equations are (for simplicity we drop the index of the element and set  $U_l(e) = U_l$ )

$$\frac{U_l^* - U_l^n}{\Delta t} - a_F f V_l^* - \tilde{a}_F f V_l^n - \frac{\partial}{\partial z} \left( v^n \frac{\partial}{\partial z} \left[ a_T U_l^* + \tilde{a}_T U_l^n \right] \right) = -gh_l \frac{\partial \eta^n}{\partial x} + F_l^x \quad (2.35)$$

$$\frac{V_l^* - V_l^n}{\Delta t} + a_F f U_l^* + \tilde{a}_F f U_l^n - \frac{\partial}{\partial z} \left( v^n \frac{\partial}{\partial z} \left[ a_T V_l^* + \tilde{a}_T V_l^n \right] \right) = -gh_l \frac{\partial \eta^n}{\partial y} + F_l^y \quad (2.36)$$

where the stresses  $\tau^{l-1}/\tau^l$  at top/bottom interface of the layer  $l$  are expressed in terms of the transports to be treated implicitly.  $F_l^x, F_l^y$  contain the terms in the equations that are treated fully explicitly (atmospheric and baroclinic pressure gradient (PG), horizontal viscosity, horizontal / vertical advection, stress induced by waves).

The viscosity  $\nu$  [ $m^2/s$ ] is the sum of molecular, turbulent and rheological viscosity,  $f$  is the Coriolis parameter,  $h_l$  is the thickness in [ $m$ ] of the  $l$ -th layer,  $\eta^n$  the elevation of sea surface at time level  $n$ .

With some algebraic manipulations, equations (2.35),(2.36) can be expressed in terms of the tendency of momentum

$$\begin{aligned} \frac{U_l^* - U_l^n}{\Delta t} - \Delta t a_F f \left( \frac{V_l^* - V_l^n}{\Delta t} \right) - \Delta t a_T \frac{\partial}{\partial z} \left( v^n \frac{\partial}{\partial z} \left( \frac{U_l^* - U_l^n}{\Delta t} \right) \right) = \\ f V_l^n + \frac{\partial}{\partial z} \left( v^n \frac{\partial U_l^n}{\partial z} \right) - g h_l \frac{\partial \eta^n}{\partial x} + F_l^x \end{aligned} \quad (2.37)$$

$$\begin{aligned} \frac{V_l^* - V_l^n}{\Delta t} + \Delta t a_F f \left( \frac{U_l^* - U_l^n}{\Delta t} \right) - \Delta t a_T \frac{\partial}{\partial z} \left( v^n \frac{\partial}{\partial z} \left( \frac{V_l^* - V_l^n}{\Delta t} \right) \right) = \\ - f U_l^n + \frac{\partial}{\partial z} \left( v^n \frac{\partial V_l^n}{\partial z} \right) - g h_l \frac{\partial \eta^n}{\partial y} + F_l^y \end{aligned} \quad (2.38)$$

All the known terms are put in the right-hand side of the equations, including the barotropic PG, that is considered entirely in this STEP I. Renaming tendencies

$$\frac{U_l^* - U_l^n}{\Delta t} = X_l \quad \frac{V_l^* - V_l^n}{\Delta t} = Y_l \quad (2.39)$$

and all the known terms as  $\text{RHS}_l^x, \text{RHS}_l^y$  leads to

$$X_l - \Delta t a_F f Y_l - \Delta t a_T \frac{\partial}{\partial z} \left( v^n \frac{\partial X_l}{\partial z} \right) = \text{RHS}_l^x \quad (2.40)$$

$$Y_l + \Delta t a_F f X_l - \Delta t a_T \frac{\partial}{\partial z} \left( v^n \frac{\partial Y_l}{\partial z} \right) = \text{RHS}_l^y \quad (2.41)$$

## 2.2.2 Viscosity Operators

Vertical derivatives in the viscosity terms in equations (2.40),(2.41) introduce dependency between adjacent layers and the equations cannot be inverted trivially. The Coriolis and the advective terms, additionally, introduce dependency between zonal and meridional momentum, requiring contemporary inversion of the 2 equations.

The discretization of vertical viscosity terms is made with centered differences. Figure (2.3) shows the 3 generic layers  $l-1, l, l+1$  separated by interfaces  $l-1, l$ . The differentiation of stresses at layer  $l$  interfaces for the zonal transport (and similarly for the meridional transport) is

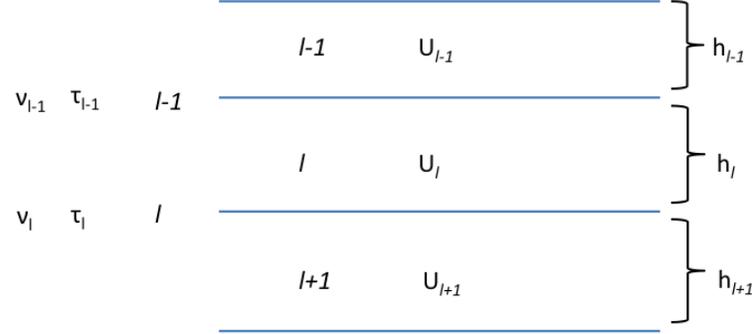


Figure 2.3: Index of vertical layers/interfaces used to define vertical viscosity operators

$$\frac{\partial}{\partial z} \left( \nu \frac{\partial U_l}{\partial z} \right) = \frac{\partial \tau_l}{\partial z} \rightarrow \frac{\tau_{l-1} - \tau_l}{h_l} \quad (2.42)$$

where the stresses  $\tau_{l-1}, \tau_l$  are functions of the transport shears

$$\tau_{l-1} = \nu_{l-1} \frac{U_{l-1} - U_l}{h_{l-1}/2 + h_l/2} \quad \tau_l = \nu_l \frac{U_l - U_{l+1}}{h_l/2 + h_{l+1}/2} \quad (2.43)$$

The vertical viscosity becomes

$$\Delta t a_T \left\{ \frac{2\nu_{l-1}}{h_l(h_{l-1} + h_l)} U_{l-1} - \frac{1}{h_l} \left[ \frac{2\nu_{l-1}}{(h_{l-1} + h_l)} + \frac{2\nu_l}{(h_l + h_{l+1})} \right] U_l + \frac{2\nu_l}{h_l(h_l + h_{l+1})} U_{l+1} \right\} \quad (2.44)$$

or, in more compact form

$$cc_l U_{l-1} - aa_l U_l + bb_l U_{l+1} \quad (2.45)$$

remarking the dependency on the upper  $(l - 1)$  and lower  $(l + 1)$  layers.

Introducing the definition  $\gamma = \Delta t a_F f$  and the factors for the discretization of vertical viscosity defined above, the system of equations (2.40),(2.41) reads



This equation cannot be solved since  $\eta^{n+1}$  is yet unknown, but can be inserted into eq. (2.49) to get

$$\eta^{n+1} - \delta \nabla (H \nabla \eta^{n+1}) = \eta^n - \delta \nabla (H \nabla \eta^n) - \Delta t \nabla (a_Z \mathbf{U}^* + \tilde{a}_Z) \mathbf{U}^n + \Delta t (P - E) \quad (2.51)$$

where  $\delta = g a_M a_Z \Delta t^2$  and  $H = \sum h_l$ . The equation that is actually solved in SHYMPI has the following shape (**subroutine hydro\_zeta**)

$$\eta^{n+1} + \delta \nabla (H \nabla \eta^{n+1}) = \eta^n + \delta \nabla (H \nabla \eta^n) + \Delta t \nabla (a_Z \mathbf{U}^* + (1 - a_Z) \mathbf{U}^n) + \Delta t (P - E) \quad (2.52)$$

, with different signs. The reason resides in the FEM formalism and the explanation goes beyond the scope of this work. The solution method for (3.3) is described in 3.3.5.

## 2.2.4 STEP 2 - Finalize Hydrodynamics

The solution of elliptic equation for free surface (eq. 3.3) allows to complete the advancement of transport applying eq. (2.53). The correction step for transport is carried in (**subroutine hydro\_transports\_final**) as follows

$$\begin{aligned} U_l^{n+1} &= U_l^* - g \Delta t a_M h_l \frac{\partial(\eta^{n+1} - \eta^n)}{\partial x} \\ V_l^{n+1} &= V_l^* - g \Delta t a_M h_l \frac{\partial(\eta^{n+1} - \eta^n)}{\partial y} \end{aligned} \quad (2.53)$$

where  $a_M$  is better known in the literature as  $\gamma$  (Campin et al. 2004) and weights the term that removes fast modes from momentum equation.

## 2.2.5 Vertical Velocities

In the case of hydrostatic model the vertical velocities are diagnosed by the continuity equation. This is done in **subroutine hydro\_vertical**. A volume equation for the layer  $l$  is first solved

$$\frac{\partial Ah_l}{\partial t} + \left[ A \nabla \left( a_z \mathbf{u}_l^{n+1} + \tilde{a}_z \right) \mathbf{u}_l^n \right] + Q_l + A \left( w_{l-1}^{n+1} - w_l^{n+1} \right) = 0 \quad (2.54)$$

where  $A$  [ $\text{m}^2$ ] is the horizontal section of the finite volume around the node,  $Q_l$  is the mass flux [in  $\text{m}^3/\text{s}$ ] from the surface or from internal sources and  $w_{l-1}^{n+1}$ ,  $w_l^{n+1}$  are the upwards mass fluxes through the top and bottom interface of the layer  $l$  respectively. The first term, called tendency term, accounts for the volume variation of the layer between time level  $n$  and  $n+1$ , and it is known quantity since new layer thicknesses are calculated after routine **hydro\_zeta**. In the case of  $z$ -coordinate model, only the upper layer volume is allowed to change.

Equation (2.54) is integrated from the bottom, using  $w_{\text{imax}}^{n+1} = 0$  as boundary condition, and the resulting values are divided by  $A$ . The mass flux across the upper surface ( $w_0$ ) is set to 0.

## 2.2.6 Advection-Diffusion

The advection-diffusion equation for any kind of tracer  $T$  is solved after hydrodynamics in **subroutine conz3d\_omp** (routine adapted to be compliant with OMP directives). The workload of solution is separated into 2 further subroutines **conz3d\_element** and **conz3d\_nodes**. The first builds the vertical system of the semi-implicit treatment of both vertical advection and diffusion <sup>1</sup>, the second actually solves the system to get the tracer at time level  $n+1$ .

The advection-diffusion equation for a generic tracer  $T$  at layer  $l$  and node  $k$  is (we omit the node index setting  $T_l(k) = T_l$ )

$$\begin{aligned} \frac{h_l^{n+1} T_l^{n+1} - h_l^n T_l^n}{\Delta t} = & - \left( \frac{\partial \bar{U}_l T^n}{\partial x} + \frac{\partial \bar{V}_l T^n}{\partial y} \right) - a_A (w_t T_t - w_b T_b)_l^{n+1} - \tilde{a}_A (w_t T_t - w_b T_b)_l^n \\ + K_H \left( \frac{\partial h_l^n T_l^n}{\partial x^2} + \frac{\partial h_l^n T_l^n}{\partial y^2} \right) + a_D \frac{\partial}{\partial z} \left( (K_V^n + \gamma) \frac{\partial h^{n+1} T^{n+1}}{\partial z} \right) + \tilde{a}_D \frac{\partial}{\partial z} \left( (K_V^n + \gamma) \frac{\partial h^n T^n}{\partial z} \right) + Q \end{aligned} \quad (2.55)$$

<sup>1</sup>in SHYFEM the semi-implicit treatment of vertical advection is possible only in case of upwind scheme

where

$$\begin{aligned}\bar{U}_l &= a_Z U_l^{n+1} + \tilde{a}_Z U_l^n \\ \bar{V}_l &= a_Z V_l^{n+1} + \tilde{a}_Z V_l^n\end{aligned}\tag{2.56}$$

and  $Q$  represents source/sink terms,  $a_A, a_D$  the weights of new time level for vertical advection and diffusion terms,  $K_H$  horizontal diffusion coefficient,  $K_V$  and  $\gamma$  turbulent and molecular diffusivities respectively. In a similar fashion to what done in STEP I of hydrodynamics for the discretization of vertical diffusion operators, the system of equations in (2.55) can be written in matrix form

$$A\mathbf{T}^{n+1} = b\tag{2.57}$$

where  $A$  is 3-diagonal matrix,  $b$  the sum of explicit terms and  $\mathbf{T}^{n+1} = [\dots T_{l-1}^{n+1}, T_l^{n+1}, T_{l+1}^{n+1} \dots]$  the solution vector. The system has size  $l_{\max} \times l_{\max}$  ( $l_{\max}$  is the number of active layers) and is solved with Thomas algorithm for tridiagonal matrices in routine **conz3d\_nodes**.

# Chapter 3

## Optimization and MPI Parallelization of SHYFEM

### Contents

---

<b>3.1</b>	<b>The Need of SpeedUp Simulations . . . . .</b>	<b>58</b>
<b>3.2</b>	<b>Dependencies Between Physical Variables on the Numerical Grid . . . . .</b>	<b>60</b>
<b>3.3</b>	<b>Parallelization Approach . . . . .</b>	<b>62</b>
3.3.1	Domain Decomposition . . . . .	63
3.3.2	Domain Decomposition Strategy . . . . .	63
3.3.3	The Domain Decomposition Algorithm . . . . .	65
3.3.4	MPI Directives . . . . .	66
3.3.5	Free Surface Solver . . . . .	69
3.3.6	I/O Management in MPI . . . . .	75
<b>3.4</b>	<b>Test Case for Validation of MPI Code . . . . .</b>	<b>76</b>
<b>3.5</b>	<b>Results . . . . .</b>	<b>78</b>
3.5.1	Performance Assessment . . . . .	80
<b>3.6</b>	<b>Extension of Restart Capabilities . . . . .</b>	<b>86</b>
<b>3.7</b>	<b>Model Diagnostics . . . . .</b>	<b>86</b>

---

This chapter unfolds the steps followed to parallelized the code of SHYFEM.

The code developments mentioned in this chapter are the result of the joint collaboration between CMCC Advanced Scientific Computing (ASC) Division and Ocean Prediction and Applications (OPA) Division.

### 3.1 The Need of SpeedUp Simulations

Last decades have seen a remarkable increase in the amount and accuracy of ocean forecast and analysis products, demanded by applications that span from military to environmental. These products benefit from the high resolution (either in time and in space) of numerical predictions but, the higher the resolution is, the more is the computational cost. In chapter 1.8 some examples of highly computationally demanding simulation were given. The computational cost is determined, in first instance, by the horizontal resolution, that is chosen depending on the physical process that are meant to be described.

Numerical simulations for global applications with the purpose of representing the mesoscale eddy activity having resolution of nearly  $1/4$  (Marsh et al. 2009) or  $1/16$  (Iovino et al. 2016) (commonly denoted in the literature as "eddy permitting" and "eddy resolving" simulations) are extremely expensive in terms of computational cost, prohibitive for conventional machines, and need to be run on state-of-the-art parallel architectures.

The development of ocean (or atmospheric) numerical modeling is strictly linked, then, with the advances of computational capabilities.

The main methods for numerical discretization of equations are Finite Differences (FDM) and Finite Elements (FEM). The first is the most straightforward way to discretize primitive equations, and it is suitable for large scale simulations (e.g. global) while the latter suits more regional applications, especially those that deal with complex coastline and bathymetry.

Finite Element's main feature is the flexibility of numerical mesh, which can be finer/coarser depending on where higher/lower resolution is needed.

Finite Elements ocean numerical modeling ever-increasing computational demand can be supported by the development of parallel architectures, that aim at reducing the time-to-solution of costly simulations. Here comes the need of

developing parallel code for finite element modeling or adapting existing FEM models for parallel execution.

The aim of this work is to parallelize SHYFEM (Umgiesser, Canu, et al. 2004), a 3D baroclinic shallow water hydrodynamic model, whose capabilities can be extended with the use of external module for sediment transport, ecological parameters, turbulence closure. SHYFEM uses a semi-implicit algorithm for integration in time, which combines the advantages of the explicit and the implicit schemes, and allows the user to select the time integration scheme.

SHYFEM has been successfully deployed on the short-term coastal ocean forecast system SANIFS (Federico et al. 2017) whose accuracy in representing either shelf-scale and coastal-harbour scale processes has been verified by comparison with observations data.

The ever more demanding needs of coastal oceanography for better accuracy of forecast products introduces some computational issues, in particular when very high-resolution grids are used. Indeed, memory requirements increase with the resolution as well as the total sequential execution time that could be prohibitive on some configurations.

Modern parallel architectures, organized into many-cores nodes, each one with its own memory capability, present two considerable advantages. If the problem is designed to be distributed into sub-problems, each part can be run independently from the other and its memory usage is proportional to its size. This creates noticeable performance speedup and the possibility to split the great memory usage of a complex problem over multiple nodes. These concepts are addressed as performance and memory scalability.

One parallel version of SHYFEM based on OpenMP exists (Eric Pascolo 2015) . However, the shared memory paradigm does not address the problem of memory scalability and inter-node computational scalability, so here comes the need of developing a version of the code to be executed on distributed architectures.

In the unstructured community, several MPI codes are freely available: MPAS (Ringler et al. 2013) and FESOM22 (Danilov, Sidorenko, et al. 2017) are devised for global applications and address the need of reproducing both mesoscale processes with the same quality of structured models and coastal processes in global simulations. MPI codes that address merely coastal processes include SLIM3D (Kärnä Tuomas et al. 2013), SCHISM (Y.J. Zhang et al. 2016), FVCOM

(Chen et al. 2003).

This chapter is organised as follows: section (3.2) introduces the SHYFEM model in and the dependencies among variables on the computational grid, the section (3.3) describes the methodology adopted to introduce the distributed memory parallelization, through the choice of the domain decomposition strategy and the understanding and handling of data dependencies among MPI processes and details on the MPI implementation. Section (3.4) introduces the test case to validate the MPI model against the serial version. The performance of the MPI code is assessed in section (3.5). Lastly, sections (3.6) and (3.7) report for enhancement of the code capabilities.

## 3.2 Dependencies Between Physical Variables on the Numerical Grid

Figure (3.1a) shows how the model variables are staggered over the computational grid. Horizontal transports ( $U, V$ ) are located in the element centers, while all the other are located on the vertexes (vertical velocity “ $w$ ”, elevation of free surface “ $\eta$ ”, scalars). Each vertex has a corresponding finite volume (delimited by dashed lines in fig 3.1a). The staggering of hydrodynamic variables is essential to have a mass conserving model (Umgiesser, Canu, et al. 2004) (Jofre et al. 2014) (Felten and Lund 2006).

Variables are staggered also in the vertical grid, as shown in figure (3.1b). The vertical grid, in a  $z$  model, is an extrusion of the horizontal grid along the vertical direction. The grid cells that can change their volume are the ones of the top-most layer. The number of active cells along the vertical direction depends on the sea depth.

Scalar variables (red) are staggered with respect vertical velocity (black), referenced in the middle and at layer interfaces respectively. The sea surface elevation is a 2D field defined only in the  $w$  points at surface.

The discretization of continuous prognostic equations in the FEM framework is realized my means of dependencies between the aforementioned variables.

The tendency in the transport equation depends on the Coriolis force, the viscosity and pressure gradients. Apart from the Coriolis term, defined in the

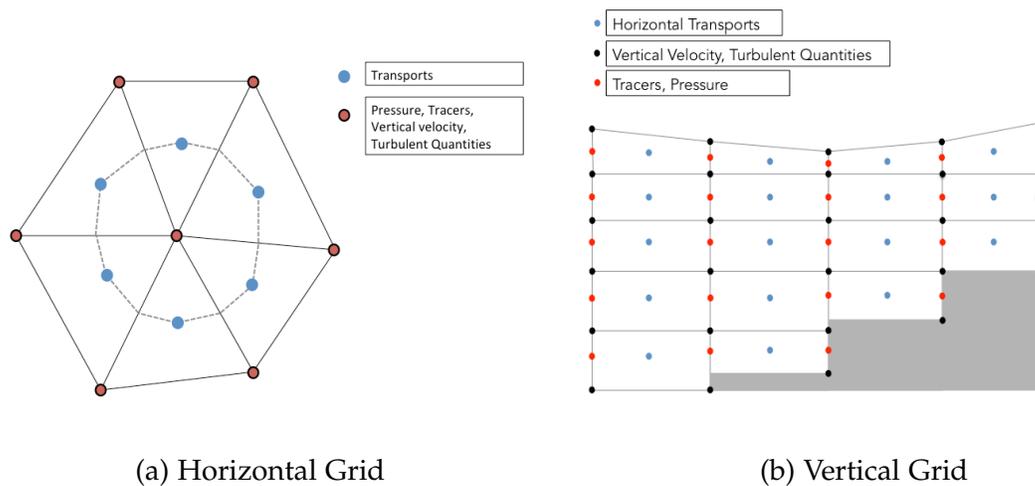


Figure 3.1: Disposition of Model Variables on the grid

center of grid cells as the transports, the other forcing terms depend on information of the surrounding variables.

The horizontal component of viscosity is realized as a viscosity operator where the contributions come from differences between the transport of current element  $U_i$  and transport  $U_j$  of surrounding elements  $j$  that share 1 side with it divided by the sum of areas  $Ae_i + Ae_j$  (see fig 3.5).  $A_H$  is the viscosity coefficient that follows Smagorinsky's formulation (see section 2.1.3). The viscosity operator consists then in an Element-to-Element dependency.

$$A_H \sum_{j=1,3} \frac{U_i - U_j}{1/2(Ae_i + Ae_j)} \quad (3.1)$$

The rate of change of transports depends on the differences of pressure around the element. The gradient of pressure field is calculated in the transport point using information of pressure on its 3 vertexes. The pressure gradient consists into a Node-to-Element dependency.

After the transports equations, the advancement of free surface elevation to the next time level comes from the solution of an elliptic equation that has variables defined only on vertexes. In this case the kind of dependency is Node-to-Node and the linear system arising from this equation is solved with the aid of Sparskit external package (Saad 1994).

The last step for the hydrodynamics is the diagnosis of the vertical velocity at the top of each finite volume using the continuity equation starting from the boundary condition imposed at the bottom. The conservation of the mass in the finite volume surrounding the node “*i*” (fig 3.1a) at level “*l*” comes from the balance of divergence of transport field around node *i* at the same level and the vertical velocity at interface *l*. The vertical velocity at interface *l*-1 is calculated in order to conserve the finite volume *i*. The divergence of transports is calculated in node *i* using the transport of the elements that have this node in common, consisting in a Elements-to-Node dependency.

Identifying the dependencies among the variables on the computational grid is the starting point to design a parallel code based on distributed memory. A more complete description of dependencies is given in section 3.3.2.

### 3.3 Parallelization Approach

Scientific and engineering numerical simulations show an ever-growing demand of computing resources due to the increasing of models resolution and complexity. Modern computer architectures are able to satisfy simulations requirements through a variety of computing hardware, often combined together into heterogeneous architectures. The design of the parallel application and the choice of the parallel paradigm/s are key factors in order to exploit computing facilities. Shared and distributed parallel programming can be mixed to better exploit heterogeneous architectures ( MPI+X allows to execute the code on clusters of NUMA nodes equipped with CPUs, GPUs, accelerators, ... ). The choice should be done taking into account the main features of the two paradigms: the shared memory approach allows multiple processing units to share data but it does not allow to scale the problem on more than one computing node, setting an upper bound to the available memory. The distributed memory approach, on the other side, allows each computing process to access its own memory space, so that bigger problems can be addressed by scaling the memory over multiple nodes. However, communications among parallel processes are needed to satisfy data dependencies. In a future perspective, marked out of problems that will be more demanding of memory and computing capabilities, the strategy developed to

parallelize the code of SHYFEM is based on the distributed memory approach and the Message Passing paradigm has been chosen to implement it. The parallelization strategy can be easily combined with existing shared memory solutions (i.e. the OpenMP version (E. Pascolo et al. 2016)) or other not yet implemented (e.g. OpenACC).

### 3.3.1 Domain Decomposition

The identification of data dependencies is a key factor for the design of the parallel algorithm, since inter-process communications have to be introduced to satisfy these dependencies.

In the case of a structured grid, each grid point usually holds information related to the cell discretized in the space and data dependencies are represented by a stencil containing the relations between each cell and its neighbours. For example, we could have five or nine points stencils which represent the dependencies of the current cell with regards to its four neighbours at north, south, east and west or that also cells along diagonals should be considered for computation.

On the other side, unstructured grid models can be characterized by dependencies among nodes (the vertexes), among elements (triangles) or among nodes and elements. These kinds of dependencies have to be taken into account when the domain decomposition strategy is defined.

### 3.3.2 Domain Decomposition Strategy

The choice between an element-based or node-based domain decomposition, graphically represented in Figure 3.2, aims at reducing data exchange among different processes.

The best domain decomposition strategy cannot be absolutely defined. Indeed, it usually depends on the code architecture and its implementation.

The analysis of the SHYFEM code shows that an element-based domain decomposition minimizes the number of communications among the parallel processes. Four types of data dependencies (graphically represented in Figure 3.3) can be identified within the code: *element-to-element* (A): the computation on each element depends on the three adjacent elements; *element-to-node* (B): the

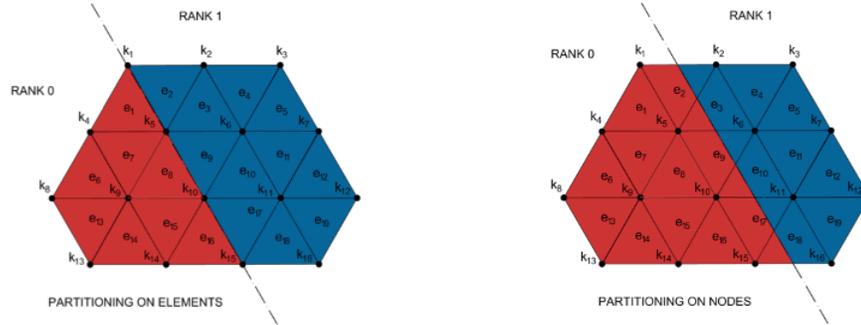


Figure 3.2: Element based domain decomposition and Node based domain decomposition.

node receives data from the incident elements (usually 6); *node-to-element* (C): the element needs data from its three nodes; *node-to-node* (D): the computation on each node depends on the adjacent nodes

The element-based domain decomposition needs data exchange when dependencies A, B and D happen, while node-based domain decomposition needs data exchange in case of dependencies C and D. The data dependency A happens only when transports are exchanged to compute the viscosity operator; the data dependency D is generated by the discretization of  $I + \delta \nabla (H \nabla)$  operator (with  $I$  identity matrix) applied to the predicted free surface  $\eta^{n+1}$  in equation 3.3. In the discrete form of equation 3.3 the value of  $\eta^{n+1}$  in a node depends on the value of  $\eta^{n+1}$  of the surrounding nodes, consisting in a *node-to-node* dependency and happens only in the context of FSE solution. The solution of FSE is obtained by means of iterative methods that see the discrete form of 3.3 in a matrix system perspective and compute the solution by means of matrix-vector products that realize the node-to-node dependency. In the context of parallel solution of FSE we use PETSc library to calculate the solution since it is devised for this kind of problems and contains both internal structures to store the matrix  $I + \delta \nabla (H \nabla)$  and the methods to calculate matrix-vector products also in the distributed memory case. Further details about the solution of FSE are provided in section (3.3.5); finally, the data dependency C is the most common in the code, more frequent than dependency B.

We can summarize that, after the analysis of the SHYFEM code, the element-based domain decomposition allows us to reduce the data dependencies that have

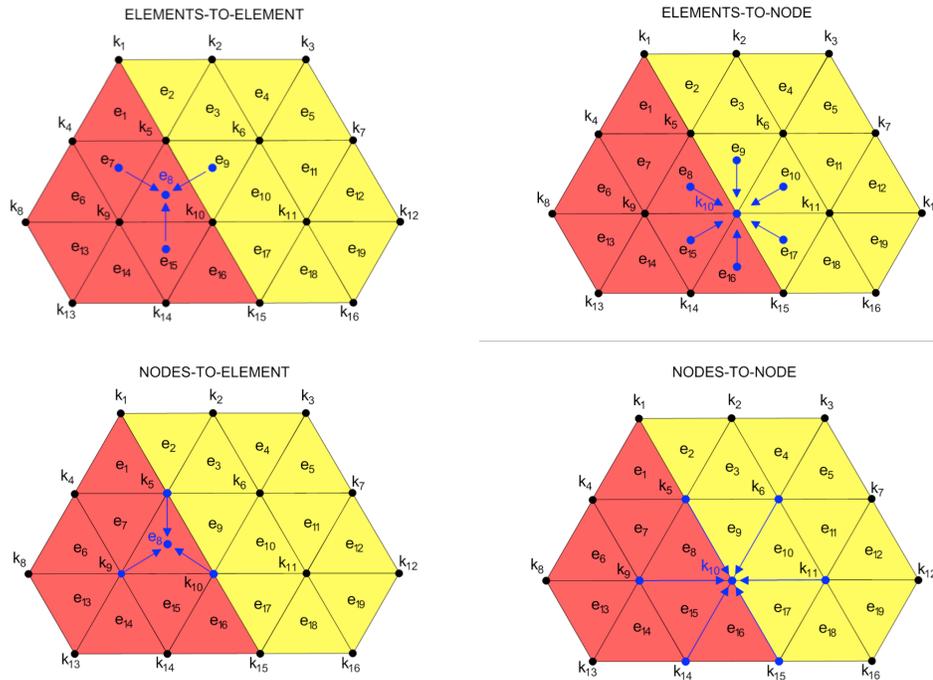


Figure 3.3: Data dependencies on the SHYFEM grid

to be solved through data exchanges among neighbouring processes. It is obvious that the computation on nodes shared among different processes is replicated.

### 3.3.3 The Domain Decomposition Algorithm

The second step, after the choice of the domain decomposition strategy, is the definition of the domain decomposition algorithm. It represents the way to distribute the workload among the processes for an efficient parallel computation. The standard approach (Hendrickson and Kolda 2000) is to consider the computational grid as a graph and to apply a graph domain decomposition strategy in order to distribute the workload. There are several parallel tools implementing solutions for the domain decomposition problem, such as ParMetis (G. Karypis 1997), the parallel extended version of Metis (G. Karypis 1999), Jostle (Walshaw and Cross 2007), PT-Scotch (Chevalier and Pellegrini 2008) and Zoltan (Hendrickson and Kolda 2000).

The scalability of the Zoltan PHG partitioner (Sivasankaran Rajamanickam

2012) has been a key factor for the choice of the domain decomposition tool to be used in the parallel version of SHYFEM. Moreover, the Zoltan library has been proved to be a good choice to simplify the development and to improve the performance of the parallel applications based on geometrically complex grids. The Zoltan framework includes parallel domain decomposition algorithms, data migration, parallel graph colouring, distributed data directories, unstructured communication services and memory management packages. It is available as open source software. An offline static partition module has been designed and implemented. It is executed once before the beginning of the simulation, when the number of parallel processes has been decided. The criteria that lead the domain decomposition phase aim at minimizing the inter-processes edge cuts and the difference among the workload assigned to the different processes. The total workload per process is computed taking into account that the number of vertical layers of each element of the grid is variable from element to element. A parametrization of the workload has been introduced since some branches of the code involve only the surface layer while others involve the whole water column. Figure 3.4 shows an example of domain domain decomposition obtained with Zoltan library.

### 3.3.4 MPI Directives

The SHYFEM code has a modular structure, and allows the user to customize the execution changing the parameters defined within a configuration file (i.e. namelist), to set up the simulation, and to activate the different modules that solve hydrodynamics, thermodynamics, turbulence.

This section describes in detail the changes made on the original code, introducing the additional data structures needed to handle the domain domain decomposition, the MPI point-to-point and collective communications, the solution of the FSE by using the external PETSc library and the I/O management.

#### Local-Global mapping

The domain decomposition over several MPI processes introduces the need to map the information of local to global entities. The entities to be mapped are the elements and the nodes. As consequence of the domain decomposition procedure,

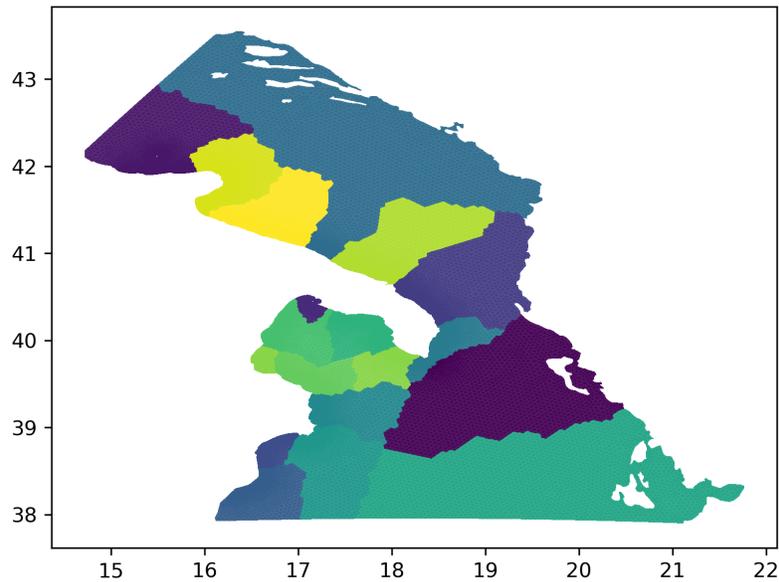


Figure 3.4: Example of domain decomposition in 18 MPI processes of the domain of SANIFS configuration (Federico et al. 2017) obtained with Zoltan

each process holds two mapping tables, one for the elements and one for the nodes. The mapping table of the elements stores the correspondence between the global identifier of the element (which is globally unique) with a local identifier of the same element which is locally unique. The same happens also with the mapping table associated to the nodes. Mapping information are stored into two local data structure, containing the Global Identification Number (GID) of respectively elements and nodes. The order of elements GIDs into local structures is natural, namely they are set in ascending order of GID. The local-global mapping is represented by the position into the local structure of the GIDs, called Local Identification Number (LID).

The GIDs of the nodes are stored in a different order. The GIDs of the nodes which belong to the edge of the local domain are stored at the end of the mapping table, this choice leads to some computational benefits: it is easy to identify all of the nodes at the edge; in most cases it is more performant executing first the computation over all of the nodes in the internal domain and after the computation over the nodes at the edge; during the data exchange it is easy to identify which nodes have to be sent and which have to be updated with the data coming from the neighbouring processes.

### Data exchange

Data exchanges are executed when Element-to-node and Element-to-element dependencies happen and MPI point-to-point communications are used. In the first case, each process receives information defined on the elements which share the target node from the processes the elements belong to. It keeps track of the shared nodes in terms of number and LID. Each process computes the local contribution and sends it to the interested neighbouring processes. The received information are stored in a temporary 3D data structure defined on nodes, vertical levels and processes. A reduction operation is performed on received information.

The Element-to-element dependency happens only once in the time loop to compute the viscosity operator. The term that appears in eq. (2.3) (and, in the same fashion, for the  $V$  transport) can be formally written as follows, without considering the dependency on levels for simplicity

$$2A_H(\mathbf{e}) \sum_{j=1,3} \frac{U(e_j) - U(\mathbf{e})}{A(e_j) + A(\mathbf{e})} \quad (3.2)$$

The viscosity operator for the element  $\mathbf{e}$  is the product of horizontal viscosity coefficient  $A_H(e)$  of the element (see 2.1.3 for the definition of  $A_H(e)$ ) times a term that mimicks the 2nd derivative, accounting for the transports  $U(e_j)$  and areas  $A(e_j)$  of the elements that share an edge with element  $\mathbf{e}$  (see fig. (3.5)).

If the element  $\mathbf{e}$  lies on the border separating its process from an adjacent one, the transports from one surrounding element is not available.

In this case, each process sends its contribution to the neighbours in terms of transport values. It keeps track of the elements to be sent/received in terms of number and LID, by using two different data structures. Only in this case, data structures are allocated including an overlap zone where the transport values are stored, as shown in Figure (3.6). Non-blocking point-to-point communications are used to overlap computation and communications.

Element-to-node dependency happens when fluxes into the finite volume associated to a node from surrounding elements need to be calculated. An example is the divergence of horizontal transport in the continuity equation (2.54). In the case of nodes that straddle the interface of two adjacent processes, each process calculates its share of the transport divergence in the finite volume (see

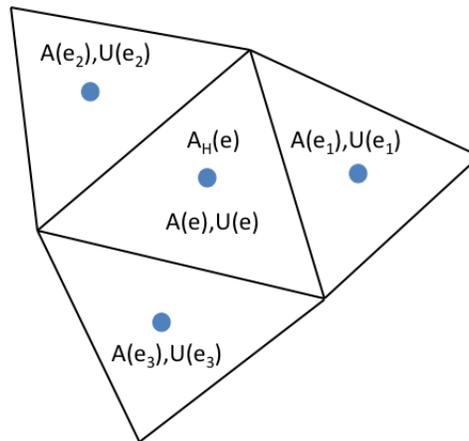


Figure 3.5: Description of variables to compute the horizontal viscosity operator in the central cell for zonal transport.  $A$  indicates the area of cells,  $U$  the zonal transports,  $A_H$  the viscosity coefficient

fig. (3.7)). A point-to-point communication is necessary so that all the processes that share these nodes have the whole value of divergence. Each process shares its contribution of the divergence on these nodes with the other neighbouring processes, receives the contributions from these and makes the sum. This operation is called briefly **exchange\_and\_sum**.

Finally, some MPI collective communications have been introduced to compute properties related to the whole domain (for instance to calculate the minimum or maximum temperature of the basin, or to calculate the total water volume). Algorithm (1), report an overview of sequential algorithm, SHYFEM, with respect to SHYMPI (2). The pseudo-code highlight also the communications steps implemented in the parallel algorithm.

### 3.3.5 Free Surface Solver

SHYMPI solves the following equation to predict the free surface elevation (**subroutine hydro\_zeta**)

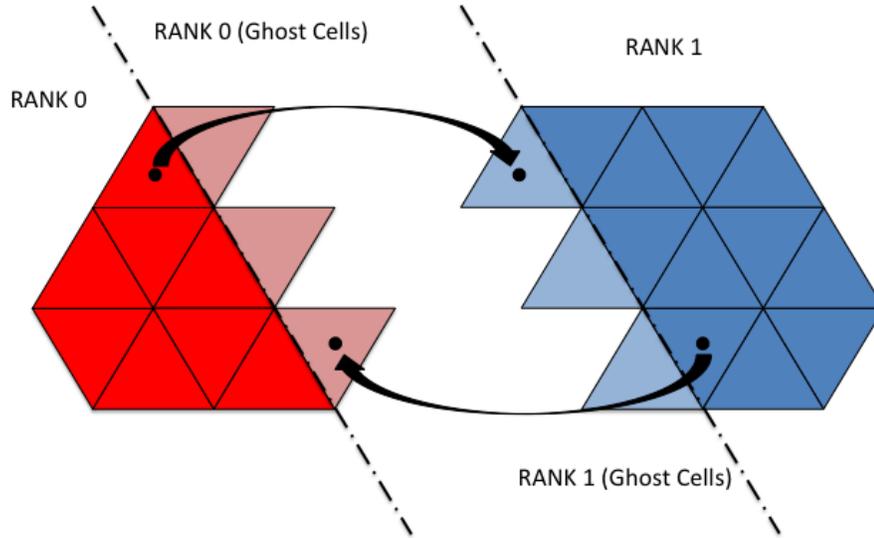


Figure 3.6: Communication pattern to satisfy Horizontal Viscosity Dependency - Faded colours indicate the ghost cells where transports from adjacent processes are stored to calculate horizontal viscosity operator

$$\eta^{n+1} + \delta \nabla (H \nabla \eta^{n+1}) = \eta^n + \delta \nabla (H \nabla \eta^n) + \Delta t \nabla (a_Z \mathbf{U}^* + \tilde{a}_Z \mathbf{U}^n) + \Delta t (P - E) \quad (3.3)$$

where where  $\delta = g a_M a_Z \Delta t^2$ ,  $H = \sum h_l$  and  $\mathbf{U}^* / \mathbf{U}^n$  the barotropic transports at time level \* and n respectively . The right hand side contains known terms and the equation (3.3) can be represented in matrix form

$$\mathbf{A} \eta^{n+1} = \mathbf{B} \quad (3.4)$$

where  $\mathbf{A}$  is the matrix of coefficients that arise from FEM discretization of derivatives of the left-hand side of eq. (3.3), with size  $n_{kn} \times n_{kn}$  and  $\mathbf{B}$  is the vector of the right-hand side of eq. (3.3). The matrix  $\mathbf{A}$  is non-singular with irregular sparse structure.

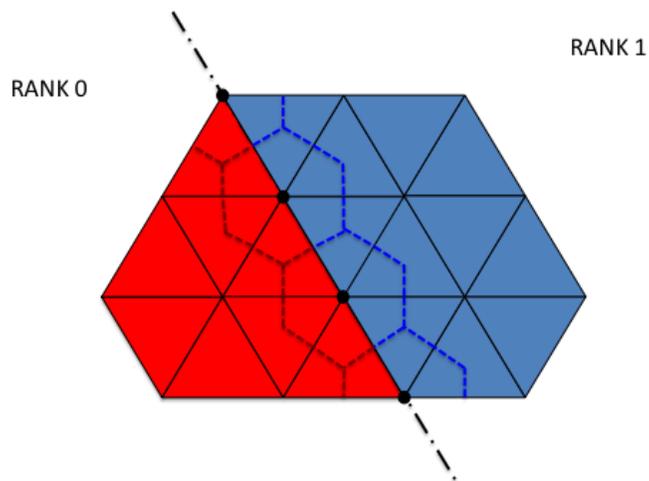


Figure 3.7: Communication pattern to satisfy Elements-to-node dependency. Black dots represents nodes shared among adjacent MPI processes, each surrounded by its finite volume (dashed lines).

---

**Algorithm 1:** SHYFEM Time loop

---

```

while  $t < num\_timesteps$  do
  SetExplicitTerms
  UpdateTransports
  SolveBarotropicEquation (SPARSKIT interface 3.8)
  CompleteTransportsUpdate
  CalcVerticalVelocity
  SolveAdvectionDiffusion
  CalcTurbulenceClosure
end while

```

---

The most convenient methods to solve large sparse system with  $\mathbf{A}$  not having particular structure properties are the iterative ones since direct inversion of  $\mathbf{A}$  would be too much expensive. These methods seek for an approximate solution for (3.4) and include Jacobi, Gauss-Seidel, Successive Over Relaxation (SOR) and Krylov Subspace Methods (KSP), the latter being considered one of the most important class of numerical methods.

Algorithms based on KSP search for an approximate solution in the space generated by the matrix  $\mathbf{A}$

$$\mathcal{K}_m(A, v_0) = span\{v_0, Av_0, A^2v_0, \dots, A^{m-1}v_0\} \quad (3.5)$$

called  $m$ th order Krylov subspace where  $v_0$  is an arbitrary vector (generally the right hand side of the system) with the property that the approximate solution  $x_m$  belongs to this subspace. In the iterative methods based on KSP the subspace  $\mathcal{K}_m(A, v_0)$  is enlarged a finite number of times ( $m$ ), where  $x_m$  represent an acceptable approximate solution, giving a residual  $r_m = \mathbf{B} - \mathbf{A}x_m$  that has the norm smaller than a certain tolerance.

In a parallel application each one of the  $m$  iterations is marked out by a computational and a communication cost to calculate the matrix-vector products in  $\mathcal{K}_m(A, v_0)$ , since both the matrix and the vector are distributed across the processes. A further cost is represented by the converge test, that generally is based on the euclidean norm of the residual

---

**Algorithm 2:** SHYMPI Time loop

---

**Require:**  $\mathcal{N}$  {Set of neighbouring processes}  
 $\mathcal{U}, \mathcal{V}$  {Set of nodes and edges defined in the current subdomain}  
**for all**  $p_i \in \mathcal{N}$  **do**  
    SendHalo( $\mathcal{U}, \mathcal{V}, p_i$ )  
**end for**  
**while**  $t < num\_timesteps$  **do**  
    **for all**  $p_i \in \mathcal{N}$  **do**  
        RecvHalo( $\mathcal{U}, \mathcal{V}, p_i$ )  
    **end for**  
    SetExplicitTerms  
    UpdateTransports  
    GlobalExchange(RHS)  
    SolveBarotropicEquation (PETSc interface 3.8)  
    CompleteTransportsUpdate  
    **for all**  $p_i \in \mathcal{N}$  **do**  
        SendHalo( $\mathcal{U}, \mathcal{V}, p_i$ )  
    **end for**  
    CalcVerticalVelocity  
    SolveAdvectionDiffusion  
    CalcTurbulenceClosure  
**end while**

---

$$\text{convergence test} = \begin{cases} |r_m| < a_{\text{tol}}, & \text{using absolute tolerance} \\ |r_m|/|B| < r_{\text{tol}}, & \text{using relative tolerance} \end{cases} \quad (3.6)$$

The calculation of the norm implies global communication to let all the processes to have the same value of the norm. Hence both point-to-point and global communication burden each one of the iteration, causing loss of efficiency for the parallel application if the number of necessary iterations is high. The number of iterations depends on the physical problem and on its size. An estimate of the problem complexity is given by the condition number  $C(\mathbf{A})$  that, for real symmetric matrices, is the ratio  $\max(\lambda)/\min(\lambda)$  between the maximum and minimum of the eigenvalues  $\lambda$  of  $\mathbf{A}$ . In general, the higher is  $C$ , the more iterations are necessary. The number of iterations, also depend on the tolerance desired, the smaller is the more are the iterations necessary.

In the case of complex systems it is convenient to modify the original linear system (3.4) to get a better Krylov subspace using a further matrix  $\mathbf{M}$ , called preconditioner, to search for an approximate solution in the modified system

$$\mathbf{M}^{-1}\mathbf{A}x = \mathbf{M}^{-1}\mathbf{B} \quad (3.7)$$

where  $\mathbf{M}^{-1} \approx \mathbf{A}^{-1}$  and is computed easily. We use the PETSc package (Balay, Gropp, et al. 1997),(Balay, Abhyankar, et al. 2020) to solve this equation efficiently.

We choose to use PETSc rather than implementing an internal parallel solver because PETSc is devised specifically to solve problems that arise from partial differential equations on parallel architecture and provides a wide variety of solver/preconditioners that can be switched through a namelist. In addition, the interface to PETSc is independent of the version, so it can be easily updated without changing the code.

Figure (3.8) shows the interfaces to the external packages for the native serial version of the program ( SHYFEM) and SHYMPI.

The interface to PETSc creates counterparts of  $\mathbf{A}$  and  $\mathbf{B}$  as objects of the package.  $\mathbf{A}$  is created as a sparse matrix in coordinate format (row,column,value) using the global ID (see fig. 3.9) of the nodes, in order to have the same non-zero pattern of the serial case regardless of the way the domain is decomposed. The same global ID are used to build the right-hand side  $\mathbf{B}$ .

To solve the free surface in SHYMPI we use Bi-Conjugate Gradient Stabilized Method (BiCGs) with incomplete LU (iLU) factorization as preconditioner. The convergence is reached when either the absolute tolerance (1e-18) or relative tolerance (1e-15) is reached.

The PETSc library uses parallel algorithm for solving the linear equations; the decomposition used inside PETSc is different from the domain decomposition used in SHYMPI (see fig. (3.10) and (3.11)) . PETSc divides the matrix  $\mathbf{A}$  in ascending order with the global ID, whilst, the SHYMPI domain decomposition is based on criteria that takes into account the geometry of the mesh and is, at any rate, different than PETSc. For this reason, after the approximate solution for  $\eta^{n+1}$  is found by PETSc, the solution vector is gathered by the master process and is redistributed across the MPI processes according to the model domain decomposition.

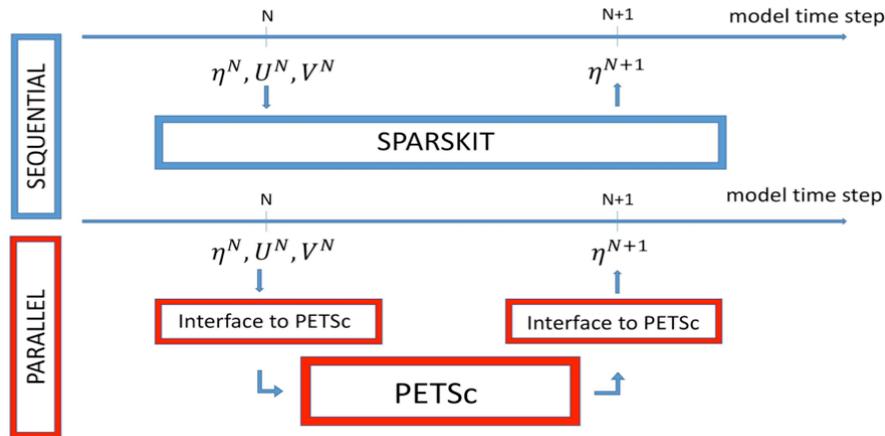


Figure 3.8: Interface to PETSc Library.

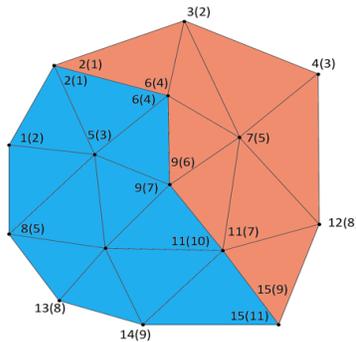


Figure 3.9: Example of partition in 2 MPI processes - Process 0 (Blue) and Process 1 (Red). Numbers represent Global (Local) ID

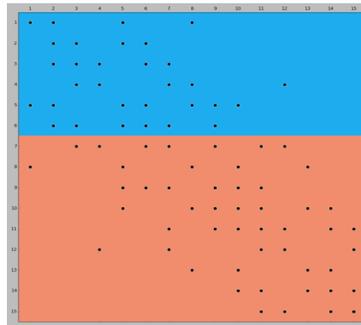


Figure 3.10: Partition of matrix A in PETSc - Rows of the matrix are sequentially appointed to the processes

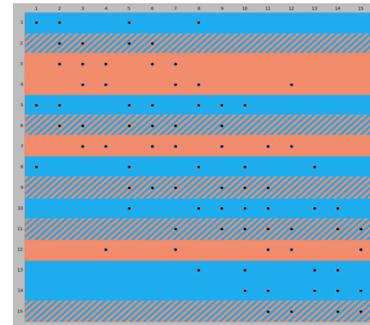


Figure 3.11: Partition of matrix A in SHYMPI - banded colors indicate rows of the matrix corresponding to shared nodes

### 3.3.6 I/O Management in MPI

I/O management usually represents a source of overhead since I/O operations often are executed sequentially. To avoid this, input and output files should be concurrently accessed by the parallel processes and each process should load its own data. However, loading the whole file for each process could affect the memory scalability. Indeed, the allocated memory should be independent by the number of parallel processes in order to ensure the memory scalability of the

code. The two issues can be addressed by distributing the I/O operations among the parallel processes. During the initialization step, SHYFEM needs to read two files: the basin geometry and the namelist. All the MPI processes perform the same operation and store common information. This phase is not scalable because each process browses the files. However, this operation is performed once and it has a limited impact on the total execution time. As second step, initial conditions and forcing (both lateral and surface) are accessed by all the parallel processes, but each one reads its portion of data, as shown in Figure 3.12. Surface atmospheric forcing is defined on a structured grid, so after reading the forcing file, each process will perform an interpolation of the data on the unstructured grid used by the model. The output writing could be performed by using external parallel libraries able to handle parallel I/O. However, up to know, SHYFEM does not use a standard format for I/O supported by parallel I/O libraries, then the adopted temporary solution is the collection of data from other processes to the master and a centralized management of the output operation. A check pointing mechanism has been implemented in the parallel version of the model. Usually, it is used when the simulation is divided into dependent chunks in order to save the status. Both phases (reading/writing) are performed in a distributed way among the processes to reduce the impact on the parallel speedup of the model, as shown in Figure 3.12; and namely each process generates its own restart file related to its sub-domain and reads its own restart file.

### 3.4 Test Case for Validation of MPI Code

We choose to run our experiments to assess the correctness of MPI implementation on the Southern Adriatic Northern Ionian coastal Forecasting System ( SANIFS) configuration that has horizontal resolution of 500 m near the coast to a minimum of 3-4 km in open waters. The total number of grid elements is 176331. Vertical resolution is 2 meters near the surface stepwise increasing towards the sea bottom, splitting the water column into 80 layers. For a detailed description of the model grid and the system settings see (Federico et al. 2017).

This experimental configuration is ideal for the comparison since is enough computationally intensive but not prohibitive for a serial run.

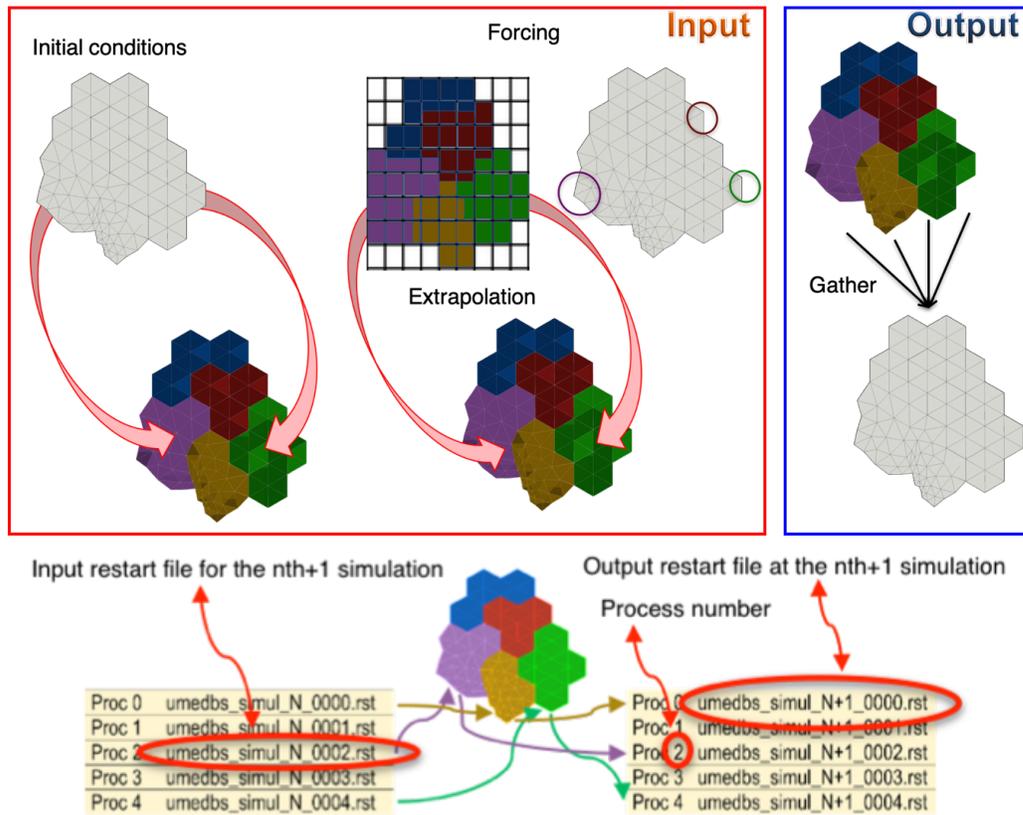


Figure 3.12: Data management related to the forcing files and restart files.

Runs are initialized with motionless velocity field and with Temperature and Salinity fields from CMEMS NRT products [<https://marine.copernicus.eu/>]. The simulations are forced hourly at lateral boundaries with fields of water level, total velocities and active tracers from the same products.

The sea level is imposed with a Dirichlet condition, while relaxation is applied to the parent model total velocities with a relaxation time of 1 hour. For tracers, a zero-gradient condition is applied if flow is outwards, while Dirichlet in case of inflow.

The boundary conditions for the upper surface follows the MFS bulk formulation (Pettenuzzo et al. 2010), that requires wind, cloud cover, air and dew point temperature, available in ECMWF analysis, with a temporal/spatial resolution of 6 hours/0.125 degrees respectively. From the same analysis we force the surface with data of precipitation.

We choose an upwind scheme for both horizontal and vertical advection in

the advection-diffusion equation. Bottom stress has a quadratic formulation. The time stepping for this simulation is fully implicit ( $a_M, a_Z = 1$ ).

We start the simulation at 2019-07-01T00:00:00 and we run for 1 days with a time step of 15 seconds, saving outputs every 1 hour. The model goes through a total of 5760 iterations.

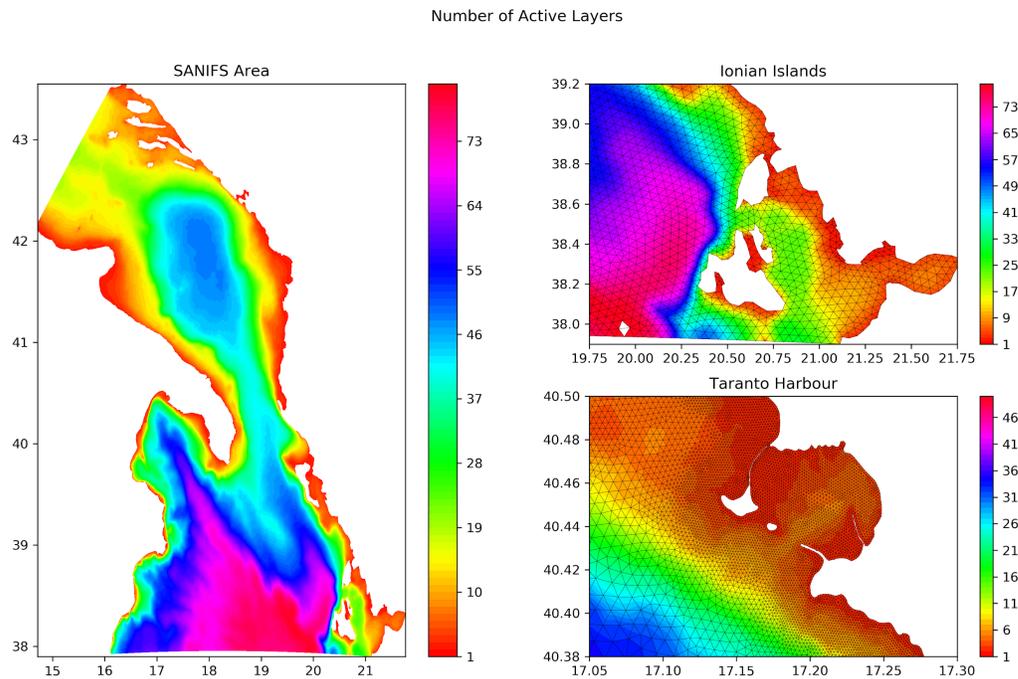


Figure 3.13: Active Layers in the domain of SANIFS configuration (Federico et al. 2017). The model grid resolution in harbors is around 50 meters (bottom-right panel)

### 3.5 Results

We have carried the same experiment using SHYMPI with 1 single process and on several domain decompositions (2,4,18,36,72,108,216). The simulations have been run on the ZEUS Super-Computer installed at CMCC HPC facility (see fig. (3.14) for the main characteristics of ZEUS architecture).

The outputs are saved every 1 simulated hour and the comparison is made using the informations of the native grid. We want assess the reproducibility

Compute node type:	Lenovo SD530
Processor:	Intel Xeon Gold 6154 (18 cores)
Processor Speed:	3,0 GHz
n° of processor per node:	2 (36 cores per node)
n° of nodes:	348
n° of processor cores:	12.528
Memory per node:	96 GB
Interconnection:	Infiniband EDR (100Gbps)
Peak Performance:	1.202 TFlops/s
Operating System:	Linux CentOS 7.6 x86_64
Batch Queuing System:	LSF v. 10.2

Figure 3.14: Main characteristics of ZEUS Super-Computer

of the result of the MPI code comparing the outputs of all decompositions with respect to the serial.

In order to evaluate the differences introduced by the parallelization, we chose a metrics the RMSE

$$\text{RMSE}(X, Y) = \sqrt{\frac{\sum_i ((X_i - Y_i)^2)}{N}} \quad (3.8)$$

We created timeseries of RMSE to compare the serial output of each variable with the MPI outputs from all the aforementioned decompositions. The RMSE timeseries have been calculated using NCO operators applied on the output on the native SHYMPI grid considering all the active cells in the domain.

Figure (3.15) shows the RMSE timeseries for all model variables (the comparison of meridional velocity not shown). The timeseries of all the MPI decomposition overlap, meaning that the program can reproduce the serial result closely to machine precision. The timeseries of the SSH are noisy but the differences against the serial remain steady and of the order of  $10^{-13}$  m. The reason for this behaviour is appointed to the non-associative properties of the sum of floating point numbers on a machine (Goldberg 1991). The coefficients of the matrix for free surface solution are built from the contribution of all elements surrounding a node. In the case of nodes shared across neighbour processes, this sum is delegated to PETSc, which does the sum with an order that is different both between different MPI decomposition and between 2 runs with the same decomposition, generating a non-deterministic behaviour. Forcing the solver to build the matrix coefficients

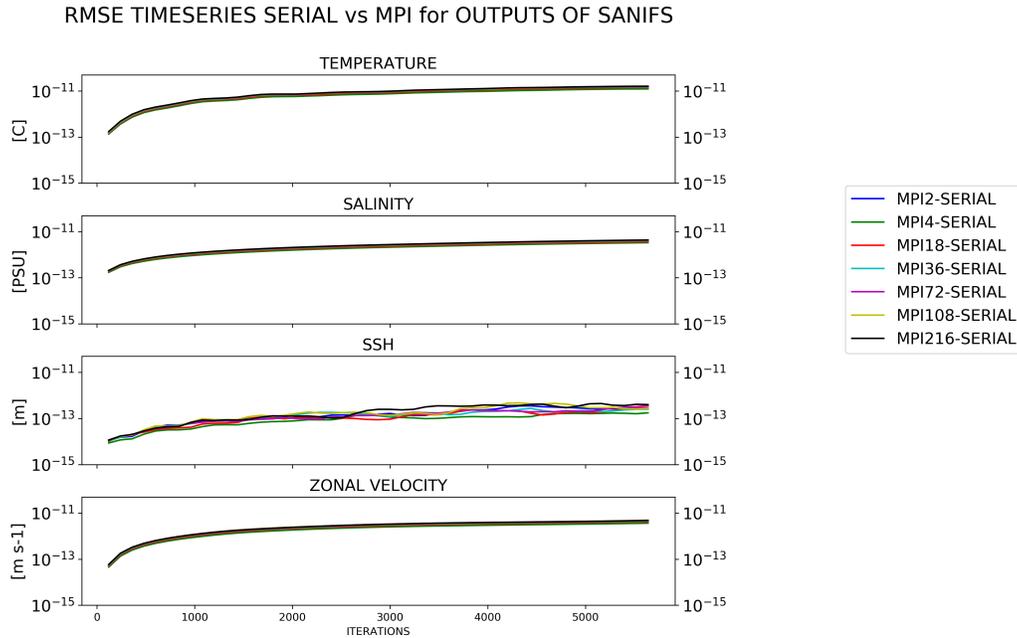


Figure 3.15: RMSEtimeseries of MPI outputs vs Serial for all variables

with the same order to eliminate this noise would lead to a loss of efficiency and is beyond our purposes. The SHYMPI program can reproduce the physical problem regardless the domain decomposition applied.

To further assess the non-deterministic behaviour of the free surface solver, we run the same simulation 5 times with the same number of processes. We use again the RMSE as metric to quantify the difference between 4 simulation with respect to the first that is taken as reference. Figure (3.16) shows the RMSE timeseries of the simulations run with 216 processes

The RMSE timeseries show that for each one of the model variables the deviations of the runs from the reference run remain close to the machine precision, with no effect on the reproducibility of the solution of the physical problem.

### 3.5.1 Performance Assessment

We make the assessment of the model performances using the common metrics of SpeedUp ( $S$ ) and Efficiency ( $E$ ). The aforementioned metrics quantify the performances (also referred as **scalability**) of a parallel application by means of comparison between the execution time of the serial program ( $T_s$ ) and the

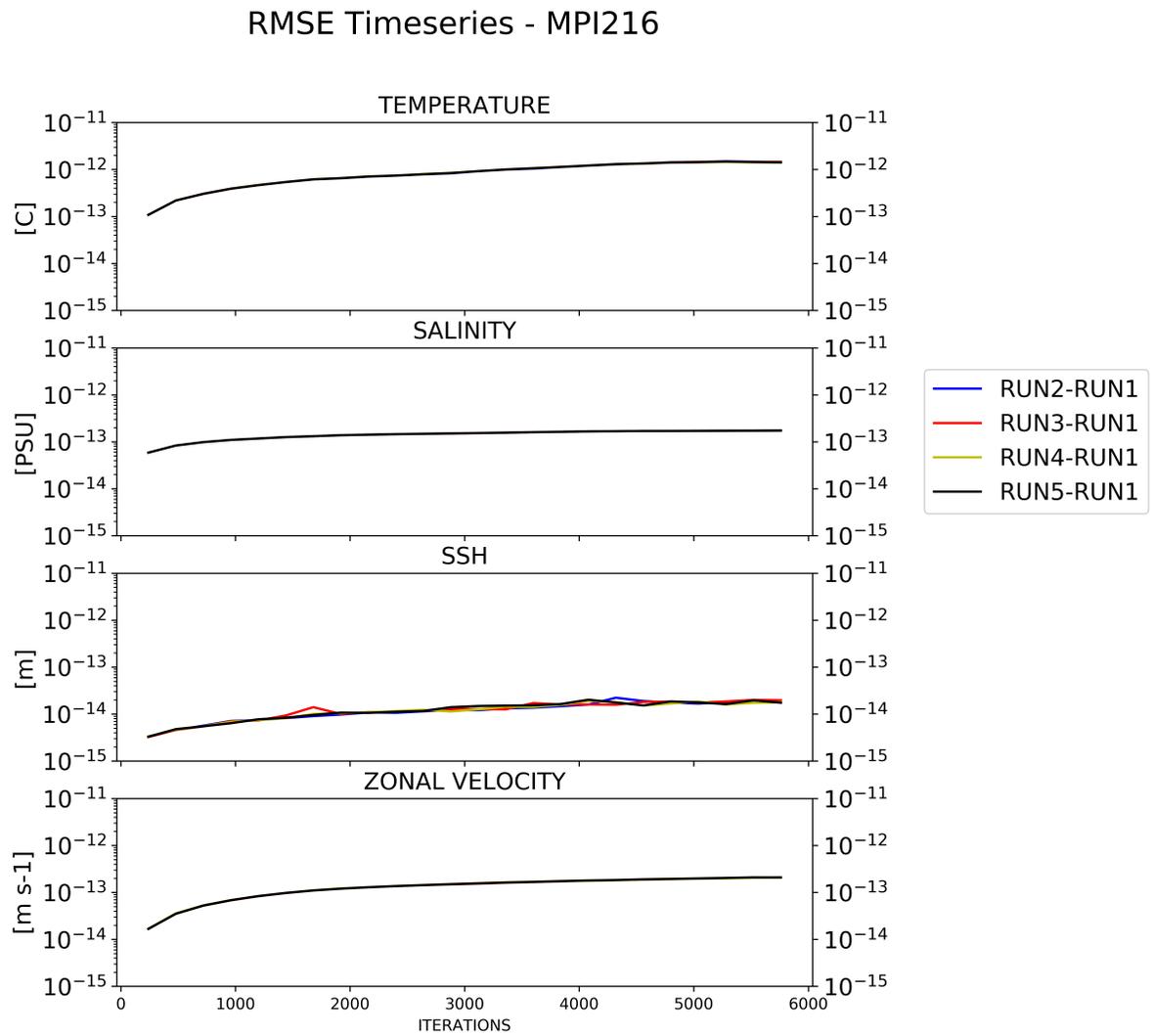


Figure 3.16: Test of Code Reproducibility with 216 MPIProcesses

parallel time ( $T_p(n)$ ), where the latter is supposed to be smaller of the former and decreasing with the amount of processes  $n$  (also named threads) deployed.

The SpeedUp and the Efficiency are defined as

$$\begin{cases} S(n) = \frac{T_s}{T_p(n)} \\ E(n) = 100 \frac{T_s}{n T_p(n)} = 100 S(n) / n \end{cases} \quad (3.9)$$

being the Efficiency generally expressed as percentage. An parallel application that satisfies the requirement  $T_p(n) = (1/n)T_s$  would have  $S(n) = n$  and  $E = 100$  %, corresponding to an ideal application that shows “linear” scalability. In real parallel applications the communication overhead plus components of the program that remain serial push the scalability away from the ideal behaviour.

To test the SHYMPI scalability on the SANIFS configuration we choose a different set of decompositions, with the aim of understanding the limit of the scalability. In practice, the limit is reached when an increase of parallel processes does not produce a reduction of the execution time or, in the worst case, when the execution time increases again.

Figure (3.17) shows the execution time on the SANIFS configuration. We drop the decomposition with few MPI processes (2,4,18) retaining the multiple of 36 (the number of cores per each node on ZEUS cluster (see fig. (3.14))) plus the decompositions with 288 and 360 processes.

The execution time decreases from nearly 8 hours for the serial run to around 3.5 minutes using 360 MPI processes. Deeper insights in the performances are provided by charts in figures (3.18) and (3.19). The SpeedUp deviates significantly from ideal case since the 36 MPI process run.

The efficiency drops below 60 % in runs with more than 216 processes. Though the performances are not excellent, the MPI implementation of the code allows to carry out demanding simulation in a time that can be deemed as acceptable. This lack of parallel efficiency, though, deserves to be investigated thoroughly, seeking for workload unbalance across the processors due to non optimal domain decomposition.

In support of this, figure (3.20) shows the scalability of SHYMPI by components of the program. The most time consuming routines of a single time step (advancement of transports, free surface solver and advection/diffusion equation)

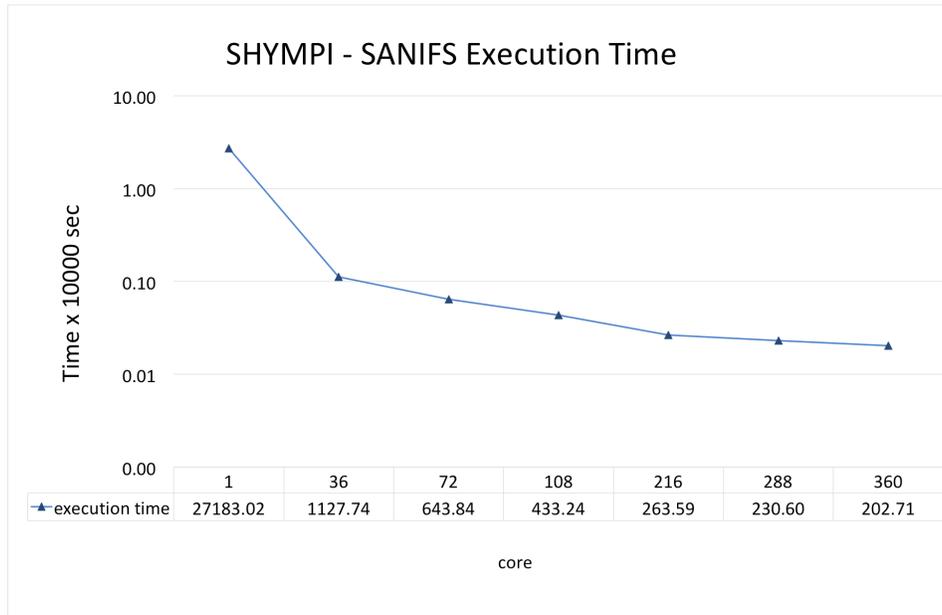


Figure 3.17: SANIFS Execution Time

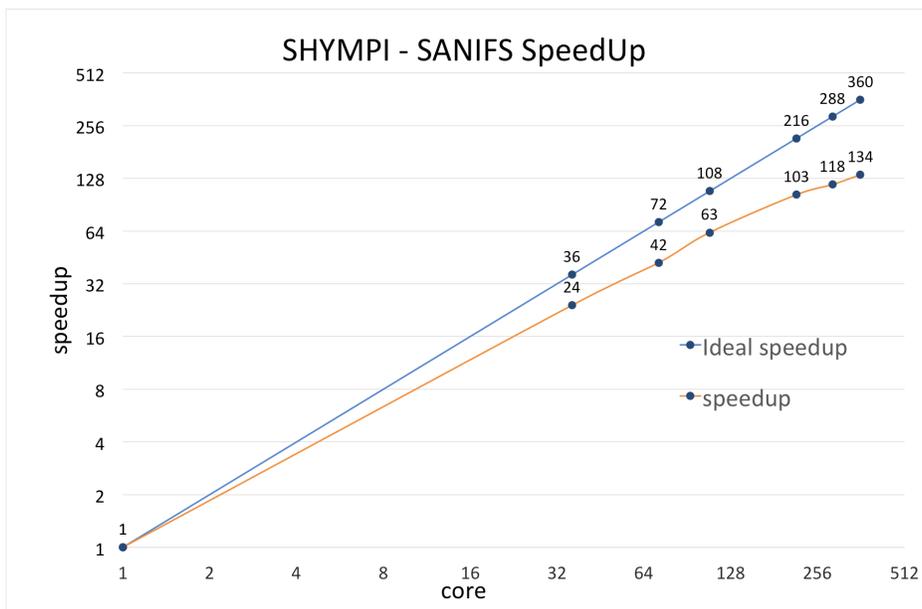


Figure 3.18: SANIFS SpeedUp

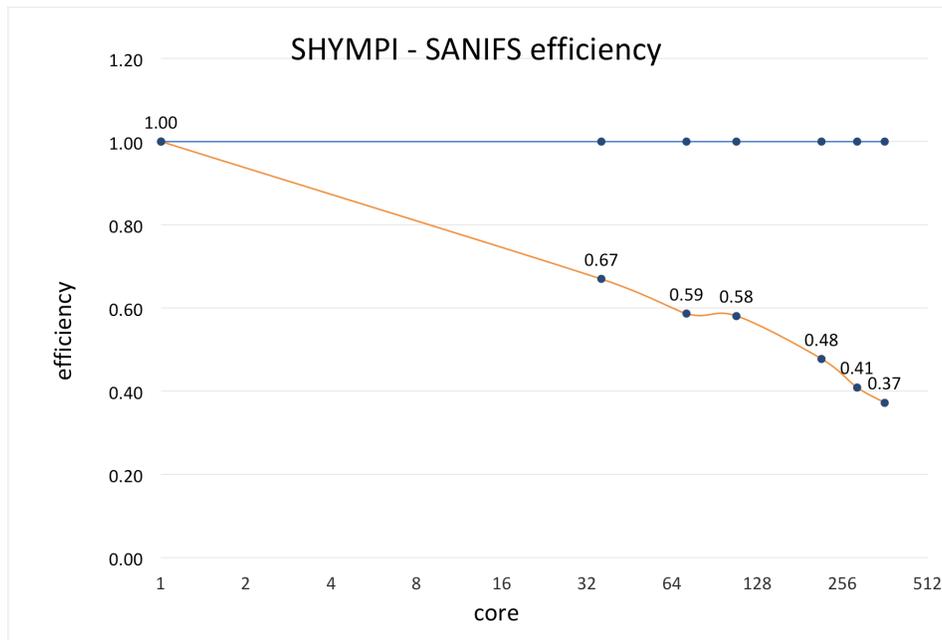


Figure 3.19: SANIFS Efficiency

are treated separately, as well as the I/O time and communication (only for parallel runs). The parts of the time step that are left, denoted as “Other” include the vertical velocity diagnostics, the solution of hydrostatic equation and turbulence closure.

The advancement of transports and the advection/diffusion equation take the most of the time step in the serial run. Communication time in this run is null and I/O is a too small fraction to be visible in the plot.

With the increase of the MPI processes, the ratios among these components change, and both I/O and communication cost become ever considerable burdening the time step. The most expensive components of the runs with highest parallelism are the “Communication+Unbalance” and the free surface solver. The communication takes account of the time spent by the program to send/receive messages, while the unbalance accounts for the time that some MPI processes remain idle. The unbalance is caused by uneven workload across the processes, with the effect that those having a smaller workload can complete the task faster than the other and remain idle as they wait for the other to finish. In an ideal case, the workload is balanced in a way so that all the MPI processes complete their own task taking all the same time. In order to get something comparable to this,

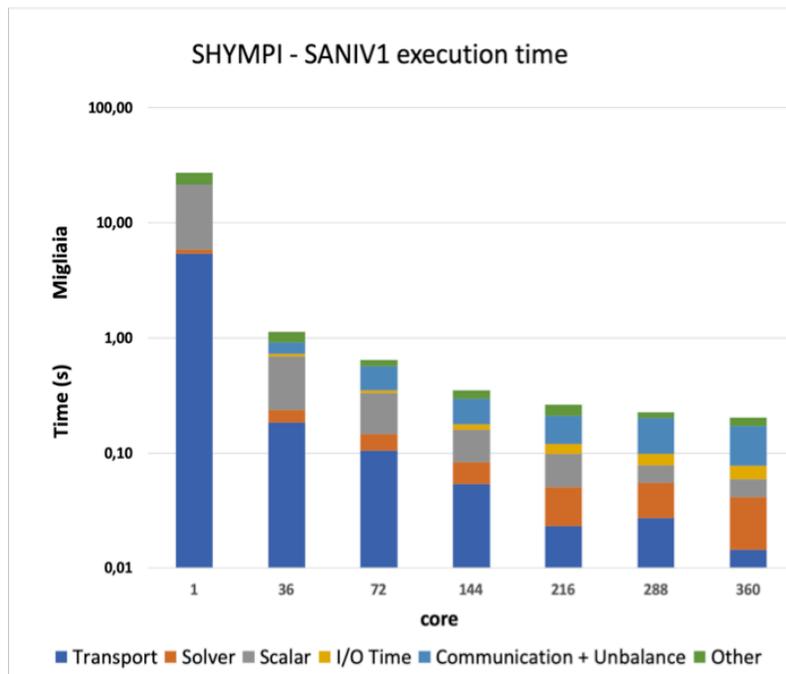


Figure 3.20: SANIFS Scalability by Components

it is essential to act on the domain decomposition for a better workload balance, keeping that the most convenient domain decomposition comes from the trade off between the balance and the limitation of the communications by limiting the inter-process edge cuts.

Almost half of the runtime in the case of 360 MPI processes is spent in the free surface solver, that shows the worst scalability among the model components. The execution time reported for the free surface solver includes the assembly of the system, the communication time of the internal routines of PETSc for each one of the solver iterations and the communication necessary to redistribute the solution onto the model grid.

The effects of a non-optimal model mesh domain decomposition on the solver efficiency hasn't been assessed yet, as well as whether the combination preconditioner/solver iLU+BiCgS is the most suitable in the context of parallel performances.

### 3.6 Extension of Restart Capabilities

The parallelization reduces significantly the execution time of simulations but, yet, some simulations can be highly onerous and is more convenient to split them into a sequence of shorter simulations. In the context of running long term simulations, we carried an enhancement of the restart capability. The restart binary that are written by a simulation can be used to start another from the fields stored inside. The fields include transports, sea level, vertical velocity, active tracers, density at level  $n$ . These fields define the starting point of the single time step of SHYMPI (fig. 2.2).

In the case of turbulence closure model, hence, the mixing coefficient, the dissipation rate and the turbulent kinetic energy at time level  $n$  that are used to initialize the prognostic equations (2.12) have been added in the restart files to avoid discontinuities in the simulation.

The improvement of restart capabilities ensures perfect reproducibility (bit-by-bit comparison) only in the case of serial run. In MPI case the behaviour is non deterministic regardless the use of restart, for the reason explained in section 3.5.1.

### 3.7 Model Diagnostics

The code of SHYMPI has been enhanced with the add of further fields that, optionally, can be cast in separate output from the hydrodynamics and active tracers in function of diagnose the model outcome or to have supplementary information to elaborate in post-processing phase.

The following 2D quantities are calculated if necessary and cast into separate files with the same frequency of the active tracers on the native model grid nodes:

- Mixed Layer Depth (MLD). The calculation is made based on a density criterion. The depth of the mixed layer is evaluated where the density drops for more than  $0.03 \text{ kg} / \text{m}^3$ . The output file has extension `.mld`
- Wind Stress. The zonal and meridional component of the wind stress with seawater density used by the model are saved into binary file having extension `.wnd`

- Surface Water Fluxes. The precipitation and the Evaporation used by the model are saved into file with extension .wfx
- Surface Heat Fluxes. The four terms contributing to the heat fluxes at the air-sea interface, described in section (2.1.4) are stored separately in a file with .qfx extension

# Chapter 4

## The Southern European Seas Modelling System

### Contents

---

<b>4.1</b>	<b>The SES System and The Domain Configuration . . . . .</b>	<b>88</b>
<b>4.2</b>	<b>Experimental Setup . . . . .</b>	<b>91</b>
<b>4.3</b>	<b>Forcings . . . . .</b>	<b>92</b>
4.3.1	Surface and Atlantic Boundaries . . . . .	92
4.3.2	Rivers . . . . .	95
<b>4.4</b>	<b>Model Integration and Performances . . . . .</b>	<b>97</b>
<b>4.5</b>	<b>Validation of Results . . . . .</b>	<b>97</b>
4.5.1	Strait Dynamics . . . . .	98
4.5.2	Sea Surface Temperature . . . . .	102
4.5.3	Circulation Patterns . . . . .	105
4.5.4	Thermohaline Structures . . . . .	108

---

### 4.1 The SES System and The Domain Configuration

This chapter addresses the second main point of this thesis, namely to overcome the lack of complete studies of the SES system, whose dynamics and complexity

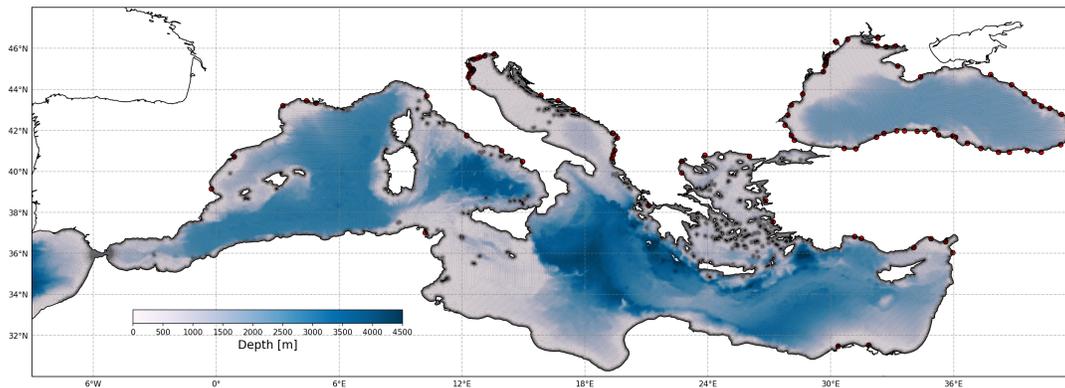


Figure 4.1: Domain and Bathymetry for the SES simulation. The computational grid is superimposed. Red dots represent river sources

has been outlined in chapter 1.8.

The added value of the modelling framework described in this chapter is a more comprehensive simulation representation of SES system including the baroclinic contribution to the dynamics and representing properly all the scales involved, with a resolution sufficient to account the contribution of coast to the general circulation and the resolution of straits.

Simulations like this are extremely computationally demanding, considering also the time scale that is meant to be covered, 1 year in this case. For this purpose the parallel version of SHYFEM, described and validated in chapter 3 is used to exploit to the most the computational resources available at the CMCC HPC facility ZEUS (see fig. 3.14).

The shoreline of the physical domain of the SES system is taken from the Global Self-consistent, Hierarchical, High-resolution Geography ( GSHHS) Database (Wessel and W. Smith 1996), and used to create the computational grid.

The SES grid resolution in the coastal areas reaches 500 meters, with the exception of focal points such as the straits, where the resolution is high enough to represent the strait.

The highest resolution of 50-60 meters is in the Bosphorus (4.2b), around 100 meters in the Dardanelles Strait (4.2c) and varying from 300 to 400 meters in the Gibraltar Straits (4.2a). The resolution decreases linearly in a 2 km wide belt from coastline, after which the resolution it is nearly constant at 4-5 km. The resulting grid has 1610284 triangular cells and 849681 grid points.

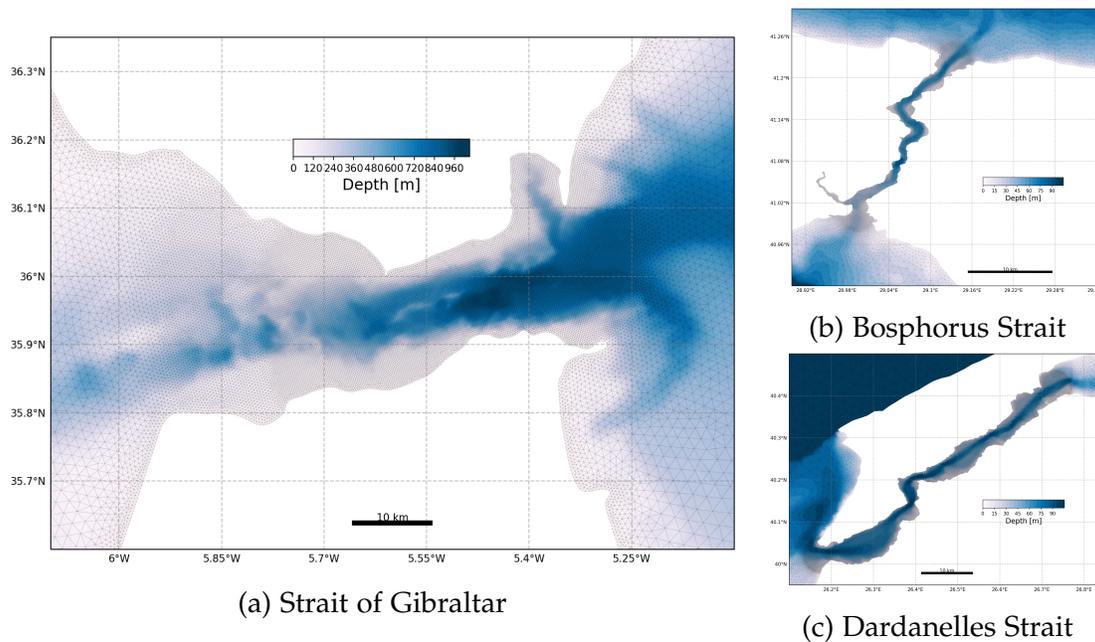


Figure 4.2: Examples of the Computational Grid in the Strait of SES system

The bathymetric information is from the European Marine Observation and Data Network (EMODnet) Digital Terrain Model service, whose digital terrain model has a resolution of around 100 meters. The EMODnet is an harmonised DTM created from selected bathymetric survey data sets, composite DTMs, and Satellite Derive Bathymetry (SDB) data products, while areas with no data coverage are addressed by integrating the GEBCO Digital Bathymetry.

A different dataset is used for the TSS, used in the work of (Gurses 2016), with the permission of the Turkish Navy, Navigation, Hydrography and Oceanography Office. High resolution (20 meters) surveys in the straits have been merged with other data sources for the Black Sea shelf area, including multi-beam surveys obtained by the R/V ALLIANCE and ADCP soundings by R/V BILIM (Özsoy et al. 2001) that have been used to correct biases and to have enhanced representation of bottom features in this area. The coverage of all these datasets The bathymetry for the remaining part of the model domain is based on the GEBCO dataset.

The final bathymetry ranges from 3 meters minimum to avoid wet and dry and to a maximum of 4500 meters. Points of the Mediterranean deeper than 4500 are concentrated along the Hellenic Trench and represent a negligible portion of the domain.

The domain considered does not include the Sea of Azov whose surface becomes frozen during the winter (Zavialov et al. 2020). No ice model is in SHYFEM, and thus this basin cannot not be properly modelled. The inclusion of the Azov Sea is, however, planned for future developments.

The water column is discretized into 104 z-layers, using a stepwise function of the the layer thickness, that has value of 2 m in the top 70 meters (fig. 4.3) to a maximum of 300 close to the bottom. Cells that are in contact with the sea floor are partial if the sea bottom is above the bottom interface of the layer. A vertical resolution at the surface of 2 meters and the minimum of bathymetry set to 3 meters results in at least 2 active layers throughout the domain.

A vertical resolution of 2 meters in the upper ocean may not be sufficient to fully capture the diurnal variability of the sea surface temperature and the intra-seasonal variability of the mixed layer (Bernie et al. 2005), but this resolution is an acceptable trade off between a good vertical resolution and a first layer that is thick enough to cope to potentially high variations in the sea level of the SES basin. Future developments will consider a more advanced choice of vertical coordinate that permits large tidal signal, as can happen at the Atlantic boundary. Of particular interest for the purpose is the  $z^*$  coordinate (A. Adcroft and Campin 2004), that distributes the free surface oscillation over the entire water column and retains the simplicity of a  $z$  coordinate.

## 4.2 Experimental Setup

We simulate the year 2016 in a free run, which is initialized from a January 2016 average of the MFS and BSFS CMEMS analysis products of the active tracers. The Sea of Marmara, and the Turkish Straits are initialized from a climatology of the simulation produced by (Aydoğdu et al. 2018) of the TSS in January. The initial free surface is flat and the fluid is motionless.

Various relevant SHYFEM parameters are reported in table 4.1 and are aimed at achieving the largest time step possible when the model dynamics selected are fully implicit ( $a_M, a_Z = 1$ ). The velocities at lateral open boundaries are relaxed with a nudging time of 1 hour. The advection schemes are centered for the vertical advection in momentum and upwind for both horizontal and vertical

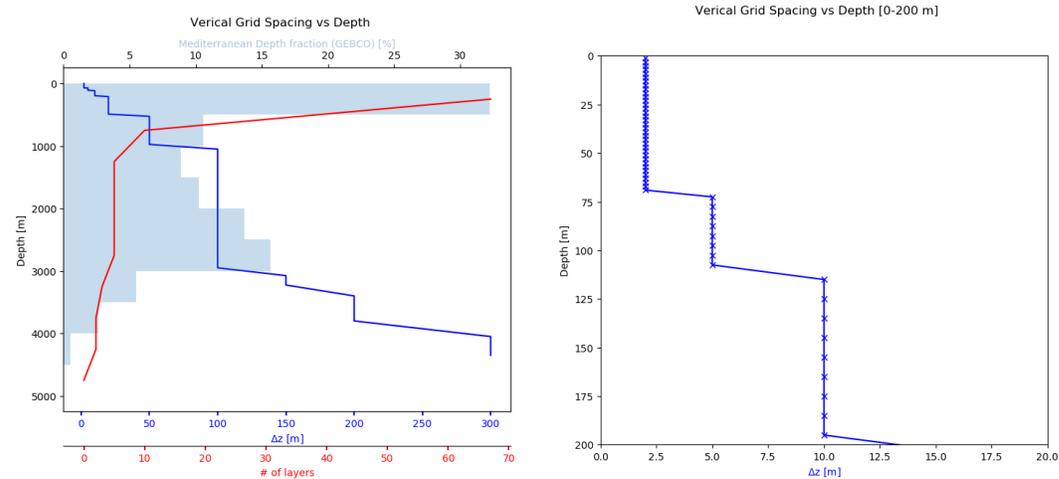


Figure 4.3: [Left Panel] Stepwise function of the layer thickness with depth (blue line). Shaded blue histogram represents the proportion of the depth of the SES system that falls in the 500 m wide bin. The red line represents the number of vertical layers in the bin. [Right Panel] Detail of layer thickness function in the first 200 meters

advection in the advection-diffusion equation (itvd,itvdv). The Smagorinsky horizontal viscosity and diffusivity coefficients are multiplied by a factor named "Horcon" parameter (Blumberg and Mellor 1987), having typical values of 0.2, assigned to both. The ratio between horizontal viscosity and horizontal-diffusivity, or the Prandtl number, is of the order of unity for this simulation. Vertical mixing coefficients are calculated with  $k - \epsilon$  method as in GOTM (Burchard and O. Petersen 1999). Further details of SHYFEM turbulent closure are provided in section 2.1.3. Bottom friction has a quadratic formulation with a bottom roughness coefficient set to  $10^{-2}$ .

## 4.3 Forcings

### 4.3.1 Surface and Atlantic Boundaries

The surface boundary conditions used to force the simulation follows the MFS Bulk formulation formally described in 2.1.4. The required atmospheric fields are from ECMWF High Resolution (HRES) global model Medium-range forecasts, providing wind speed at 10 meters (U10m,V10m), mean sea level pressure (MSL),

name	description	value
$a_M$	weight of n+1 time level for barotropic pressure in momentum	1
$a_Z$	weight of n+1 time level for transport in continuity equation	1
idtout	frequency of output	3 hours
$t_{\text{nudge}}$	relaxation time for velocities a lateral open boundaries	1 hour
ihwadv	scheme for vertical advection in momentum equation	2
itvd	scheme for horizontal advection in advection-diffusion equation	0
itvdv	scheme for vertical advection in advection-diffusion equation	0
ahpar	Horcon parameter for horizontal viscosity	0.2
thpar,shpar	Horcon parameter for horizontal diffusivity	0.2

Table 4.1: Relevant SHYMPI Parameters for SES Simulation

	Source	freq	fields	ref.
Initial condition	CMEMS MFS + uTSS + CMEMS BS	n.a.	T,S	(Clementi et al. 2019) (Aydoğdu et al. 2018) (Ciliberti et al. 2020)
Atmospheric forcing	IFS HRES (0.125) forecast	6h	T2M,D2M,TCC U10m,V10m,MSL rain	ECMWF
Atlantic boundary	CMEMS MFS	1d	T,S, $\eta$ ,U,V	(Clementi et al. 2019)

Table 4.2: Parent models used to initialize and force the simulation

Total cloud cover (TCC), Air temperature (T2M) and Dew point temperature at 2 meters (D2M) and precipitation with spatial and temporal resolution of 0.125 and 6h respectively. The atmospheric fields are bi-linearly interpolated onto the model grid at runtime.

The lateral boundary conditions are formally described in section 2.1.5. The Atlantic boundary is forced with daily fields of sea level, total velocities and active tracers from CMEMS MEDSEA\_ANALYSIS\_FORECAST\_PHY\_006\_013 products (Clementi et al. 2019). Table 4.2 gives a summary of parent model fields used to force the simulation. The parent model fields are interpolated onto the SHYFEM grid in pre-processing phase.

The atmospheric forcing fields are corrected according to the creeping sea-fill methodology described in (Kara, Wallcraft, and Hurlburt 2007) to overcome the problem of land contamination. Land and sea have significantly different properties, such as roughness, and an abrupt transition can consist in an improper representation of the atmospheric forcing near coastline, especially when difference of resolution between ocean and atmospheric model is substantial. The creeping algorithm extends the atmospheric sea points values over the land point (this procedure is also known as Sea-Over-Land). The same procedure is applied to CMEMS temperature and salinity fields used to initialize the simulation.

In this simulation the relaxation time for total velocities is set to 1 hour. The CMEMS products suit the need of this simulation because they are extensively validated and benefit from data assimilation of SST, in situ profiles and sea level.

### 4.3.2 Rivers

River release has been found to significantly affect the ocean circulation and dynamics from a coastal scale ((Simpson et al. 1993); (V. H. Kourafalou 1999); (Garvine 1999); (Schiller and V. Kourafalou 2010) among the others) to a basin scale ((Skirris et al. 2007); (Coles et al. 2013); (Verri et al. 2018)). It significantly the buoyancy budget and the general circulation of marginal seas such as the Adriatic and Black Sea. One third of the Mediterranean discharge is located in the Adriatic which makes it a dilution basin. Rivers reduce the volume of dense water formed in the Southern Adriatic Sea due to increased water stratification. This results in a weakened downwelling branch of the anti-estuarine meridional overturning cell from the Central Mediterranean Sea (Verri et al. 2018).

River inputs also play a key role in the estuarine circulation of the Black Sea.

Here, the freshwater fluxes from river discharge and precipitation are not balanced by evaporation resulting in a net outflow of brackish waters through the Bosphorus Strait towards the Marmara. Estimates of this net transport over the long term range from 200 to 300 km<sup>3</sup>/year (Jarosz et al. 2011).

Thus a dataset of river discharge was created based on monthly mean climatologies. The dataset is the result of merging of various pre-existing datasets and of excluding all those rivers that have less than 50 m<sup>3</sup>/s in terms of mean annual discharge.

We implement in the SES simulation a total of 73 rivers (see fig. 4.1) in the SES simulation corresponding to 94 river sources. Some rivers have more than one source such as the Danube, whose delta has 3 branches and 8 mouths. The Danube is the biggest river source with a discharge of around 6000 m<sup>3</sup>/s with a maximum of 9000 m<sup>3</sup>/s in May (Kara, Wallcraft, Hurlburt, and E.V. Stanev 2008), whose distribution among the branches has been estimated by (Panin 2003)(Panin 2004). The Danube discharge of freshwater provides nearly 60 % of all river input into the Black Sea (Kara, Wallcraft, Hurlburt, and E.V. Stanev 2008).

The largest freshwater discharge in the Mediterranean is provided by the Po river with an annual average of 1584 m<sup>3</sup>/s (Raicich 1996). The Po Delta has been modeled in SES as distributing its discharge among its 9 mouths (fig. 4.4) according the percentages provided by (Provini et al. 1992)

The information on river monthly mean discharges is obtained from the various



Figure 4.4: Location of Po Delta Mouths and grid points where discharges are applied. Blue dots are the actual river mouths locations and red ones are the discharge points in the model. Black dots are the SES model coastline

datasets as listed in table A.1.

Figure 4.5 provides a graphical illustration of the choice of boundary conditions for the rivers. The blue cells indicates the one where the river water is discharged in the domain, and the red cells represent the model's active cells. Normal velocity is imposed to the segment (or segments) corresponding to the river mouth. In the SES simulation the coastal are around 500 m in size, which is comparable with the width of the largest deltas considered in the simulation, so there is always 1 element cell per river discharge. A Dirichlet condition is imposed at each river open boundary with a fixed value of 15 PSU . This is selected due to the lack of observations of salinity in the proximity of river mouths.

Including the river in the model computations is desirable as it then accounts for all of the salt wedge, and can be implemented in a future version of the SES modelling system. A zero-gradient temperature condition is applied.

The left-side grid cells in fig. 4.5 indicate the model's active cells next to the river. To enhance the mixing, the freshwater is introduced only in the top layer or in the first 2 active grid cells of the domain. The SHYFEM parameter "levmax" is used for this, and is set to 1 or 2 depending on the depth of the element closest to the river.

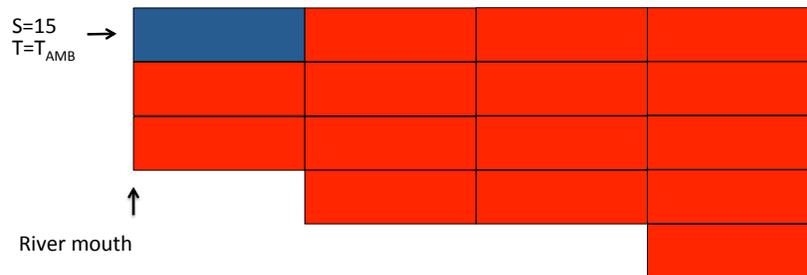


Figure 4.5: Lateral Open Boundary Condition in the case of rivers. Normal Velocities are imposed to the boundary segment. A Dirichlet condition is set for Salinity, equal to 15 PSU for all the rivers. Zero-gradient condition is set for the Temperature

## 4.4 Model Integration and Performances

The simulation of one year required 125000 core hours on the ZEUS Supercomputer (fig. 3.14) with 432 MPI processes. SHYMPI performances is not outstanding with a throughput of around 0.08 SYPD, but it is in part justified by the small timestep of the simulation, which is a variable with a maximum of 40 s. The model goes through 3200 iterations on average to simulate 1 day.

The simulation is separated into 1-day-long chunks, each one producing a restart used by the following chunk. The output is released every 3 simulated hours, giving a total amount of 3.6Tb.

## 4.5 Validation of Results

In this section the output of simulated 2016 is analysed. The quantity of data is huge as well as the amount of informations that can be gathered. The analysis is concentrated then on the features addressed by the simulation considering the strait dynamics, the representation of SST and stratification and the general circulation. SST and stratification are compared against satellite and in-situ observations respectively while circulation patterns are compared with the outcome of

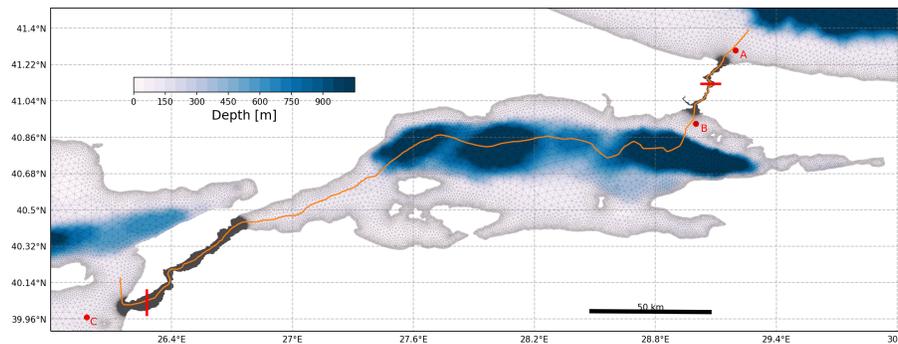


Figure 4.6: Bathymetry of the Turkish Strait System. The grid of SES is superimposed. Timeseries of velocity and sea level are extracted in the points A,B,C. Orange line is the thalweg. Red bars in Dardanelles and Bosphorus indicate where net water transports are computed. Resolution is 100 m and 70 m in Dardanelles and Bosphorus respectively

previous studies.

### 4.5.1 Strait Dynamics

The SES simulation is aimed at accurately representing of the circulation and dynamics of extremely complex systems in semi-enclosed basins connected by narrow straits. Figure 4.6 shows the Turkish Straits System with the locations of A,B and C points where timeseries of water level and velocity are extracted and the transects where net water transports are calculated (red bars) and the path of the thalweg.

The timeseries of the difference of water level between points A and C and the norm of velocity in B are shown in fig. 4.7. The unity vectors at the bottom represent the direction of velocity at B. The  $\eta_A - \eta_C$  difference (green line) is variable during the year and attains the highest values of nearly 0.5 meters during the spring, when the Black Sea river discharge is plentiful. In his simulation (Aydoğdu et al. 2018) estimates this sea level difference as 24 cm in the period 2009-2013. The summer exhibits a decrease in the sea level difference up to the middle of September, with a subsequent increase. The timeseries of velocity in B (blue line) is significantly correlated to the sea level difference, with a Pearson correlation coefficient equal to 0.74, indicating the dependence of the Bosphorus Jet velocity on the Black Sea / Aegean sea level difference. The direction of the

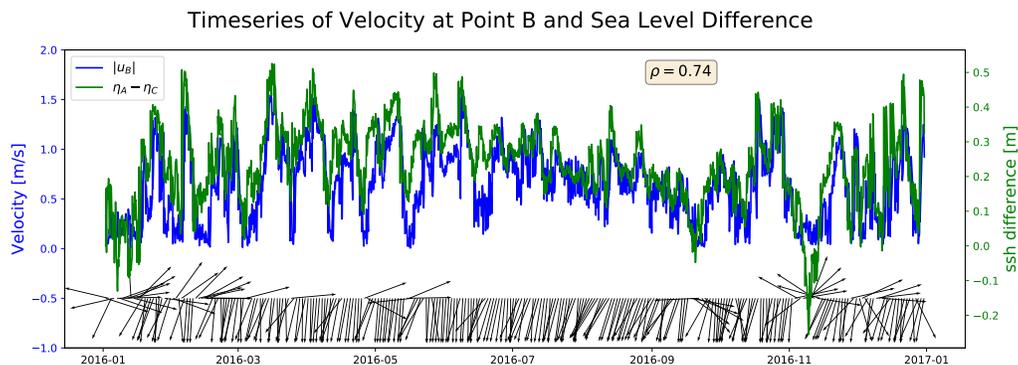


Figure 4.7: Timeseries of difference of sea level between point A and point C in TSS against timeseries of velocity in Point B (see fig. 4.6). Arrows represent the direction of velocity in B

flow is predominantly southwards. Other directions are associated with periods where the velocity is close to 0. The Bosphorus Jet is more intense during spring, with a maximum value of 1.5 m/s.

Figures 4.8, 4.9 show the averaged salinity along the TSS thalweg for July and December 2016 respectively. The main difference is the value of salinity at the surface of Marmara, which is nearly 5 PSU fresher during the summer period because of the higher runoff of the Black Sea during the previous months. The salinity at the surface during the winter period of 2016 is greater when compared long term measurement in Marmara Sea conducted between 1986 and 1992 (Beşiktepe et al. 1994).

In the work of (Aydoğdu et al. 2018), the depth of the interface between the top and bottom layers in the Marmara Sea is located at around 20 meters. In SES this interface is significantly deeper and located at around 40 meters (fig. 4.9). A further observed difference observed is the thickness of the interface between these two layers. It appears sharp and a few meters thick in (Aydoğdu et al. 2018), while in the case of SES the salinity gradient is milder and concentrated between 60 and 100 meters depths. The Marmara salinity is mixed below 100 meters.

The timeseries of the net transports through the main SES straits are reported in fig. 4.10. The cross section for the calculation of TSS transports is presented in fig. 4.6. The net transport through Bosphorus and Dardanelles (defined as positive from the Black Sea to the Aegean) are almost perfectly in phase, showing both high variability during the year. The Dardanelles transport has a higher level of

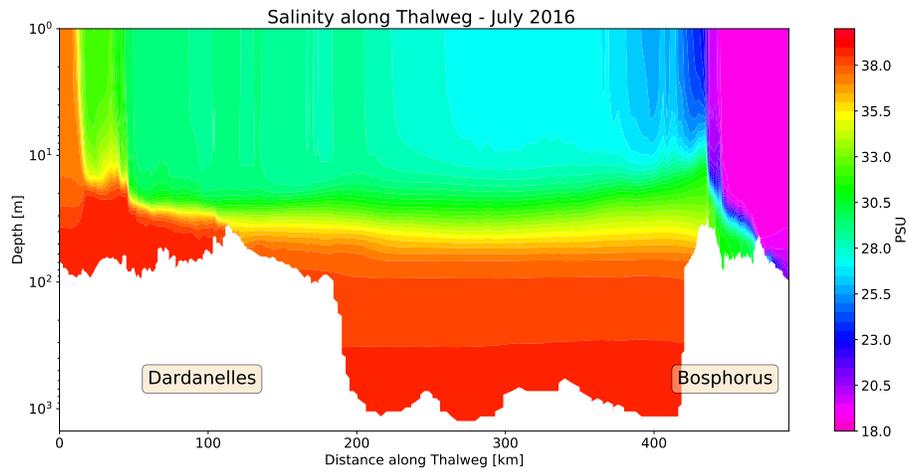


Figure 4.8: Averaged July 2016 Salinity along TSS Thalweg (see fig. 4.6)

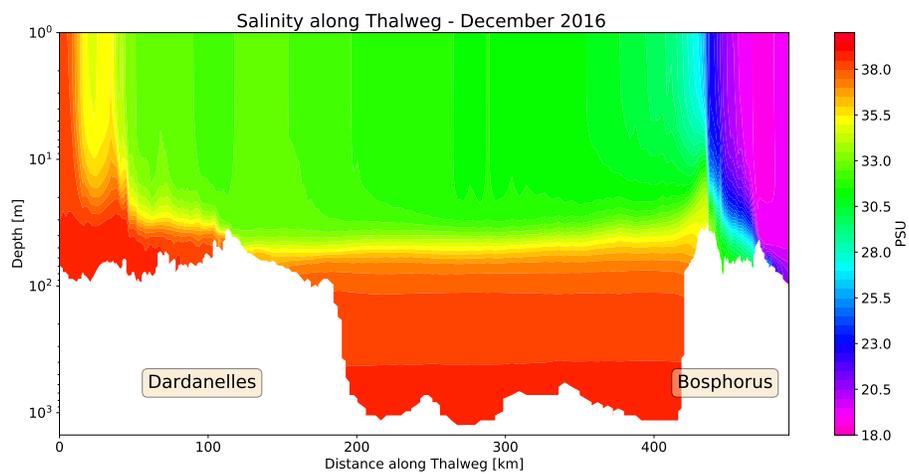


Figure 4.9: Same as in fig 4.8 for December 2016

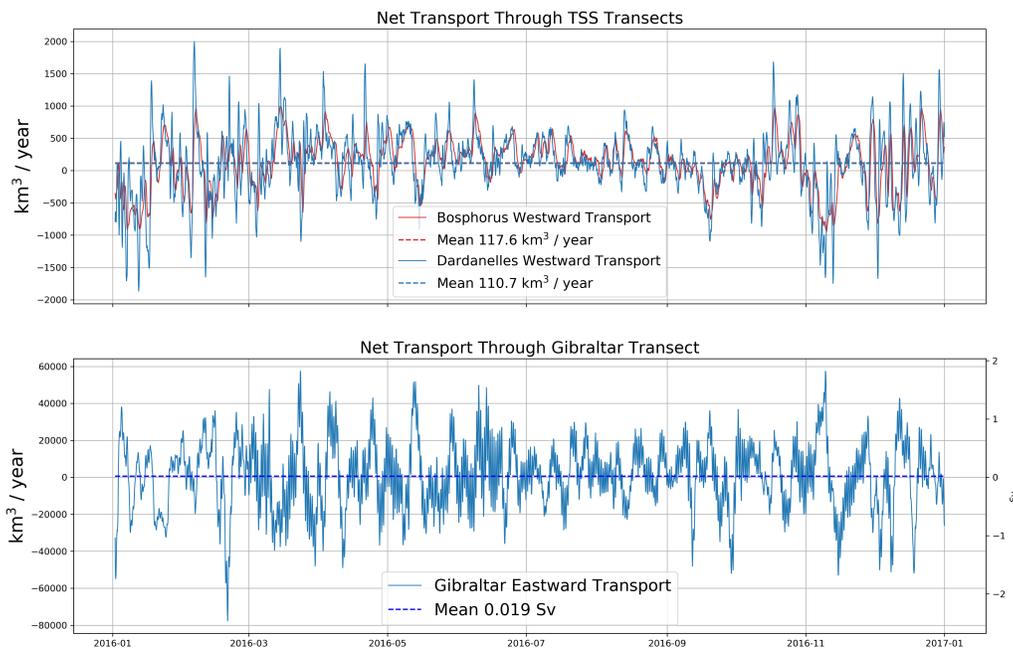


Figure 4.10: Timeseries of Net Transports at TSS Straits (top panel) and Gibraltar Strait (bottom panel). Dashed horizontal lines represent averages.

variability than then Bosphorus but then annual means differ by a few  $\text{km}^3/\text{year}$ . The annual averages differ significantly from previous estimates derived from an analysis based over the long term (Ünlülata et al. 1990b), which are around  $300 \text{ km}^3/\text{year}$ . This difference can be explained by the high variability of the transports over time.

Previous estimates of net transport in TSS also differ. For example, in the simulated period 2008-2013 (Aydoğdu et al. 2018) obtains values of net transports of  $148$  and  $104 \text{ km}^3/\text{year}$  for the Bosphorus and Dardanelles. The model for the Black Sea developed by (Gunduz et al. 2020) considers a reduced box in Marmara and estimates the net Bosphorus transport as  $209 \text{ km}^3/\text{year}$ , nearly twice the observed ADCP measurement value from (Jarosz et al. 2011) in the period 09/2008 - 02/2009, of around  $110 \text{ km}^3/\text{year}$ .

The Bottom panel of fig. (4.10) shows the net transports through the strait of Gibraltar, considered as positive eastward. This is highly variable in this short period. Estimates of the transport at Gibraltar over a longer period have been provided by (Soto-Navarro et al. 2010) from an analysis of both satellite and in-situ observations on the period October 2004 - January 2009, resulting in  $0.038 \text{ Sv}$ , a

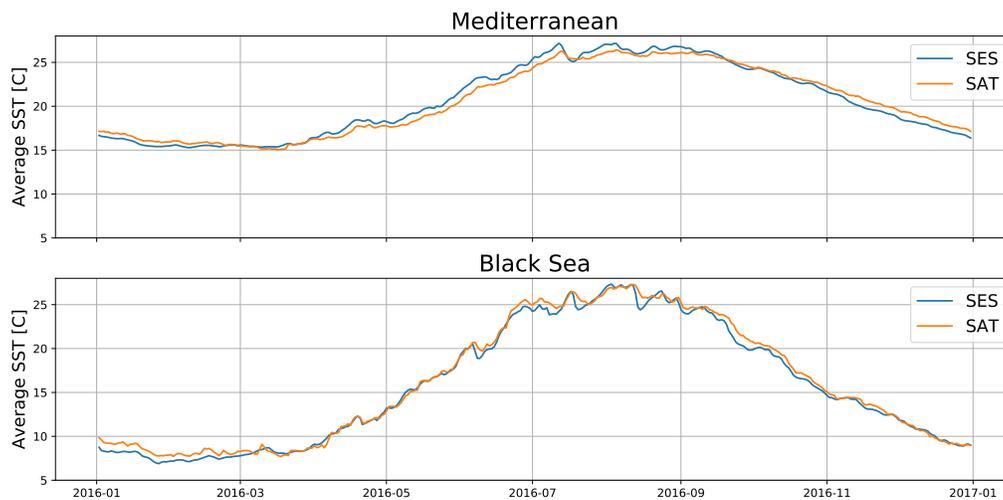


Figure 4.11: Timeseries of Basin Averaged SST for SES in the Mediterranean region (top panel) and Black Sea (bottom) against AVHRR satellite SST (Pisano et al. 2016) (denoted as SAT)

value that is double that of SES for 2016.

## 4.5.2 Sea Surface Temperature

Figure 4.11 shows the timeseries of the basin averaged sea surface temperature (SST) for the Mediterranean Region and the Black Sea, both separately and compared with the observed SST from the satellite.

The satellite temperature dataset is from an analysis over the long term (1981–2017) from Pathfinder v.5.2 (PFV52) Advanced Very High Resolution Radiometer (AVHRR) sea surface temperature (SST) data was been used to build nighttime, high-resolution, reprocessed gap-free SST maps (REP L4) covering the Mediterranean Sea (Pisano et al. 2016).

The comparison is made separately for the Mediterranean and Black sea as the bulk formulation for the air-sea interaction used in this simulation is specifically tailored for the Mediterranean. The average SST for the Mediterranean in the cold period of the year is nearly  $7\text{ C}^\circ$  higher than in the Black Sea, while values for the summer are comparable.

The temporal SST variability is reproduced well in both regions, including some variations in the summer period. The SES sea surface is generally warmer

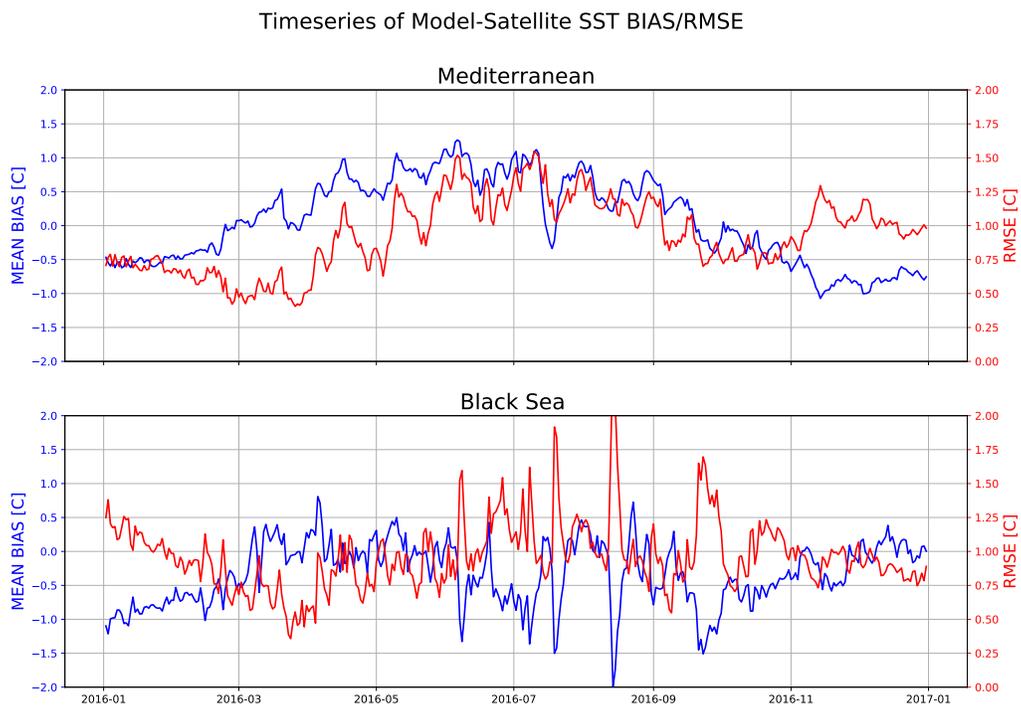


Figure 4.12: Timeseries of mean SES-Satellite SST difference and RMSE for the Mediterranean and Black Sea Region

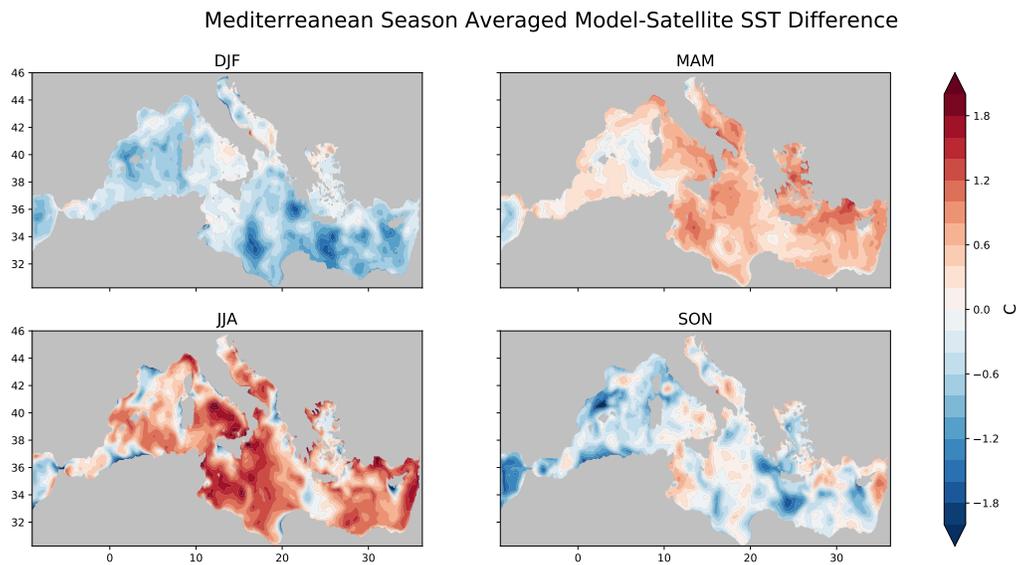


Figure 4.13: Seasonal Averaged SES-Satellite SST BIAS for the Mediterranean

than the satellite in its center in spring and summer and colder in January and in late Autumn. The same cold bias is observed at the beginning of 2016 in the Black Sea.

The SES-Satellite Temperature BIAS and RMSE timeseries are in fig. 4.12. Bias is generally less than  $1\text{ C}^\circ$  in absolute value for the Mediterranean with RMSE giving similar indications of the SST representation. The MFS bulk formulation appears to be effective for the Black Sea, but the Bias timeseries are more noisy, especially in the warm periods of the year.

Figures (4.13),(4.14) show the seasonally averaged BIAS between SES and Satellite SST for the Mediterranean and Black Sea regions respectively.

Bottom left panel of figure (4.13) gives more insights on the bias timeseries visible in fig. (4.12) for the summer period, that shows the highest departure from observations. The warm bias is distributed throughout the basin with particularly hot spots in the central Tyrrhenian and Ionian. The SST in SES is also warmer than observations in the eastern Mediterranean, particularly along the southern coast of Turkey. Cold biases are found in the upwelling regions such as the Sicily Strait (Lermusiaux and Robinson 2001), south of Cyprus (Mauri et al. 2019) and along the Algerian coast.

The bottom right panel of fig. (4.13) shows the bias for the autumn, that has

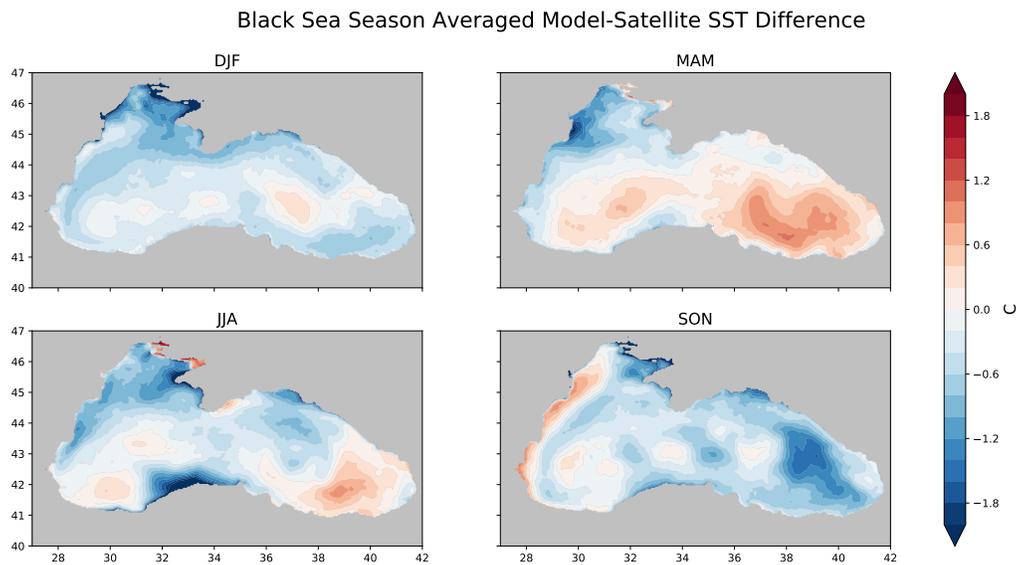


Figure 4.14: Same as in fig. 4.13 for the Black Sea

values around  $-1\text{ C}^\circ$  in the timeseries. Main sources of this bias are located in the Eastern Mediterranean and north of Balearic Islands.

The analogous maps of seasonally averaged bias for the Black Sea (fig. 4.14) explain some features of the bias timeseries of fig. (4.12). The warm bias of around  $0.5\text{ C}^\circ$  visible in the spring in the timeseries is concentrated in the eastern sub-basin. The SST is lower than that of the satellite in correspondence of the Danube Delta, when the river discharge is higher.

As highlighted by the BIAS timeseries, also for the Black Sea the SST for the summer period is also the most difficult to represented for the Black Sea. The cold overshoots in the timeseries of fig. (4.12) are appointed to misrepresented local processes as can be seen in bottom left panel of fig. (4.14). In this case the main deviations from the observed SST are in the shelf areas in the North West and along the coast of Turkey.

### 4.5.3 Circulation Patterns

The general circulation of the SES system is shown in fig. 4.15, showing the average of 2016 at 5 meters depth. The intensity of currents is represented by both thickness of the streamlines and the filled contour plot.

Several circulation patterns identified by (Adani et al. 2011)(Pinardi, Zavatarelli,

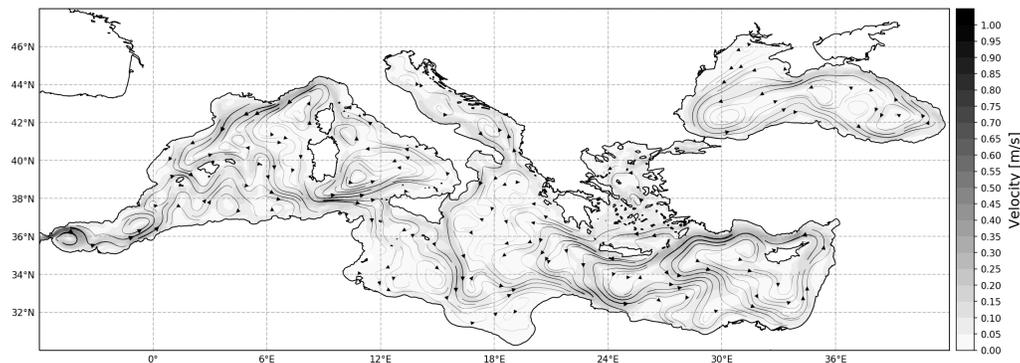


Figure 4.15: Average currents for 2016 at 5 meters depth. The streamlines thickness is proportional to the current intensity, displayed in the contour plot.

et al. 2015) are visible. The Western Alboran gyre and the Almera-Oran cyclonic eddy appear definite, while the Eastern does not. The anticyclonic eddies of the Algerian current are also visible.

Other intense currents of the WMED such as the Liguro-Provençal-Catalan current, edging the Gyre in the Gulf of Lyon, and the Western-Mid-Mediterranean current are clearly visible along with other subtle features such as the Northern Tyrrhenian Gyre.

The path of the Sicily Strait Tunisian current is directed southward, and continues along the Libyan coast. The path of the Atlantic-Ionian stream diverts southwards before flowing to the EMED, after flanking an anticyclonic structure that lies in the Gulf of Sirte and appears smaller than in the analysis of (Adani et al. 2011).

The gyres of the Adriatic sea are present, with the Western Adriatic Current flowing into the Northern Ionian.

The EMED's most intense currents is the Cretan Passage Southern current which continues as the Southern Levantine current and the Asian Minor current along the rim of the Levantine basin. Various eddies in this area are similar to those in the 1987-2007 reanalysis.

The rim current of the Black Sea covers the whole basin, showing a clear Western Gyre and a more complex system of gyres in the East of the basin, similar to (J. Staneva et al. 2001) (see fig. 1.7)

The SES circulation at 300 m is displayed in fig. 4.16 with the contour limit set to 20 cm/s. The pattern most closely resembles that with weaker surface currents

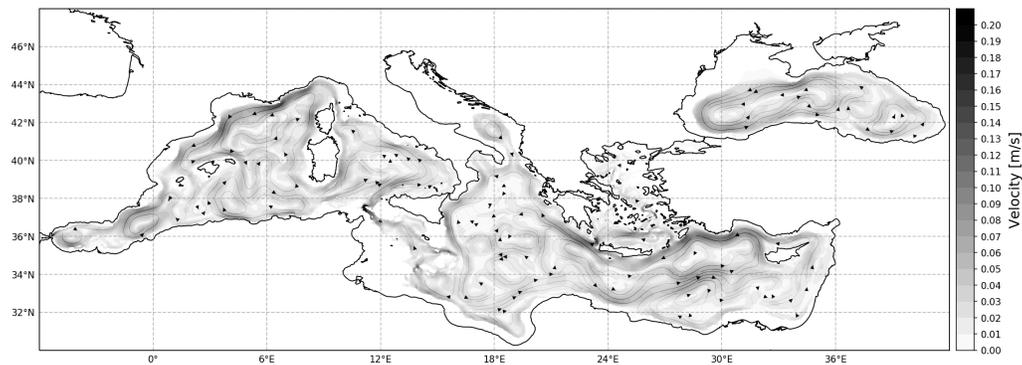


Figure 4.16: As in fig. 4.15 for 300 meters depth

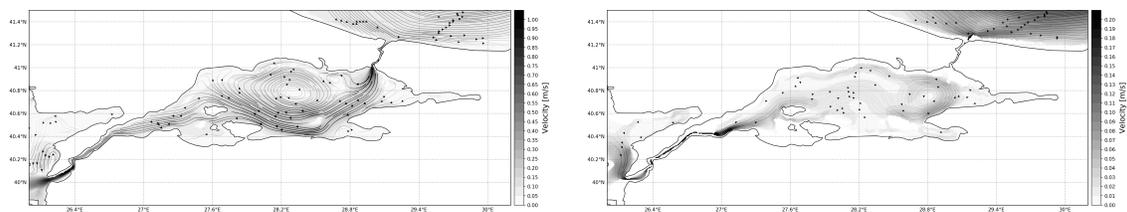


Figure 4.17: Average currents for 2016 in the Turkish Strait System. Streamlines thickness is proportional to current intensity. [Left Panel] Surface Currents [Right Panel] Currents at 50 meters

near the surface. The opposite direction of the Levantine Intermediate water with respect to AIS is visible in the Strait of Sicily.

The circulation in the Sea of Marmara is given in fig. 4.17 using a different scale to show the magnitude of currents. The surface circulation is dominated by the Jet of Bosphorus that enforces the basin-wide anticyclone. Currents are particularly intense at the exit of the Bosphorus and Dardanelles, with average values close to 1 m/s.

This pattern partially matches those in the previous studies, with an S-shaped current in the surface Marmara circulation (Beşiktepe et al. 1994). The first part borders the south flank of the central anticyclone and diverts northward and along the North coast of Marmara before diverting westward into the Dardanelles. The last part of this quasi-permanent feature is not observed in this simulation as it is in the works of (Sannino et al. 2017) and (Aydoğdu et al. 2018). The central part of the Marmara Sea has a rather coarse resolution in SES, around 3.5 km, possibly explaining this shortcoming.

The Marmara circulation pattern at 50 meters of depth is predominantly eastward with highest velocities in the Dardanelles strait. The current flows into a cyclone located in the east of the basin. A small cyclonic structure is present in the North of Marmara.

#### 4.5.4 Thermohaline Structures

To examine the SES representation of the baroclinic structure of the water column we first consider the spatial and temporal distribution of the Mixed Layer Depth (MLD), which is directly released directly by SHYMPI as a diagnostic field. The MLD is calculated as the depth at which the density is  $0.03 \text{ kg m}^{-3}$  higher than the density at surface. A comprehensive list of methods of calculation of MLD is given in (Boyer Montégut et al. 2004).

The maps of monthly averaged MLD in SES are given in fig. 4.18. The result is compared with MLD climatology obtained by (D'Ortenzio et al. 2005) with a temperature criterion of  $0.2 \text{ C}^\circ$  applied to a combined dataset from the MEDAR/MEDATLAS project and the World Ocean Database that considers profiles from 1940 to 2004.

Similarly to (D'Ortenzio et al. 2005) inhomogeneities in the MLD are observed in the MLD during the first months of the year in the Mediterranean, where shallow values of around 30-50 meters in the WMED are found, while MLD is deeper in the EMED. The MLD in the EMED is shallower when compared to the longer climatology result, but some spots with values higher than 200-300 meters are found in the South Adriatic and in Levanine Basin.

The MLD appears rather homogeneous during the central months of the year, with particularly shallow values that only exceed 10 meters in some shallow areas in September. These values are systematically shallower than the climatology.

The correspondent 2016 MLD climatology from parent models is in fig. 4.19 calculated with a density criterion after (Boyer Montégut et al. 2004). The maps for cold months are qualitatively similar to the SES case while differences during the warm part of the year are evident with CMEMS estimate nearly 20 meters deeper. The misfit with both the observations on the long period and the parent model leads to conclude that SES underestimates the MLD in the case of high stratification due to, apparently, too modest mixing. The last consideration is

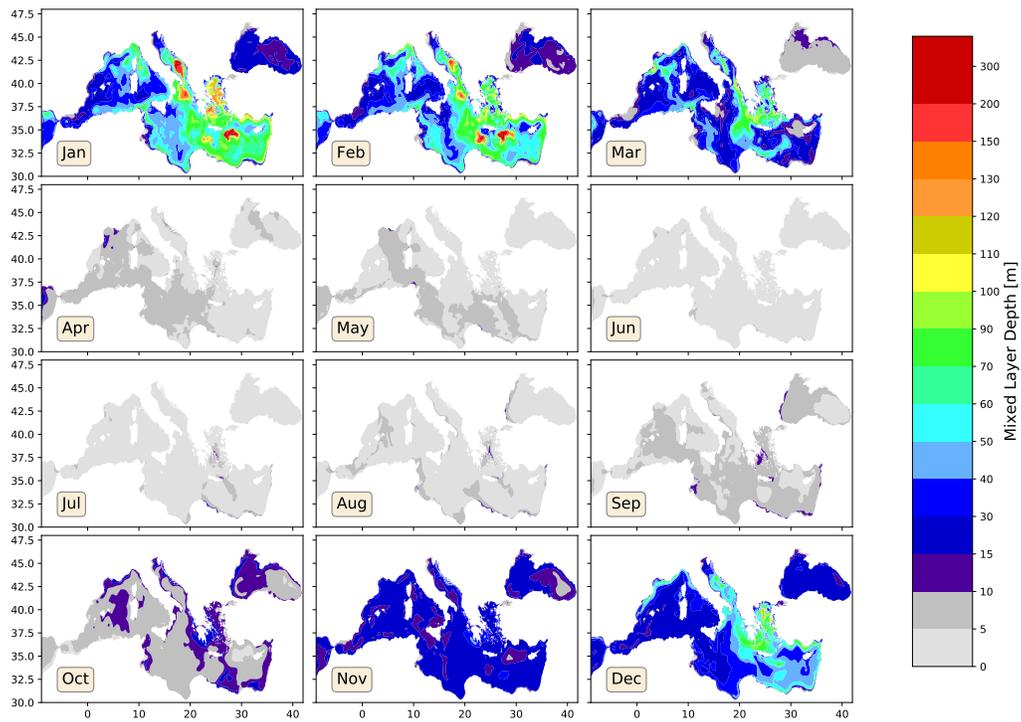


Figure 4.18: Maps of Mixed Layer Depth calculated with Density Criterion for 2016

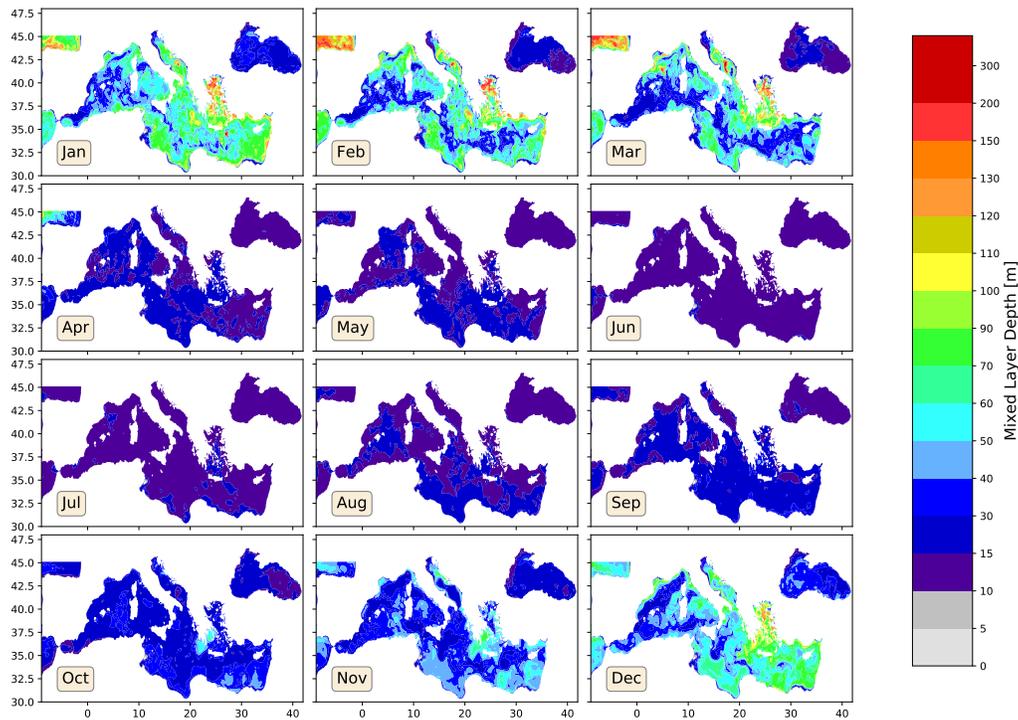


Figure 4.19: As in fig. 4.18 for the CMEMS analysis products

made under the assumption that the parent model with the benefit of data assimilation of both SST and Temperature/Salinity profiles, gives a more realistic representation of stratification.

We further compare the temperature and salinity stratification with the observed in-situ profiles available at CMEMS portal <sup>1</sup>. Figure 4.20 shows the locations of the profiles used for the comparison colored by season. The SES system covers a relatively large area and a detailed analysis by sub-area is preferable, so the SES domain is divided into 18 sub-areas in the same fashion as in the Quality Information Document (QUID) for the validation of the MFS Analysis and forecast. Two further sub-areas have been added in the Black Sea, divided into West and East regions. The analysis is not conducted in the North Adriatic as it generally has no data due to its shallowness. The general circulation also makes drifters unlikely to be found there.

<sup>1</sup>INSITU\_MED\_NRT\_OBSERVATIONS\_013\_035, INSITU\_BS\_NRT\_OBSERVATIONS\_013\_034

The vertical structure is validated by means of the CMCC internally developed Product Quality tool (PQTool) for the Black Sea Forecast System ( BSFS ) that extracts RMSE and BIAS metrics along a selected vertical subdivision of the water column. The PQTool has been adapted to analyse datasets based on the SHYFEM native grid.

The sector analysis of the RMSE of SES vs the observed temperature is given in fig. (4.21) with a line for each season. Sub-areas for the summer and autumn season show the most noticeable deviations from the observations.

In all the sectors the bulk of RMSE is concentrated in the first 100 meters with maximum between 20-50 both in summer and autumn. The Ionian sectors, differently from the others, show a significantly larger RMSE in autumn compared to summer, with maximum at 50 meters.

In most of the sub-regions the RMSE is below  $2\text{ C}^\circ$ , with the sub-regions of the EMED and Black Sea showing values of nearly  $3\text{ C}^\circ$ .

The sector averaged BIAS is shown in figure (4.22). Larger biases are always in correspondence of the thermocline in JJA and SON season at around 20 meters. SES is generally colder than in the observations, with cold biases that can reach  $2\text{ C}^\circ$ , such as in the North Tyrrhenian Sea (tyrr\_area6).

The inversion of the BIAS is common among all the Mediterranean sub-areas, which have slightly warm biases near the surface and more pronounced cold biases in the thermocline. This feature is particularly noticeable in summer and indicates that the surface boundary condition for the temperature is correctly modeled.

The Black Sea SES temperature is also cold biased the surface vicinity. During the colder seasons, in which columns are largely mixed, the BIAS is narrow with values not exceeding  $1\text{ C}^\circ$ .

Some sub-areas show substantial results, as it is for Adriatic area 10, showing relatively small values of the two metrics for temperature ( $1\text{ C}^\circ$ ) and salinity ( $0.2\text{ PSU}$ ). It can be argued that this is due to the realistic representation of the river runoff in the Adriatic, being a non negligible factor in the dynamics of the region (Verri et al. 2018) (Guarnieri et al. 2013).

The RMSE and BIAS analyses were conducted for the salinity in all the sub-

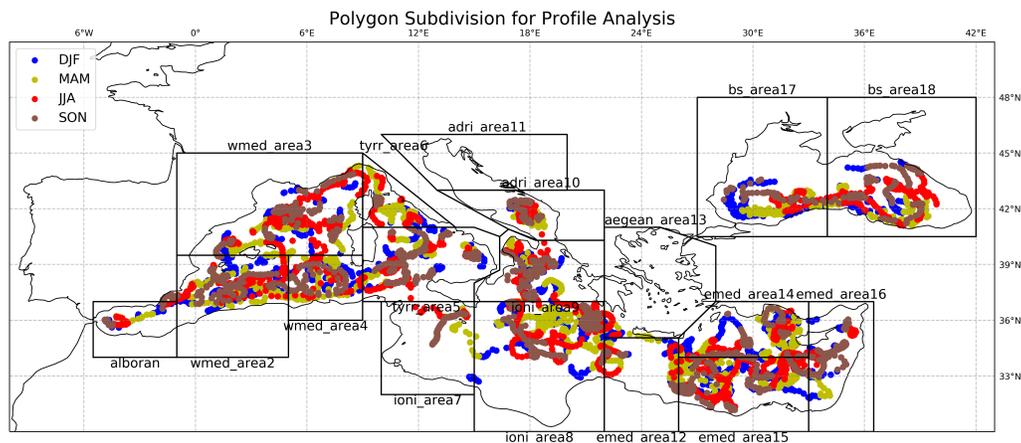


Figure 4.20: Profiles used to validate the representation of stratification. The analysis is subdivided into areas represented by the black polygons and into seasons represented by the colors

regions. Large values of RMSE are confined to the first 100 meters with higher values below the surface, as in the Ionian sub-areas. The Alboran shows a relatively high RMSE for all the season except for the winter between 100 and 200 meters, where the more saline Mediterranean waters wedge under the inflowing Atlantic waters. The reason of this deviation is not clear and deserves more investigation as the representation of the straits and their influence on the dynamics should be a point of strength in the SES simulation. Investigations for this shortcoming can also include a non optimal implementation of the Atlantic open boundary. The BIAS for the alboran sub-region, however, shows values of -0.2 PSU which is smaller than the BIAS obtained for other close areas of the WMED, such as wmed\_area2 and wmed\_area4, and comparable to the other subregions. This leads to argue that the variability caused by the Atlantic boundary is not perfectly reproduced by the model and the differences are exacerbated by the RMSE.

The remaining sub-areas of SES do not show significant BIAS against the observations, with regular functions, indicating that the RMSE exacerbates differences related to isolated profiles.

The Black Sea, however, performs worse than the other sub-areas in the representation of Salinity, where SES is more saline in the top 80 meters in summer and autumn with a BIAS that is around 0.8 PSU. Below 100 meters of depth the SES is fresher in all the seasons of approximately 0.3-0.4 PSU.

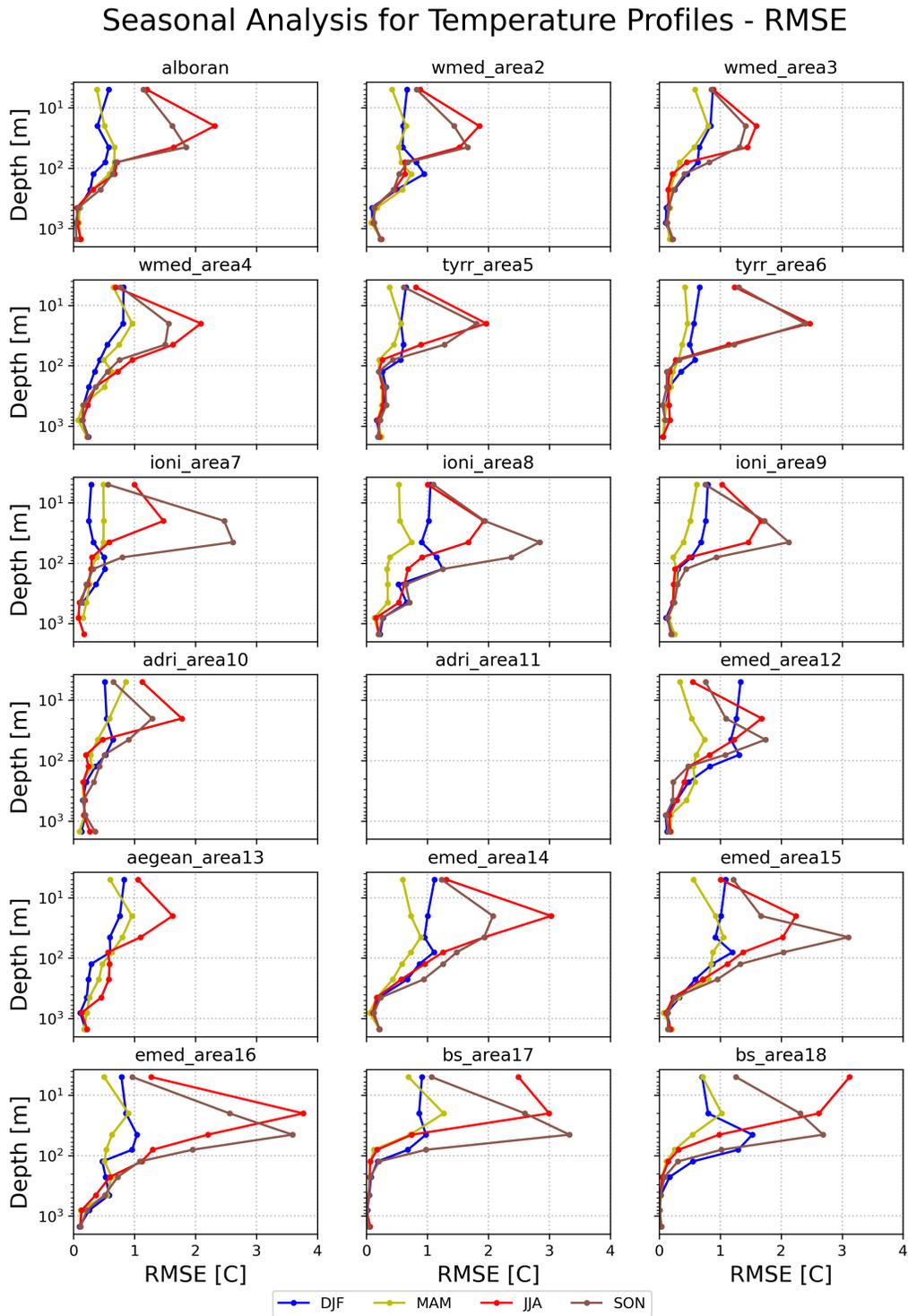


Figure 4.21: Seasonal Analysis of the Temperature stratification using RMSE metric in the SES areas

Seasonal Analysis for Temperature Profiles - BIAS

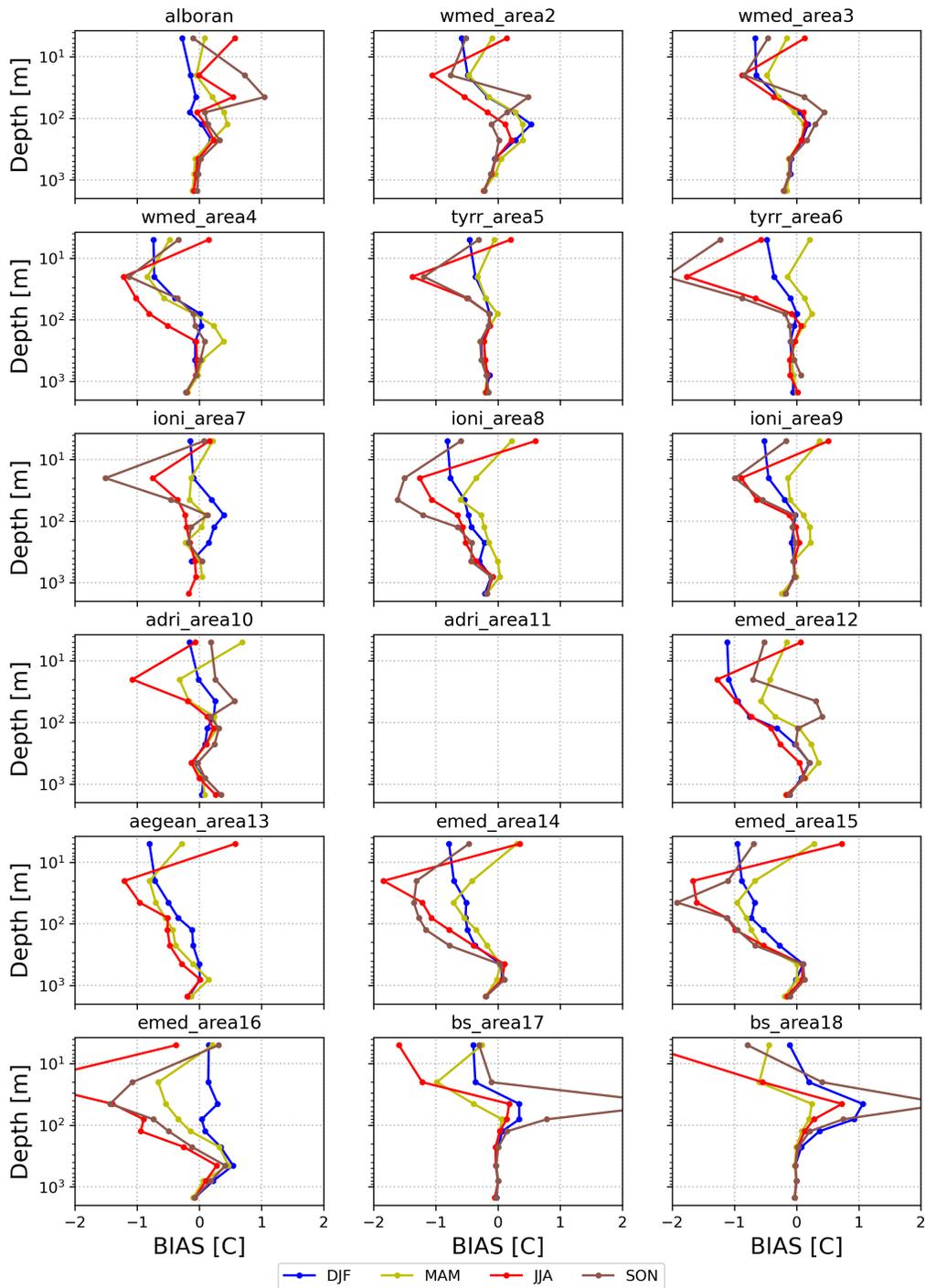


Figure 4.22: As in fig. 4.21 for the BIAS

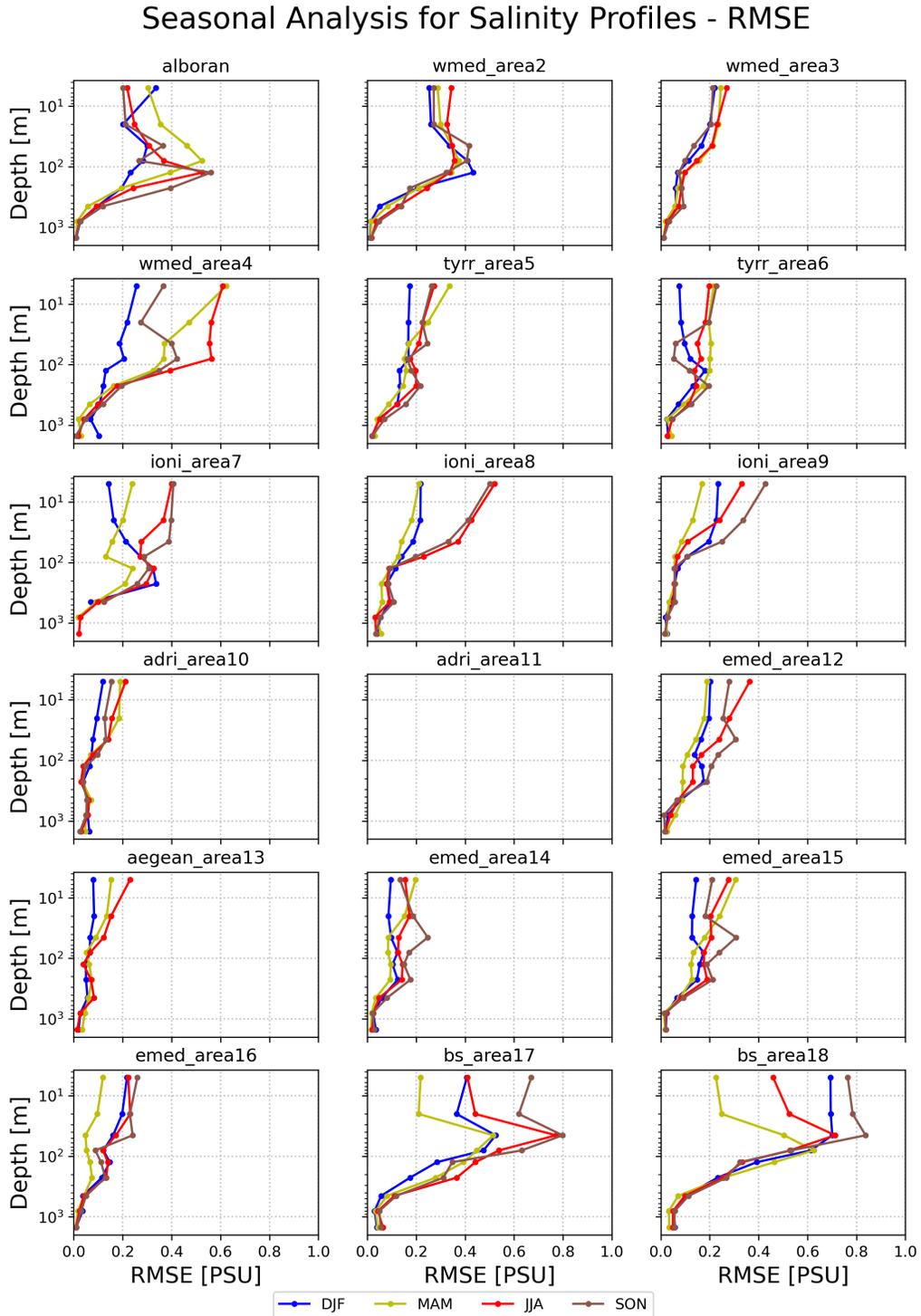


Figure 4.23: Seasonal Analysis of the Salinity stratification using RMSE metric in the SES areas

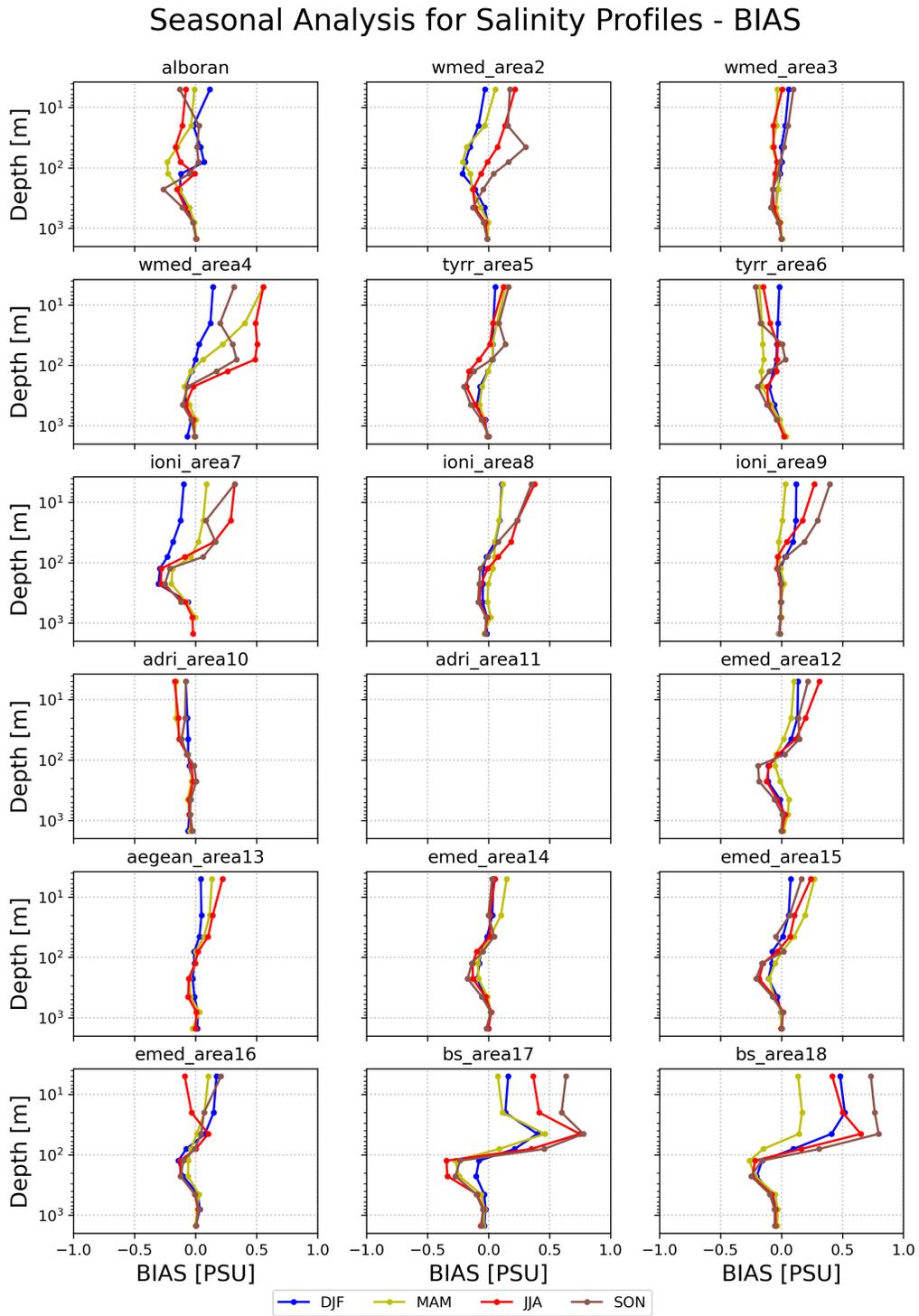


Figure 4.24: As in fig. 4.21 for the BIAS

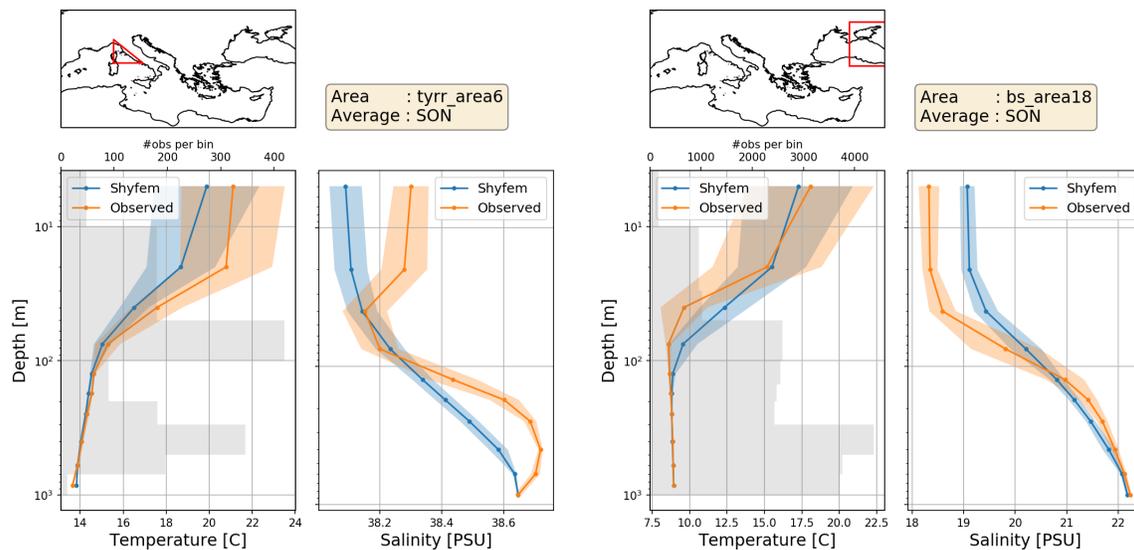


Figure 4.25: [Left Panel] Autumn Area-averaged model and observed profiles for `tyrr_area6`. Thick lines represent the average, shaded area the  $[-\sigma, \sigma]$  range. Shaded gray histogram represents the number of observations per depth bin [Right Panel] The same as left panel for `bs_area18`

In order to investigate the source of this departure from the observations, we focus on the sectors where the BIAS is larger. The area-averaged profiles for `tyrr_area6` and `bs_area18` are in fig. (4.25). In both cases the BIAS of the temperature is attributed to a milder thermocline with respect to the observations. The BIAS for the salinity is around 0.2 PSU in the Tyrrhenian sector but is larger in the Black Sea case with nearly 0.8 PSU more saline water in the first 50 meters. This difference can be hardly appointed to lack of freshwater input, since almost the Black Sea rivers are included in the simulation.

To understand to what extent this misrepresentation of the thermohaline structure depends on the model setup, namely to the choice of the advection schemes, the lock exchange experiment has been carried with SHYFEM to test its capability of reproducing the theoretical results obtained by (Benjamin 1968). The details of the experiment, provided in Appendix B, show that using the low order advection schemes available, the solution for gravity adjustment current velocity obtained with SHYFEM does not match with the theory in on the scales represented in the SES simulation.

It can be argued, then, that the shortcomings in the SES representation of the

vertical structures can be due to some features of the numerics of SHYFEM that deserve more examinations.

# Chapter 5

## Conclusions and Future Work

The complexity and heterogeneous nature of the water bodies composing the Southern European Seas has been identified in various previous studies.

We report the most significant studies that have used modelling to describe the physics of such water bodies, either considered singularly or in a broader context, in which their interactions with other components of SES are considered. Most of these modelling efforts focus on a single or a few physical processes, or in some cases, the representation of the physics is detailed but only considers a restricted component of SES. To the author's knowledge, none of these past works address the following main points: 1) Modelling the SES system as a whole; 2) a complete representation of its physics; or 3) examining the effects of the physical processes over the medium to long-term. The main aim of this study is to simulate the SES system as a whole, considering the most important processes that drive its dynamics over a time-scale that is as long as possible. Thus, it is essential to apply code that is optimized and can exploit current computational capabilities. Therefore realistic and comprehensive simulations can only be conducted using parallel codes.

The unstructured grid model SHYFEM is found to be suitable for the study of SES dynamics, but its code is not parallel. Thus, a detailed intervention in the SHYFEM code to extend its capabilities is required. The hydrodynamical core of SHYFEM is parallelized with a distributed memory strategy, allowing for both calculation and memory scalability. The implementation of the parallel version includes external libraries for domain partitioning and the solution of

the free surface equation. The validation of the parallel code was conducted using a realistic configuration as a benchmark. The parallel model does not exactly reproduce the output of the serial code but can reproduce the physics of the problem without significant differences in terms of the serial program, and are close to machine precision. The source of these differences has been to the different order of operations in each of the domain decompositions since, due to roundoff errors, the associative laws of algebra do not necessarily hold for floating-point numbers. Forcing the code to exactly reproduce the order of the operation in the serial code has been found to lead to a dramatic loss of efficiency, and is therefore not considered in this work.

Our assessment reveals that scalability in the parallel code is poor, with parallel efficiency dropping below 40% when massive parallelism is attained. The analysis of the parallel performance reveals that with the high level of MPI processes used, the burden of communication and the cost to solve the free surface equation take up a huge proportion of the single model time step. The non-optimal distribution of the workload among the MPI processes is the main cause of this, which affects both the 3D calculation and the 2D solver for the free surface. The workload balance needs to be improved, with a more suitable solution for domain partitioning. The parallel code, however, enables one of the main tasks of this work to be accomplished, namely to obtain the results of the simulation in a time that is reasonable and significantly faster than the serial case. The benchmark has demonstrated that the execution time is reduced from nearly 8 hours for the serial program to less than 4 minutes with 360 MPI cores.

The parallel code has also been used to model the Southern European Seas system, demonstrating that this simulation is now feasible. The enhanced throughput of the model means that the production of massive output raises the issue of post-processing, and that specific tools aimed at extracting information from the created dataset are required. We have only partly addressed this problem, mainly by gathering information qualitatively. A river discharge dataset was created to model the river influence on the SES circulation, mainly in the Black Sea, and excluding small rivers that are not deemed as important, particularly for the Mediterranean.

The boundary conditions set for the rivers imply the same value of salinity imposed at river mouths, and thus is a relatively rough approximation. Modelling

the river-sea continuum by extending the model grid into rivers, which is the most desirable option, or by combining the simulation with a model that simulates river-sea exchanges, would be of benefit. These features can be considered in future developments. The analysis of one simulated year's output focuses on the strait dynamics, the representation of the surface temperature, the circulation patterns and the stratification.

The sea surface temperature comparison shows good agreement with the satellite, both in the Mediterranean and in the Black Sea, indicating that the bulk formulas for air-sea interaction devised for Mediterranean also work for the Black Sea, at least in the medium terms.

The values for the transport in the strait are in line with those in the literature, but no observations are available for the period of the simulation. Some discrepancies with previous studies are found in the stratification of Marmara, but these do not show the role of the Black Sea runoff in terms of the surface salinity.

The general circulation of SES is well represented by the model, where most of the relevant patterns depicted in previous works are visible.

A thorough analysis of the stratification has been conducted for all of the SES when compared with the ARGO profiles. This reveals that the largest errors are associated with the temperature in the thermocline. This error is greater in summer and fall in particular, when higher stratification is present. The sector connotation of the analysis leads us to identify the Black Sea as the most problematic and showing the greatest BIAS. The temperature BIAS for the other SES sub-regions remains confined to the  $[-1,1]$  C° range, and thus we can conclude that the simulation can sufficiently reproduce the temperature stratification. The salinity behaves similarly, with BIAS in the range  $[-0.5,0.5]$  PSU throughout the SES with the exception of the Black Sea, where significant biases are also present below 100 meters.

The departures from observed profiles have been further analyzed with comparison between sector averaged model/observation profiles in sectors where the temperature BIAS is more prominent. In these cases the thermocline in SES is significantly milder than the observations. In order to understand to what extent this difference can be attributed to experimental setup or numerical shortcomings, the lock exchange experiment has been carried following the recent related literature. The experiment has highlighted relevant discrepancies with respect to the

literature, leading to the conclusion that some aspects of SHYFEM numerics need to be analyzed better.

These considerations suggest the following future developments:

- The scalability of SHYMPI has room for improvement, and an optimal workload balance across the MPI processes should be sought. Also the performance of the FSE solver can be improved with a tailored combination of preconditioner / iterative solver.
- The numerics of SHYFEM is meant to be thoroughly evaluated, especially for what concerns the tracer advection, with the aim of reproduce correctly the solution of idealized test case such as the lock exchange experiment. The possibility of implementing higher order advection schemes will be considered. In the context of realistic simulations, we aim to assess better the turbulent mixing schemes parameterization, by means of sensitivity experiments. These, however, need to be carried on a simple configuration, since SES is computationally too demanding.
- The  $z$  coordinates used in the SES simulation, with resolution of 2 meters near the surface, is not the best choice to fully capture the diurnal variability of the sea surface temperature and the intra-seasonal variability of the mixed layer. To increase the vertical resolution and to cope with the large sea surface variability that are possible in a broad system like SES we aim at using a different choice for the vertical discretization such as  $z^*$  coordinate, with the possibility of including the Atlantic tides in the computations.
- The configuration of the SES system can be improved through the proper implementation of the river boundary conditions, including the river-sea continuum in the computational grid, with the aim of better representing the mixing processes.
- The simulation of one year produces reasonable results, but a longer run, though, is necessary to assess whether the above considerations still hold over a longer period. We aim at extending the length of the run to, at least, 3 years.

- We aim at using the SES configuration as a baseline to provide short term forecasting and analysis on the longer term. To this end, some further code developments suit this need such as the implementation of additional diagnostics (20 C° isotherm depth, vorticity etc.) and, notably, the inclusion of observations in the computations by means of data assimilation techniques.

# Appendix A

## River Dataset

The dataset of river discharge implemented in SES considers various sources, and it has been built according the following criteria:

- All the rivers of the CMEMS BLKSEA\_ANALYSIS\_FORECAST\_PHYS\_007\_001 are included in the dataset, with estimate of the inflow using monthly mean inflow provided by SESAME project (Ludwig et al. 2009)
- The rivers included in the Mediterranean have a mean annual discharge of  $50 \text{ m}^3/\text{s}$  at least.

Table A.1 gathers the source of the rivers with the annual mean of discharge. BSFS indicate rivers flowing into the Black Sea and the threshold of  $50 \text{ m}^3/\text{s}$  is not applied. River marked with MFS are the same implemented in the CMEMS MEDSEA\_ANALYSIS\_FORECAST\_PHY\_006\_013 that come from the PERSEUS project, if not specified differently. The locations of the rivers sources are shown in fig. 4.1.

Other sources are the Albanian Hydrometeorological Institute (AHI) for Mat, Seman and Vijose. The Marecchia and Po river discharge are from (Raicich 1996), the latter divided among the branches according to (Provini et al. 1992).

Dataset provide upstream estimates of the river discharge, such as in the case of Nile, which has two branches. In these cases, the final discharge is divided among the the river mouths proportionally to their width. Its discharge informations are from Global River Discharge Database (RivDIS) (Vörösmarty et al. 1998)(Vörösmarty et al. 1998).

Name	Source	mean annual discharge [m <sup>3</sup> /s]	Name	Source	mean annual discharge [m <sup>3</sup> /s]	Name	Source	mean annual discharge [m <sup>3</sup> /s]
Aapsta	BSFS	53.0	Evros	MFS	165.0	Po_scirocco	Raicich_1996	49.1
Adige	GRDC	223.8	Gediz	MFS	52.0	Po_tolle	Raicich_1996	242.5
Aksu	BSFS	15.0	Goksu	MFS	202.0	Po_tramontana	Raicich_1996	169.5
Arno	CREST_snc	99.3	Isonzo	Malacic_2009	110.4	Reka_mzymta	BSFS	61.0
Aspropotamo	MFS	108.0	Jucar	RivDIS	125.9	Reka_nekepsukho	BSFS	65.0
Aude	MFS	59.0	Karacennembongazi	BSFS	19.0	Reka_shakhe	BSFS	55.0
Ayancik	BSFS	14.0	Karasu	BSFS	286.0	Reka_tsemes	BSFS	14.0
Bartin	BSFS	136.0	Kinitrishi	BSFS	275.0	Reno	ARPA_EMR	49.3
Brenta	ARPA_VENETO	93.2	Kizilirmak	BSFS	240.0	Rhone1	MFS	170.0
Buene1	MFS	337.0	Kizilirmak2	BSFS	10.0	Rhone2	MFS	1536.0
Buene2	MFS	337.0	Krka	Pasartic_2004	56.5	Rioni	RivDIS	408.0
Bzipi	BSFS	33.0	Kurna	BSFS	19.0	Sakarya	BSFS	193.0
Catakli	BSFS	39.0	Lallya	BSFS	13.0	Sele	CREST_snc	69.6
Cayagzi	BSFS	17.0	Livenza	ARPA_VENETO	88.3	Seman	AHI	86.0
Cetina	Pasartic_2004	88.3	Manavgat_Nehri	MFS	121.0	Seyhan	MFS	200.0
Ceyhan	MFS	231.0	Marecchia	Raicich_1996	121.9	Shkumbi	AHI	58.7
Cilingoz	BSFS	11.0	Mat	AHI	87.4	Sile	ARPA_VENETO	52.9
Danube1	BSFS	570.0	Meandro	RivDIS	98.5	Sogutu	BSFS	16.0
Danube2	BSFS	570.0	Mejerda	MFS	58.0	Struma	MFS	81.0
Danube3	BSFS	570.0	Neretva	Pasartic_2004	332.1	Sumer	BSFS	27.0
Danube4	BSFS	570.0	Nile_east	RivDIS	253.1	Tabaguri	BSFS	148.0
Danube5	BSFS	570.0	Nile_west	RivDIS	1000.4	Tagliamento	ARPA_VENETO	96.9
Danube6	BSFS	570.0	Orontes	MFS	94.0	Terme	BSFS	24.0
Danube7	BSFS	1315.0	Ozluce	BSFS	42.0	Tevere1	MFS	90.0
Danube8	BSFS	1841.0	Peneo	MFS	66.0	Tevere2	MFS	90.0
Dnepr1	BSFS	260.0	Piave	ARPA_VENETO	54.3	Vardar	MFS	96.0
Dnepr2	BSFS	1170.0	Po_bastimento	Raicich_1996	66.5	Vaya	BSFS	14.0
Dnestr1	BSFS	156.0	Po_bonifazi	Raicich_1996	98.2	Vjiose	AHI	189.0
Dnestr2	BSFS	156.0	Po_dritta	Raicich_1996	477.0	Volturno	CREST_snc	103.1
Dyavolska	BSFS	15.0	Po_gnocca	Raicich_1996	251.9	Yesilirmak	BSFS	159.0
Ebro	RivDIS	1242.1	Po_goro	Raicich_1996	158.5			
Eurimedonte	MFS	85.0	Po_maistra	Raicich_1996	71.3			

Table A.1: Rivers included in the SES System

Italian regional agencies for environmental protection (Arpa Emilia Romagna and Arpa Veneto) have provided values for Brenta, Piave, Livenza, Sile and Reno.

The resulting water discharge is equally distributed in the Mediterranean and in Black Sea, as can be seen in fig. A.1. The rivers provide an average of 672 km<sup>3</sup>/year as the sum of 345 km<sup>3</sup>/year for the Mediterranean and 327 km<sup>3</sup>/year for the Black Sea.

All the river discharge data are corrected in pre-processing. Figure A.2 shows the climatology of the Saint George Branch of Danube River discharge with a maximum of nearly 2600 m<sup>3</sup>/s in May. Horizontal blue bars indicate the mean monthly values and the blue spline the linear interpolation. The correction to the linear interpolation, represented by the cyan line, is applied according to (Killworth 1996) and has more effect on the minimum and maximum values of discharge accounting for under/over estimation due to the interpolation. Taking the month of May as example the discharge is underestimated by the linear interpolation, that considers only lower values. The correction creates a new function (cyan line) whose monthly values match the climatology.

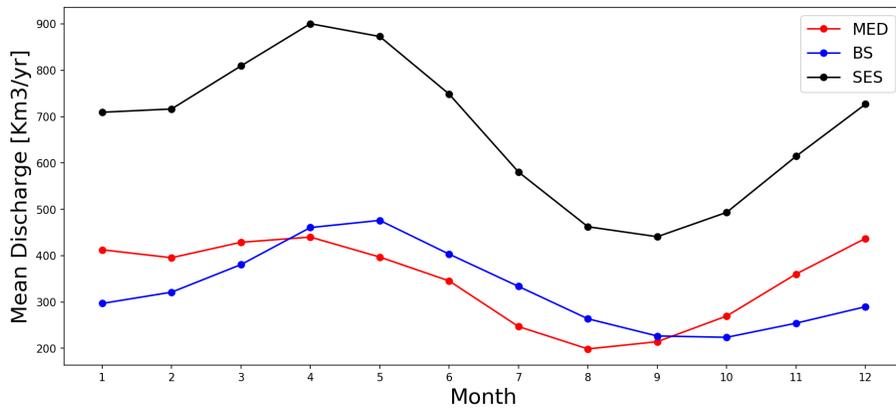


Figure A.1: Monthly Mean Climatology of River Discharge for Mediterranean and Black Sea

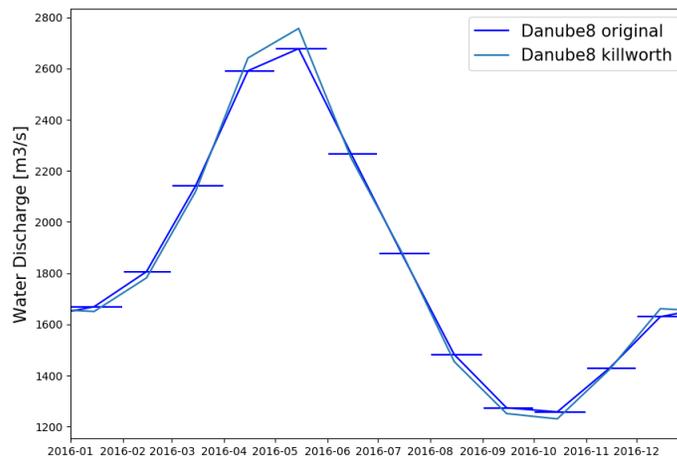


Figure A.2: Climatological Monthly Mean Discharge of the Saint George Branch of Danube River (Blue Line and horizontal bars). Cyan line represent the correction applied to the climatology according to (Killworth 1996)

# Appendix B

## Lock Exchange Experiment

The Lock Exchange Experiment (also referred to as Dam Break Experiment) is common in the literature to test the advection schemes and the baroclinic force in numerical ocean circulation models in case of modelling of buoyancy-driven flows or currents.

Two incompressible fluids of slightly different density confined in a rectangular channel, with impermeable walls and a free surface, are initially divided by a wall and motionless.

The fluid of higher density fills the left part of the channel, the rest of the channel by the lower density fluid.

When the wall is removed the buoyancy-driven flow develops with the high density fluid intruding the bottom right part of the channel and the light fluid mutually intruding the top of the channel in the opposite direction (see fig. B.1).

Eventually, after the interfacial wave has dissipated its energy, an equilibrium state will be achieved with two layers of equal thickness, the lighter fluid over the denser one.

A steady state solution for the problem has been obtained by (Benjamin 1968), where the adjustment current is driven solely by the gravity, without considering any other forces like rotation of Earth nor surface/bottom/lateral friction. The adjustment current depends on the initial density difference

$$u_g = \frac{1}{2} \sqrt{gH \frac{\delta\rho}{\rho_0}} \quad (\text{B.1})$$

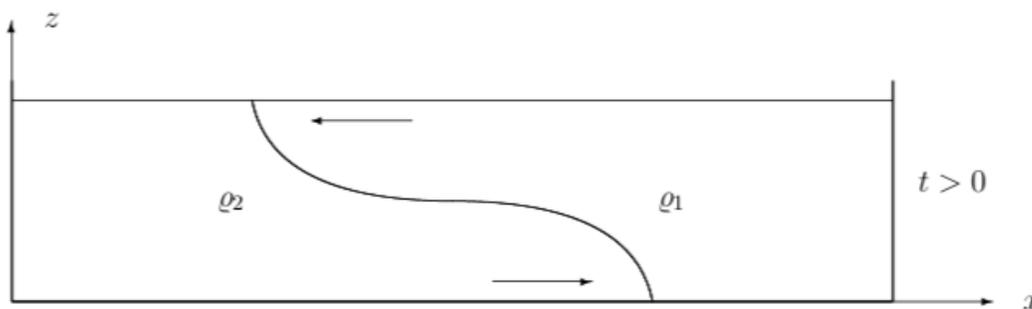


Figure B.1: Graphical representation of the gravity adjustment current in the Lock Exchange Experiment. Taken from (Jankowski 1999). The denser fluid with density  $\rho_2$  intrudes under the lighter fluid with density  $\rho_1$ .

where  $H$  is the depth of the channel.

The experiment has been carried with regular grid models (Ilicak, A.J. Adcroft, et al. 2012) and, more recently, with unstructured models (M. R. Petersen, Jacobsen, et al. 2015) (Y.J. Zhang et al. 2016) (Kärnä et al. 2018).

To test the capabilities of SHYFEM to reproduce the flow we created a 20 meters deep 3D channel, with an horizontal resolution of approximately 250 m (fig. B.2). For the geometry, with the exception of the  $y$  dimension, and the setup of the experiment we refer to (Ilicak, A.J. Adcroft, et al. 2012).

The channel is 64 km long, 2.5 km wide and 20 meters deep. The initial temperature  $T_0$  is set as follows

$$T_0 = \begin{cases} 5 \text{ C}^\circ & \text{if } x \leq 32 \text{ km} \\ 30 \text{ C}^\circ & \text{if } x > 32 \text{ km} \end{cases} \quad (\text{B.2})$$

and the equation of state is linear so that this temperature gap results into a density difference of  $5 \text{ kg/m}^3$ . Using the theoretical result for the adjustment current leads to a value of  $u_g \approx 0.5 \text{ m/s}$ .

The vertical is discretized into 20 evenly spaced 1 m thick layers. The simulation does not consider other forcings than gravity (Coriolis, bottom/lateral stress) nor explicit diffusion. Horizontal and vertical viscosity coefficients are set to  $10^{-2} \text{ m}^2/\text{s}$  and  $10^{-4} \text{ m}^2/\text{s}$  respectively.

Results obtained in the recent literature are commonly compared after 17 simulated hours, when the front location is theoretically at  $x = 62 \text{ km}$ . Results

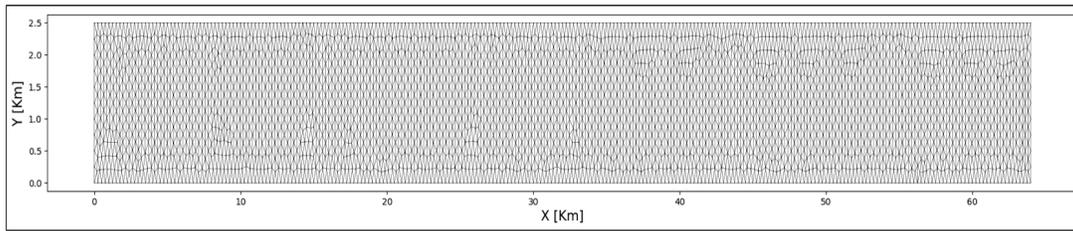


Figure B.2: Horizontal mesh for the Lock Exchange Experiment. The mesh has been created with Delaunay triangulation forced with a constant spatial resolution of 250 meters. The channel extends for 64 km along the x direction and is 2.5 km large. The depth is constant at 20 meters.

obtained with ROMS, MITgcm and MOM models by (Ilicak, A.J. Adcroft, et al. 2012) are shown in fig. (B.3). The gravity flow propagates with a velocity that almost equals the theoretical value. This is obtained with 3rd order accurate tracer advection schemes (7th in the case of MITgcm). The experiments carried with the unstructured grid models (M. R. Petersen, Jacobsen, et al. 2015) (Y.J. Zhang et al. 2016) (Kärnä et al. 2018) show analogous results.

SHYFEM has 3 types of horizontal advection schemes for tracers (Upwind, TVD and TVD2) and 2 choices for the vertical (Upwind and TVD). The TVD schemes for the horizontal advection use the average gradient (TVD) and the gradient of the upwind node (TVD2). In all TVD case a Superbee flux limiter is applied.

Figure (B.4) shows the results for SHYFEM after 17 hours with all the 6 possible combinations of advection schemes. In all cases the adjustment current has reached the position  $x \approx 52$  km, corresponding to a velocity of  $u \approx 0.32$  m/s, significantly smaller than the theoretical result or the results obtained in the literature and it is independent of the advection scheme used.

A study carried by (Bellafiore, McKiver, et al. 2018) that runs SHYFEM in both hydrostatic and non-hydrostatic mode in the same configuration used by (Lai et al. 2010) for the lock exchange, that considers a rectangular tank with size (x:0.8 m, y:0.008 m, z:0.1 m), obtains a result for the gravity adjustment current velocity that matches the theoretical value. The reason why SHYFEM does not reproduce the theoretical result on a larger scale has yet to be investigated extensively.

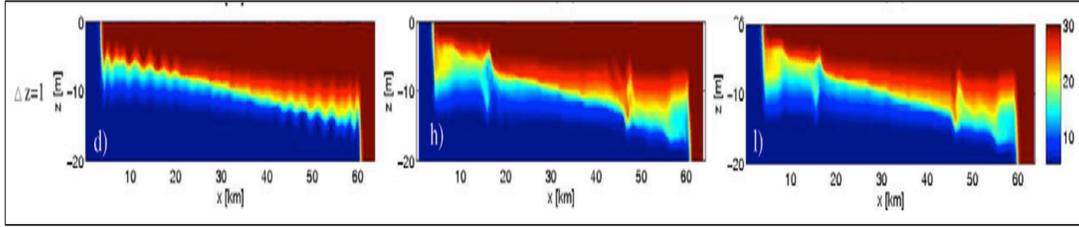


Figure B.3: Results of the lock exchange experiment from (Ilicak, A.J. Adcroft, et al. 2012) after 17 simulated hours. The 3 simulations have been carried with the models ROMS, MITgcm and MOM models respectively

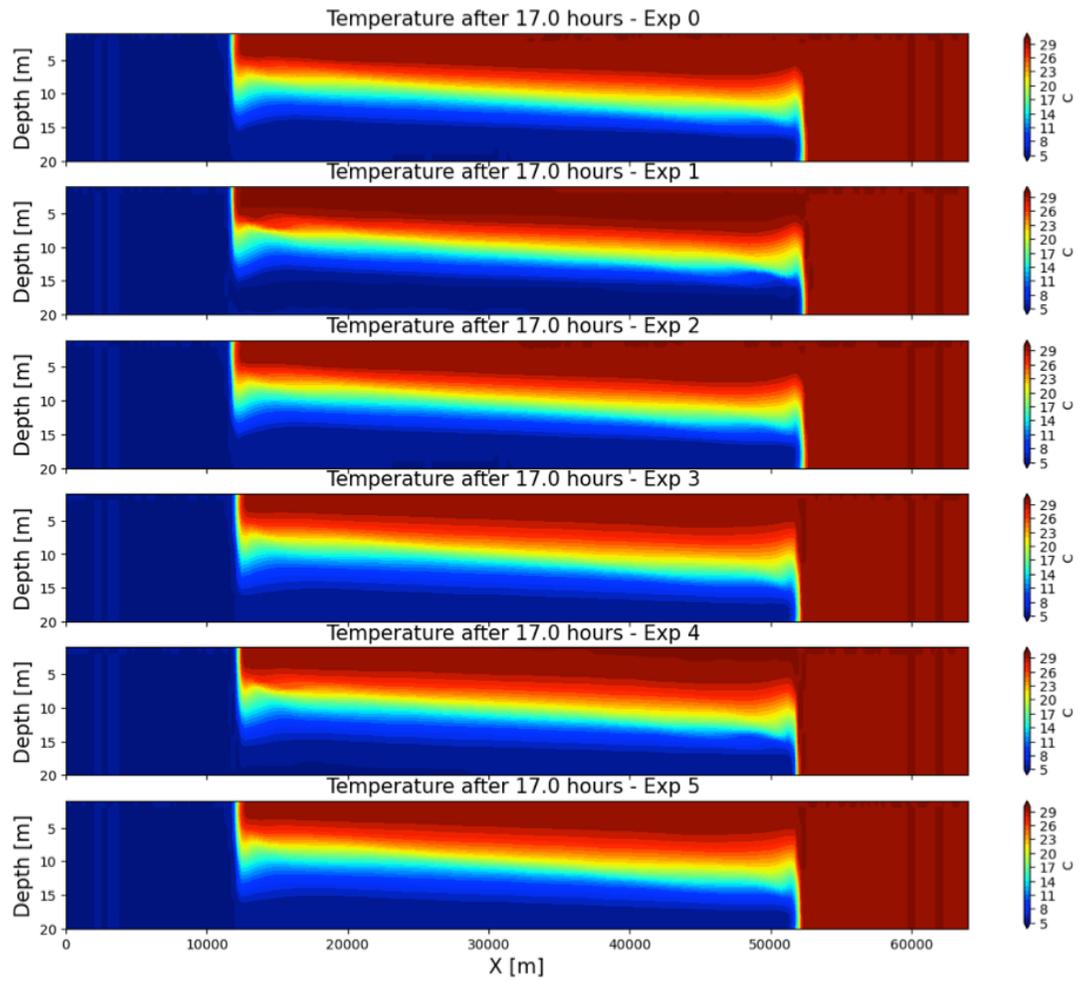


Figure B.4: Results of the lock exchange experiment with SHYFEM after 17 simulated hours. The number of the experiment indicates the combination of advection schemes used (see table B.1)

Name	Horizontal ADV	Vertical ADV
EXP0	TVD2	TVD
EXP1	TVD	TVD
EXP2	Upwind	TVD
EXP3	TVD2	Upwind
EXP4	TVD	Upwind
EXP5	Upwind	Upwind

Table B.1: Tracer Advection Schemes used in SHYFEM for lock exchange

# Acronyms

**BSFS** Black Sea Forecast System. 91, 111

**CIL** Cold Intermediate Layer. 13, 22

**CMEMS** Copernicus Marine Environment Monitoring Service. 17, 20, 21, 22, 91, 94, 108, 110

**CPU** Central Processing Unit. 30, 31, 62

**CUDA** Compute Unified Device Architecture. 30

**ECMWF** European Centre for Medium-range Weather Forecasts. 77

**FEM** Finite EleMent. 70

**FESOM2** Finite Element Sea-ice Ocean Model 2. 27, 59

**FSE** Free Surface Equation. 64, 66, 122

**FVCOM** Finite Volume Community Ocean Model. 59

**GEBCO** GEneral Bathymetric Chart of the Oceans. 90

**GPU** Graphics Processing Unit. 30, 31, 62

**GSHHS** Global Self-consistent, Hierarchical, High-resolution Geography. 89

**MFS** Mediterranean Forecast System. 18, 20, 91, 92, 94, 104, 110

**MITgcm** Massachussets Institute of Technology general circulation model. 17, 24, 33

- MLD** Mixed Layer Depth. 108
- MMBA** Mediterranean-Marmara-Black Sea-Azov system modelled in (C. Ferrarin, Bellafiore, et al. 2018). 26, 34
- MOM** Modular Ocean Model. 18, 21, 33
- MPI** Message Passing Interface. vi, 30, 31, 32, 33, 34, 57, 59, 60, 62, 66, 67, 68, 69, 71, 74, 75, 76, 79, 80, 81, 82, 84, 85, 86, 120, 122
- NCO** NetCDF Operators. 79
- NEMO** Nucleus for European Modeling of Oceans. 20, 21, 33
- OpenACC** Open ACCelerators. 30
- OpenMP** Open MultiProcessing. 59
- PDE** Partial Differential Equation. 31
- PETSc** Portable Extensible Toolkit for Scientific computations. 64, 66, 74, 75, 79, 85
- PSU** Practical Salinity Unit [0.001]. 96, 97, 99, 111, 112, 117, 121
- RMSE** Root Mean Squared Error. 79, 80, 111
- ROMS** Regional Ocean Modeling System. 23
- SANIFS** Southern Adriatic Northern-Ionian Forecast System. 59, 67, 76, 78, 82, 83, 84, 85
- SCHISM** Semi-implicit Cross-scale Hydroscience Integrated System Model. 59
- SES** Southern European Seas. v, vi, 2, 3, 4, 5, 6, 16, 17, 26, 34, 35, 88, 89, 90, 91, 92, 93, 95, 96, 98, 99, 102, 104, 105, 106, 108, 110, 111, 112, 113, 115, 117, 119, 120, 121, 122
- SHYFEM** System of Hydrodynamic Finite Element Modules. v, 17, 22, 26, 28, 35, 36, 37, 38, 41, 43, 44, 45, 48, 55, 58, 59, 60, 63, 64, 66, 69, 72, 74, 76, 117, 119

**SHYMPI** SHYfem MPI. 69, 73, 74, 75, 78, 79, 80, 82, 86, 122

**SYPD** Simulated Years Per Day. 34, 97

**TSS** Turkish Strait System. v, 2, 14, 15, 23, 24, 25

# Bibliography

- Adani, M., S. Dobricic, and N. Pinardi (2011). "Quality assessment of a 1985-2007 Mediterranean sea reanalysis". In: *Journal of Atmospheric and Oceanic Technology* 28.4, pp. 569–589.
- Adcroft, A. and J.-M. Campin (2004). "Rescaled height coordinates for accurate representation of free-surface flows in ocean circulation models". In: *Ocean Modelling* 7.3-4, pp. 269–284.
- Alpar, B. and H. Yüce (1998). "Sea-level Variations and their Interactions Between the Black Sea and the Aegean Sea". In: *Estuarine, Coastal and Shelf Science* 46.5, pp. 609–619. ISSN: 0272-7714. DOI: <https://doi.org/10.1006/ecss.1997.0285>. URL: <http://www.sciencedirect.com/science/article/pii/S0272771497902857>.
- Ambrose, Robert, Tim Wool, and James Martin (Jan. 1993). "The water quality analysis simulation program, WASP5, Part A: Model documentation". In:
- Artegiani, A. et al. (1997). "The Adriatic Sea General Circulation. Part I: Air?Sea Interactions and Water Mass Structure". In: *Journal of Physical Oceanography* 27.8, pp. 1492–1514. DOI: 10.1175/1520-0485(1997)027<1492:TASGCP>2.0.CO;2. URL: [https://journals.ametsoc.org/view/journals/phoc/27/8/1520-0485\\_1997\\_027\\_1492\\_tasgcp\\_2.0.co\\_2.xml](https://journals.ametsoc.org/view/journals/phoc/27/8/1520-0485_1997_027_1492_tasgcp_2.0.co_2.xml).
- Aydoğdu, A. et al. (2018). "Circulation of the Turkish Straits System under inter-annual atmospheric forcing". In: *Ocean Science* 14.5, pp. 999–1019.
- Bajo, M. et al. (2014). "The water circulation near the danube delta and the romanian coast modelled with finite elements". In: *Continental Shelf Research* 78, pp. 62–74.
- Balay, Satish, Shrirang Abhyankar, et al. (2020). *PETSc Users Manual*. Tech. rep. ANL-95/11 - Revision 3.13. Argonne National Laboratory. URL: <https://www.mcs.anl.gov/petsc>.

- Balay, Satish, William D. Gropp, et al. (1997). "Efficient Management of Parallelism in Object Oriented Numerical Software Libraries". In: *Modern Software Tools in Scientific Computing*. Ed. by E. Arge, A. M. Bruaset, and H. P. Langtangen. Birkhäuser Press, pp. 163–202.
- Beckers, J.-M. (1991). "Application of the GHER 3D general circulation model to the Western Mediterranean". In: *Journal of Marine Systems* 1.4, pp. 315–332.
- Beckers, J.M. et al. (2002). "Modelling the Danube-influenced North-western Continental Shelf of the Black Sea. I: Hydrodynamical Processes Simulated by 3-D and Box Models". In: *Estuarine, Coastal and Shelf Science* 54.3, pp. 453–472. ISSN: 0272-7714. DOI: <https://doi.org/10.1006/ecss.2000.0658>. URL: <http://www.sciencedirect.com/science/article/pii/S0272771400906589>.
- Bellafore, D., W.J. McKiver, et al. (2018). "The importance of modeling nonhydrostatic processes for dense water reproduction in the Southern Adriatic Sea". In: *Ocean Modelling* 125, pp. 22–38.
- Bellafore, D. and G. Umgiesser (2010). "Hydrodynamic coastal processes in the North Adriatic investigated with a 3D finite element model". In: *Ocean Dynamics* 60.2, pp. 255–273.
- Benjamin, T.B. (1968). "Gravity currents and related phenomena". In: *Journal of Fluid Mechanics* 31.2, pp. 209–248.
- Bernie, D. J. et al. (2005). "Modeling Diurnal and Intraseasonal Variability of the Ocean Mixed Layer". In: *Journal of Climate* 18.8, pp. 1190–1202. DOI: 10.1175/JCLI3319.1. URL: <https://journals.ametsoc.org/view/journals/clim/18/8/jcli3319.1.xml>.
- Beşiktepe, Şükrü T. et al. (1994). "The circulation and hydrography of the Marmara Sea". In: *Progress in Oceanography* 34.4, pp. 285–334. ISSN: 0079-6611. DOI: [https://doi.org/10.1016/0079-6611\(94\)90018-3](https://doi.org/10.1016/0079-6611(94)90018-3). URL: <http://www.sciencedirect.com/science/article/pii/0079661194900183>.
- Blumberg, A.F. and G.L. Mellor (1987). "A description of a three-dimensional coastal ocean circulation model". In: *Three-dimensional Coastal Ocean Models*.
- Boyer Montégut, Clément de et al. (2004). "Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology". In: *Journal of Geophysical Research: Oceans* 109.C12. DOI: <https://doi.org/10.1029/2004JC002378>.

- Burchard, H. and O. Petersen (1999). "Models of turbulence in the marine environment - A comparative study of two-equation turbulence models". In: *Journal of Marine Systems* 21.1-4, pp. 29–53.
- Campin, J.-M. et al. (2004). "Conservation of properties in a free-surface model". In: *Ocean Modelling* 6.3-4, pp. 221–244.
- CHARNEY, J. G., R. FJÖRTOFT, and J. Von NEUMANN (1950). "Numerical Integration of the Barotropic Vorticity Equation". In: *Tellus* 2.4, pp. 237–254. DOI: <https://doi.org/10.1111/j.2153-3490.1950.tb00336.x>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.2153-3490.1950.tb00336.x>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.2153-3490.1950.tb00336.x>.
- Chen, Changsheng, Hedong Liu, and Robert C. Beardsley (Jan. 2003). "An Unstructured Grid, Finite-Volume, Three-Dimensional, Primitive Equations Ocean Model: Application to Coastal Ocean and Estuaries". In: *Journal of Atmospheric and Oceanic Technology* 20.1, pp. 159–186. ISSN: 0739-0572. DOI: 10.1175/1520-0426(2003)020<0159:AUGFVT>2.0.CO;2. eprint: [https://journals.ametsoc.org/jtech/article-pdf/20/1/159/3316448/1520-0426\(2003\)020\\_0159\\_augfvt\\_2\\_0\\_co\\_2.pdf](https://journals.ametsoc.org/jtech/article-pdf/20/1/159/3316448/1520-0426(2003)020_0159_augfvt_2_0_co_2.pdf). URL: [https://doi.org/10.1175/1520-0426\(2003\)020%3C0159:AUGFVT%3E2.0.CO;2](https://doi.org/10.1175/1520-0426(2003)020%3C0159:AUGFVT%3E2.0.CO;2).
- Chevalier, C. and F. Pellegrini (2008). "PT-Scotch: A tool for efficient parallel graph ordering". In: *Parallel Computing* 34.6. Parallel Matrix Algorithms and Applications, pp. 318–331. ISSN: 0167-8191. DOI: <https://doi.org/10.1016/j.parco.2007.12.001>. URL: <http://www.sciencedirect.com/science/article/pii/S0167819107001342>.
- Chiggiato, J. et al. (2012). "Dynamics of the circulation in the Sea of Marmara: Numerical modeling experiments and observations from the Turkish straits system experiment". In: *Ocean Dynamics* 62.1, pp. 139–159.
- Ciliberti, S. A. et al. (2020). *Black Sea Analysis and Forecast (CMEMS BS-Currents, EAS3 system) (Version 1)*. DOI: [https://doi.org/10.25423/CMCC/BLKSEA\\_ANALYSIS\\_FORECAST\\_PHYS\\_007\\_001\\_EAS3](https://doi.org/10.25423/CMCC/BLKSEA_ANALYSIS_FORECAST_PHYS_007_001_EAS3).
- Clementi, E. et al. (2019). *Mediterranean Sea Analysis and Forecast (CMEMS MED-Currents, EAS5 system)*. DOI: [https://doi.org/10.25423/CMCC/MEDSEA\\_ANALYSIS\\_FORECAST\\_PHY\\_006\\_013\\_EAS5](https://doi.org/10.25423/CMCC/MEDSEA_ANALYSIS_FORECAST_PHY_006_013_EAS5).

- Coles, Victoria J. et al. (2013). "The pathways and properties of the Amazon River Plume in the tropical North Atlantic Ocean". In: *Journal of Geophysical Research: Oceans* 118.12, pp. 6894–6913. DOI: <https://doi.org/10.1002/2013JC008981>.
- Covelli, S. et al. (2007). "Transport and dispersion of particulate Hg associated with a river plume in coastal Northern Adriatic environments". In: *Marine Pollution Bulletin* 55.10-12, pp. 436–450.
- Cowles, Geoffrey (May 2008). "Parallelization of the FVCOM Coastal Ocean Model." In: *IJHPCA* 22, pp. 177–193. DOI: 10.1177/1094342007083804.
- Cucco, A. et al. (2012). "A high-resolution real-time forecasting system for predicting the fate of oil spills in the Strait of Bonifacio (western Mediterranean Sea)". In: *Marine Pollution Bulletin* 64.6, pp. 1186–1200.
- D'Ortenzio, Fabrizio et al. (June 2005). "Seasonal variability of the mixed layer depth in the Mediterranean Sea as derived from in situ profiles". In: *Geophysical Research Letters* 32. DOI: 10.1029/2005GL022463.
- Danilov, S., G. Kivman, and J. Schröter (2004). "A finite-element ocean model: Principles and evaluation". In: *Ocean Modelling* 6.2, pp. 125–150.
- Danilov, S., D. Sidorenko, et al. (2017). "The finite-volume sea ice-Ocean model (FESOM2)". In: *Geoscientific Model Development* 10.2, pp. 765–789.
- Demyshev, S.G., S.V. Dovgaya, and V.A. Ivanov (2012). "Numerical modeling of the influence of exchange through the Bosphorus and Dardanelles Straits on the hydrophysical fields of the Marmara Sea". In: *Izvestiya - Atmospheric and Ocean Physics* 48.4, pp. 418–426.
- Dietrich, David (Jan. 1998). "Application of a modified Arakawa 'a' grid ocean model having reduced numerical dispersion to the Gulf of Mexico circulation". In: *Dynamics of Atmospheres and Oceans* 27, pp. 201–217. DOI: 10.1016/S0377-0265(97)00009-2.
- Dukowicz, J.K. and R.D. Smith (1994). "Implicit free-surface method for the Bryan-Cox-Semtner ocean model". In: *Journal of Geophysical Research* 99.C4, pp. 7991–8014.
- Fairall, C.W. et al. (2003). "Bulk parameterization of air-sea fluxes: Updates and verification for the COARE algorithm". In: *Journal of Climate* 16.4, pp. 571–591.
- Federico, I. et al. (2017). "Coastal ocean forecasting with an unstructured grid model in the southern Adriatic and northern Ionian seas". In: *Natural Hazards*

- and Earth System Sciences* 17.1, pp. 45–59. ISSN: 15618633. DOI: 10.5194/nhess-17-45-2017.
- Felten, Frédéric N. and Thomas S. Lund (2006). “Kinetic energy conservation issues associated with the collocated mesh scheme for incompressible flow”. In: *Journal of Computational Physics* 215.2, pp. 465–484. ISSN: 0021-9991. DOI: <https://doi.org/10.1016/j.jcp.2005.11.009>. URL: <http://www.sciencedirect.com/science/article/pii/S0021999105004997>.
- Fernandez Montblanc, Tomas et al. (Dec. 2019). “A Pan-European high resolution storm surge hindcast”. In: *Environment international* 135, p. 105367. DOI: 10.1016/j.envint.2019.105367.
- Ferrarin, C., D. Bellafiore, et al. (2018). “Tidal dynamics in the inter-connected Mediterranean, Marmara, Black and Azov seas”. In: *Progress in Oceanography* 161, pp. 102–115.
- Ferrarin, C., S. Davolio, et al. (2019). “Cross-scale operational oceanography in the Adriatic Sea”. In: *Journal of Operational Oceanography* 12.2, pp. 86–103.
- Ferrarin, Christian et al. (2010). “Modelling fluxes of water and sediment between Venice Lagoon and the sea”. In: *Continental Shelf Research* 30.8. The Coastal Morphodynamics of Venice Lagoon and its Inlets, pp. 904–914. ISSN: 0278-4343. DOI: <https://doi.org/10.1016/j.csr.2009.08.014>. URL: <http://www.sciencedirect.com/science/article/pii/S0278434309002544>.
- Fofonoff, N. and R. Millard (Jan. 1983). “Algorithms for Computation of Fundamental Properties of Seawater”. In: *UNESCO Tech. Pap. Mar. Sci.* 44.
- G., Madec and NEMO System Team (n.d.). *NEMO ocean engine*. Scientific Notes of Climate Modelling Center 27. Zenodo. DOI: 10.5281/zenodo.1464816.
- G. Karypis, V. Kumar (1999). *METIS: A software package for partitioning unstructured graphs, partitioning meshes, and computing fill-reducing orderings of sparse matrices, version 4.0*. University of Minnesota, Department of Computer Science and Engineering.
- G. Karypis K. Schloegel, V. Kumar (1997). *PARMETIS: Parallel graph partitioning and sparse matrix ordering library*. University of Minnesota, Department of Computer Science and Engineering.
- Garvine, Richard W. (1999). “Penetration of Buoyant Coastal Discharge onto the Continental Shelf: A Numerical Model Experiment”. In: *Journal of Physical Oceanography* 29.8, pp. 1892–1909. DOI: 10.1175/1520-0485(1999)029<1892:

- POBCDO>2.0.CO;2. URL: [https://journals.ametsoc.org/view/journals/phoc/29/8/1520-0485\\_1999\\_029\\_1892\\_pobcdo\\_2.0.co\\_2.xml](https://journals.ametsoc.org/view/journals/phoc/29/8/1520-0485_1999_029_1892_pobcdo_2.0.co_2.xml).
- Goldberg, David (Mar. 1991). "What Every Computer Scientist Should Know about Floating-Point Arithmetic". In: *ACM Comput. Surv.* 23.1, pp. 5–48. ISSN: 0360-0300. DOI: 10.1145/103162.103163. URL: <https://doi.org/10.1145/103162.103163>.
- Grégoire, M., C. Raick, and K. Soetaert (2008). "Numerical modeling of the central Black Sea ecosystem functioning during the eutrophication phase". English. In: *Progress in Oceanography* 76.3, pp. 286–333. ISSN: 0079-6611. DOI: 10.1016/j.pocean.2008.01.002.
- Griffies, Stephen M. (2012). *GFDL Ocean Group Technical Report No. 7 - Elements of the Modular Ocean Model (MOM)*. NOAA/Geophysical Fluid Dynamics Laboratory.
- Group, The Wamdi (Dec. 1988). "The WAM Model—A Third Generation Ocean Wave Prediction Model". In: *Journal of Physical Oceanography* 18.12, pp. 1775–1810. ISSN: 0022-3670.
- Guarnieri, A. et al. (2013). "Impact of tides in a baroclinic circulation model of the Adriatic Sea". In: *Journal of Geophysical Research: Oceans* 118.1, pp. 166–183.
- Gunduz, M., E. Özsoy, and R. Hordoir (2020). "A model of Black Sea circulation with strait exchange (2008-2018)". In: *Geoscientific Model Development* 13.1, pp. 121–138.
- Gurses, O. (2016). "Dynamics of the Turkish straits system :A numerical study with a finite element ocean model based on an unstructured grid approach". PhD Thesis. PhD thesis. Middle East Technical University.
- Haidvogel, D.B. et al. (Mar. 2008). "Ocean forecasting in terrain-following coordinates: Formulation and skill assessment of the Regional Ocean Modeling System". In: *Journal of Computational Physics* 227, pp. 3595–3624. DOI: 10.1016/j.jcp.2007.06.016.
- Hellerman, Sol and Mel Rosenstein (July 1983). "Normal Monthly Wind Stress Over the World Ocean with Error Estimates". In: *Journal of Physical Oceanography* 13.7, pp. 1093–1104. ISSN: 0022-3670. DOI: 10.1175/1520-0485(1983)013<1093:NMWSOT>2.0.CO;2. eprint: [https://journals.ametsoc.org/jpo/article-pdf/13/7/1093/4406266/1520-0485\(1983\)013\1093\\\_nmwsot](https://journals.ametsoc.org/jpo/article-pdf/13/7/1093/4406266/1520-0485(1983)013\1093\_nmwsot)

- \_2\\_0\\_co\\_2.pdf. URL: [https://doi.org/10.1175/1520-0485\(1983\)013%3C1093:NMWSOT%3E2.0.CO;2](https://doi.org/10.1175/1520-0485(1983)013%3C1093:NMWSOT%3E2.0.CO;2).
- Hendrickson, Bruce and Tamara G Kolda (2000). "Graph partitioning models for parallel computing". In: *Parallel Computing* 26.12. Graph Partitioning and Parallel Computing, pp. 1519–1534. ISSN: 0167-8191. DOI: [https://doi.org/10.1016/S0167-8191\(00\)00048-X](https://doi.org/10.1016/S0167-8191(00)00048-X). URL: <http://www.sciencedirect.com/science/article/pii/S016781910000048X>.
- Huret, Martin et al. (Oct. 2007). "Dispersal modeling of fish early life stages: Sensitivity with application to Atlantic cod in the western Gulf of Maine". In: *Marine Ecology Progress Series (0171-8630) (Inter-Research), 2007-10 , Vol. 347 , P. 261-274* 347. DOI: 10.3354/meps06983.
- Iacono, R. et al. (2013). "Seasonal variability of the tyrrhenian sea surface geostrophic circulation as assessed by altimeter data". In: *Journal of Physical Oceanography* 43.8, pp. 1710–1732.
- Ilicak, M., A.J. Adcroft, et al. (2012). "Spurious diapycnal mixing and the role of momentum closure". In: *Ocean Modelling* 45-46, pp. 37–58.
- Ilicak, M., T.M. Özgökmen, et al. (2009). "Non-hydrostatic modeling of exchange flows across complex geometries". In: *Ocean Modelling* 29.3, pp. 159–175.
- Iovino, D. et al. (2016). "A 1/16° eddy simulation of the global NEMO sea-ice-ocean system". In: *Geoscientific Model Development* 9.8, pp. 2665–2684.
- Ivanov, L. I., Ş. Beşiktepe, and E. Özsoy (1997). "The Black Sea Cold Intermediate Layer". In: *Sensitivity to Change: Black Sea, Baltic Sea and North Sea*. Ed. by Emin Özsoy and Alexander Mikaelyan. Dordrecht: Springer Netherlands, pp. 253–264. ISBN: 978-94-011-5758-2. DOI: 10.1007/978-94-011-5758-2\_20. URL: [https://doi.org/10.1007/978-94-011-5758-2\\_20](https://doi.org/10.1007/978-94-011-5758-2_20).
- Jankowski, Jacek (Jan. 1999). "A non-hydrostatic model for free surface flows". PhD thesis.
- Jarosz, E. et al. (2011). "Observed volume fluxes in the Bosphorus Strait". In: *Geophysical Research Letters* 38.21. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-81155124357&doi=10.1029>.
- Jerlov, N. (1968). *Optical Oceanography*. Vol. 5. C, p. 7.
- Jofre, Lluís et al. (2014). "Conservation Properties of Unstructured Finite-Volume Mesh Schemes for the Navier-Stokes Equations". In: *Numerical Heat Transfer*,

- Part B: Fundamentals* 65.1, pp. 53–79. DOI: 10.1080/10407790.2013.836335. URL: <https://doi.org/10.1080/10407790.2013.836335>.
- Kara, A.B., A.J. Wallcraft, and H.E. Hurlburt (2007). “A correction for land contamination of atmospheric variables near land-sea boundaries”. In: *Journal of Physical Oceanography* 37.4, pp. 803–818.
- Kara, A.B., A.J. Wallcraft, H.E. Hurlburt, and E.V. Stanev (2008). “Air-sea fluxes and river discharges in the Black Sea with a focus on the Danube and Bosphorus”. In: *Journal of Marine Systems* 74.1-2, pp. 74–95.
- Kärnä, T. et al. (2018). “Thetis coastal ocean model: discontinuous Galerkin discretization for the three-dimensional hydrostatic equations”. In: *Geoscientific Model Development* 11.11, pp. 4359–4382. DOI: 10.5194/gmd-11-4359-2018. URL: <https://gmd.copernicus.org/articles/11/4359/2018/>.
- Kärnä Tuomas, T., V. Legat, and E. Deleersnijder (2013). “A baroclinic discontinuous Galerkin finite element model for coastal flows”. In: *Ocean Modelling* 61, pp. 1–20.
- Katavouta, Anna and Keith R. Thompson (2016). “Downscaling ocean conditions with application to the Gulf of Maine, Scotian Shelf and adjacent deep ocean”. In: *Ocean Modelling* 104, pp. 54–72. ISSN: 1463-5003. DOI: <https://doi.org/10.1016/j.ocemod.2016.05.007>. URL: <http://www.sciencedirect.com/science/article/pii/S1463500316300373>.
- Killworth, Peter D. (Jan. 1996). “Time Interpolation of Forcing Fields in Ocean Models”. In: *Journal of Physical Oceanography* 26.1, pp. 136–143. ISSN: 0022-3670. DOI: 10.1175/1520-0485(1996)026<0136:TIOFFI>2.0.CO;2. eprint: [https://journals.ametsoc.org/jpo/article-pdf/26/1/136/4424780/1520-0485\(1996\)026\\_0136\\_tioffi\\_2\\_0\\_co\\_2.pdf](https://journals.ametsoc.org/jpo/article-pdf/26/1/136/4424780/1520-0485(1996)026_0136_tioffi_2_0_co_2.pdf). URL: [https://doi.org/10.1175/1520-0485\(1996\)026%3C0136:TIOFFI%3E2.0.CO;2](https://doi.org/10.1175/1520-0485(1996)026%3C0136:TIOFFI%3E2.0.CO;2).
- Knudsen, Martin (1900). “Ein hydrographischer lehrsatz”. In: *Annalen der Hydrographie und Maritimen Meteorologie* 28.7, pp. 316–320.
- Kondo, J. (1975). “Air-sea bulk transfer coefficients in diabatic conditions”. In: *Boundary-Layer Meteorology* 9.1, pp. 91–112.
- Korotenko, K.A. (2015). “Modeling mesoscale circulation of the Black Sea”. In: *Oceanology* 55.6, pp. 820–826.

- Kourafalou, Vassiliki H. (1999). "Process studies on the Po River plume, North Adriatic Sea". In: *Journal of Geophysical Research: Oceans* 104.C12, pp. 29963–29985. DOI: <https://doi.org/10.1029/1999JC900217>.
- Lai, Zhigang et al. (2010). "A nonhydrostatic version of FVCOM: 1. Validation experiments". In: *Journal of Geophysical Research: Oceans* 115.C11. DOI: <https://doi.org/10.1029/2009JC005525>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2009JC005525>. URL: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JC005525>.
- Lermusiaux, P.F.J. and A.R. Robinson (2001). "Features of dominant mesoscale variability, circulation patterns and dynamics in the Strait of Sicily". In: *Deep Sea Research Part I: Oceanographic Research Papers* 48.9, pp. 1953–1997. ISSN: 0967-0637. DOI: [https://doi.org/10.1016/S0967-0637\(00\)00114-X](https://doi.org/10.1016/S0967-0637(00)00114-X). URL: <http://www.sciencedirect.com/science/article/pii/S096706370000114X>.
- Li, Siqi et al. (2020). "Impacts of Oceanic Mixed Layer on Hurricanes: A Simulation Experiment With Hurricane Sandy". In: *Journal of Geophysical Research: Oceans* 125.11. e2019JC015851 2019JC015851, e2019JC015851. DOI: <https://doi.org/10.1029/2019JC015851>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019JC015851>. URL: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JC015851>.
- Ludwig, Wolfgang et al. (Mar. 2009). "River discharges of water and nutrients to the Mediterranean and Black Sea: Major drivers for ecosystem changes during past and future decades?" In: *Progress In Oceanography* 80, pp. 199–217. DOI: 10.1016/j.pocean.2009.02.001.
- Madec, Gurvan et al. (Dec. 1998). "OPA 8.1 Ocean General Circulation Model reference manual". In: *Notes du pôle de modélisation, laboratoire d'océanographie dynamique et de climatologie, Institut Pierre Simon Laplace des sciences de l'environnement global* 11.
- Maderich, Vladimir, Yuriy Ilyin, and Evgeny Lemeshko (July 2015). "Seasonal and interannual variability of the water exchange in the Turkish Straits System estimated by modelling". In: *Mediterranean Marine Science* 16, pp. 444–459. DOI: 10.12681/mms.1103.
- Maderich, Vladimir and S. Konstantinov (Aug. 2002). "Seasonal Dynamics of the System Sea-Strait: Black Sea–Bosphorus Case Study". In: *Estuarine Coastal and Shelf Science* 55, pp. 183–196. DOI: 10.1006/ecss.2001.0895.

- Marchesiello, P., J.C. McWilliams, and A. Shchepetkin (2001). "Open boundary conditions for long-term integration of regional oceanic models". In: *Ocean Modelling* 3.1-2, pp. 1–20.
- Marsh, Robert et al. (2009). "Recent changes in the North Atlantic circulation simulated with eddy-permitting and eddy-resolving ocean models". In: *Ocean Modelling* 28.4, pp. 226–239. ISSN: 1463-5003. DOI: <https://doi.org/10.1016/j.ocemod.2009.02.007>. URL: <http://www.sciencedirect.com/science/article/pii/S1463500309000274>.
- Marshall, John et al. (1997). "Hydrostatic, quasi-hydrostatic, and nonhydrostatic ocean modeling". In: *Journal of Geophysical Research: Oceans* 102.C3, pp. 5733–5752. DOI: <https://doi.org/10.1029/96JC02776>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/96JC02776>. URL: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JC02776>.
- Mauri, E. et al. (2019). "On the variability of the circulation and water mass properties in the eastern Levantine Sea between September 2016-August 2017". In: *Water (Switzerland)* 11.9.
- Miladinova, S. et al. (2018). "Formation and changes of the Black Sea cold intermediate layer". In: *Progress in Oceanography* 167, pp. 11–23. ISSN: 0079-6611. DOI: <https://doi.org/10.1016/j.pocean.2018.07.002>. URL: <http://www.sciencedirect.com/science/article/pii/S0079661117303117>.
- Nadia, Pinardi et al. (May 2006). "The physical, sedimentary and ecological structure and variability of shelf areas in the Mediterranean Sea." In:
- Neumeier, U. et al. (2008). "Sedtrans05: An improved sediment-transport model for continental shelves and coastal waters with a new algorithm for cohesive sediments". In: *Computers and Geosciences* 34.10, pp. 1223–1242.
- Nittis, Kostas and L Perivoliotis (Dec. 2002). "Circulation and hydrological characteristics of the North Aegean Sea: A contribution from real-time buoy measurements". In: *Medit. Mar. Sci* 3, pp. 21–31. DOI: 10.12681/mms.255.
- Oddo, P. et al. (2009). "A nested Atlantic-Mediterranean Sea general circulation model for operational forecasting". In: *Ocean Science* 5.4, pp. 461–473.
- Özsoy, Emin et al. (2001). "Mixing in the Bosphorus Strait and the Black Sea continental shelf: observations and a model of the dense water outflow". In: *Journal of marine systems* 31.1-3, pp. 99–135.

- Pacanowski, RC, Keith Dixon, and Anthony Rosati (Jan. 1990). "The GFDL modular ocean model users guide: Version 1.0". In: *GFDL Ocean Group Tech. Rep 2*, pp. 08542–0308.
- Palazov, A. et al. (2019). "Black sea observing system". In: *Frontiers in Marine Science* 6.JUN.
- Palma, M. et al. (2020). "Short-term, linear, and non-linear local effects of the tides on the surface dynamics in a new, high-resolution model of the Mediterranean Sea circulation". In: *Ocean Dynamics* 70.7, pp. 935–963.
- Panin, Nicolae (2003). "The Danube Delta. Geomorphology and Holocene Evolution: a Synthesis / Le delta du Danube. Géomorphologie et évolution holocène : une synthèse". eng. In: ISSN: 1266-5304. DOI: 10.3406/morfo.2003.1188. URL: [https://www.persee.fr/doc/morfo\\_1266-5304\\_2003\\_num\\_9\\_4\\_1188](https://www.persee.fr/doc/morfo_1266-5304_2003_num_9_4_1188).
- (2004). *Danube river, Giurgiu, daily average water level and discharge, 2002*. data set. DOI: 10.1594/PANGAEA.136566. URL: <https://doi.org/10.1594/PANGAEA.136566>.
- Pascolo, E. et al. (2016). "OpenMP tasks: Asynchronous programming made easy". In: *2016 International Conference on High Performance Computing Simulation (HPCS)*. Institute of Electrical and Electronics Engineers Inc., pp. 901–907. ISBN: 9781509020881. DOI: 10.1109/HPCSim.2016.7568430.
- Pascolo, Eric (2015). "Shyferm parallelization: an innovative task approach for coastal environment FEM software". Master Thesis in High Performance Computing. URL: <https://iris.sissa.it/handle/20.500.11767/5127?mode=full.369#.X5WuY1NKjGI>.
- Peneva, E. et al. (2001). "Water transport in the Bosphorus Straits estimated from hydro-meteorological and altimeter data: Seasonal to decadal variability". In: *Journal of Marine Systems* 31.1-3, pp. 21–33.
- Pessini, F. et al. (2018). "Mesoscale eddies in the algerian basin: Do they differ as a function of their formation site?" In: *Ocean Science* 14.4, pp. 669–688.
- Petersen, Mark R., Xylar S. Asay-Davis, et al. (2019). "An Evaluation of the Ocean and Sea Ice Climate of E3SM Using MPAS and Interannual CORE-II Forcing". In: *Journal of Advances in Modeling Earth Systems* 11.5, pp. 1438–1458. DOI: <https://doi.org/10.1029/2018MS001373>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2018MS001373>. URL: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001373>.

- Petersen, Mark R., Douglas W. Jacobsen, et al. (2015). "Evaluation of the arbitrary Lagrangian–Eulerian vertical coordinate method in the MPAS-Ocean model". In: *Ocean Modelling* 86, pp. 93–113. ISSN: 1463-5003. DOI: <https://doi.org/10.1016/j.ocemod.2014.12.004>. URL: <https://www.sciencedirect.com/science/article/pii/S1463500314001796>.
- Pettenuzzo, D., W.G. Large, and N. Pinardi (2010). "On the corrections of ERA-40 surface flux products consistent with the Mediterranean heat and water budgets and the connection between basin surface total heat flux and NAO". In: *Journal of Geophysical Research: Oceans* 115.6.
- Pinardi, N., I. Allen, et al. (2003). "The Mediterranean ocean forecasting system: First phase of implementation (1998-2001)". In: *Annales Geophysicae* 21.1 PART I, pp. 3–20.
- Pinardi, N., A. Rosati, and R.C. Pacanowski (1995). "The sea surface pressure formulation of rigid lid models. Implications for altimetric data assimilation studies". In: *Journal of Marine Systems* 6.1-2, pp. 109–119.
- Pinardi, N., M. Zavatarelli, et al. (2015). "Mediterranean Sea large-scale low-frequency ocean variability and water mass formation rates from 1987 to 2007: A retrospective analysis". In: *Progress in Oceanography* 132, pp. 318–332.
- Pisano, A. et al. (2016). "The new Mediterranean optimally interpolated pathfinder AVHRR SST Dataset (1982–2012)". In: *Remote Sensing of Environment* 176, pp. 107–116. ISSN: 0034-4257. DOI: <https://doi.org/10.1016/j.rse.2016.01.019>. URL: <http://www.sciencedirect.com/science/article/pii/S0034425716300190>.
- Provini, A., G. Crosa, and R. Marchetti (1992). "Nutrient export from the Po and Adige river basins over the last 20 years". In: *Science of the Total Environment* SUPPL. Pp. 291–313.
- Raicich, F. (1996). "On the fresh water balance of the Adriatic Sea". In: *Journal of Marine Systems* 9.3-4, pp. 305–319.
- Ringler, T. et al. (2013). "A multi-resolution approach to global ocean modeling". In: *Ocean Modelling* 69, pp. 211–232.
- Robinson, A.R. et al. (2001). *Mediterranean Sea Circulation*, pp. 710–725.
- Rodi, Wolfgang (1987). "Examples of calculation methods for flow and mixing in stratified fluids". In: *Journal of Geophysical Research: Oceans* 92.C5, pp. 5305–5328. DOI: 10.1029/JC092iC05p05305. eprint: <https://agupubs.onlinelibrary>.

- wiley.com/doi/pdf/10.1029/JC092iC05p05305. URL: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JC092iC05p05305>.
- Roland, Aron (Jan. 2008). "Development of WWM II: Spectral wave modeling on unstructured meshes". PhD thesis.
- Saad, Youcef (1994). *SPARSKIT: a basic tool kit for sparse matrix computations - Version 2*.
- Sannino, G., A. Sözer, and E. Özsoy (2017). "A high-resolution modelling study of the Turkish Straits System". In: *Ocean Dynamics* 67.3-4, pp. 397–432.
- Schiller, Rafael and Villy Kourafalou (Dec. 2010). "Modeling river plume dynamics with the HYbrid Coordinate Ocean Model". In: *Ocean Modelling* 33, pp. 101–117. DOI: 10.1016/j.ocemod.2009.12.005.
- Schumann, U. and T. Gerz (Jan. 1995). "Turbulent Mixing in Stably Stratified Shear Flows". In: *Journal of Applied Meteorology* 34.1, pp. 33–48. ISSN: 0894-8763. DOI: 10.1175/1520-0450-34.1.33. eprint: [https://journals.ametsoc.org/jamc/article-pdf/34/1/33/3896657/1520-0450-34\\_1\\_33.pdf](https://journals.ametsoc.org/jamc/article-pdf/34/1/33/3896657/1520-0450-34_1_33.pdf). URL: <https://doi.org/10.1175/1520-0450-34.1.33>.
- Simpson, John et al. (Jan. 1993). "Periodic stratification in the Rhine ROFI in the North Sea". In: *Oceanologica Acta* 16, pp. 23–32.
- Sivasankaran Rajamanickam, Erik G. Boman (2012). *An evaluation of the zoltan parallel graph and hypergraph partitioners*. Technical report. Sandia National Laboratories (SNL-NM), Albuquerque, NM (United States).
- Skliris, Nikolaos, Sarantis Sofianos, and A. Lascaratos (Mar. 2007). "Hydrological changes in the Mediterranean Sea in relation to changes in the freshwater budget: A numerical modelling study". In: *Journal of Marine Systems* 65, pp. 400–416. DOI: 10.1016/j.jmarsys.2006.01.015.
- Smagorinsky, J. (1963). "GENERAL CIRCULATION EXPERIMENTS WITH THE PRIMITIVE EQUATIONS". In: *Monthly Weather Review* 91.3, pp. 99–164. DOI: 10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2.
- Soto-Navarro, J. et al. (2010). "Estimation of the Atlantic inflow through the Strait of Gibraltar from climatological and in situ data". In: *Journal of Geophysical Research: Oceans* 115.10.
- Stanev, E.V. (2005). "Understanding Black Sea dynamics: An overview of recent numerical modeling". In: *Oceanography* 18.SPL.ISS.2, pp. 56–75.

- Stanev, E.V., M.J. Bowman, et al. (2003). "Control of Black Sea intermediate water mass formation by dynamics and topography: Comparison of numerical simulations, surveys and satellite data". In: *Journal of Marine Research* 61.1, pp. 59–99.
- Stanev, E.V., S. Grashorn, and Y.J. Zhang (2017). "Cascading ocean basins: numerical simulations of the circulation and interbasin exchange in the Azov-Black-Marmara-Mediterranean Seas system". In: *Ocean Dynamics* 67.8, pp. 1003–1025.
- Stanev, Emil and Joanna Staneva (Feb. 2000). "The impact of the baroclinic eddies and basin oscillations on the transitions between different quasi-stable states of the Black Sea circulation". In: *Journal of Marine Systems* 24, pp. 3–26. DOI: 10.1016/S0924-7963(99)00076-7.
- Staneva, J. V. and E. V. Stanev (1997). "Cold Intermediate Water Formation in the Black Sea. Analysis on Numerical Model Simulations". In: *Sensitivity to Change: Black Sea, Baltic Sea and North Sea*. Ed. by Emin Özsoy and Alexander Mikaelyan. Dordrecht: Springer Netherlands, pp. 375–393. ISBN: 978-94-011-5758-2. DOI: 10.1007/978-94-011-5758-2\_29. URL: [https://doi.org/10.1007/978-94-011-5758-2\\_29](https://doi.org/10.1007/978-94-011-5758-2_29).
- Staneva, Joanna et al. (Nov. 2001). "Rim Current and coastal eddy mechanisms in an eddy-resolving Black Sea general circulation model". In: *Journal of Marine Systems* 31, pp. 137–157. DOI: 10.1016/S0924-7963(01)00050-1.
- Tolman, H.L. (2008). "A mosaic approach to wind wave modeling". In: *Ocean Modelling* 25.1-2, pp. 35–47.
- Tonani, M., N. Pinardi, S. Dobricic, et al. (2008). "A high-resolution free-surface model of the Mediterranean Sea". In: *Ocean Science* 4.1, pp. 1–14.
- Tonani, M., N. Pinardi, C. Fratianni, et al. (2009). "Mediterranean Forecasting System: Forecast and analysis assessment through skill scores". In: *Ocean Science* 5.4, pp. 649–660.
- Umgiesser, G., D.M. Canu, et al. (2004). "A finite element model for the Venice Lagoon. Development, set up, calibration and validation". In: *Journal of Marine Systems* 51.1-4 SPEC. ISS. Pp. 123–145.
- Umgiesser, G., C. Ferrarin, et al. (2014). "Comparative hydrodynamics of 10 Mediterranean lagoons by means of numerical modeling". In: *Journal of Geophysical Research: Oceans* 119.4, pp. 2212–2226.

- Umgiesser, G., D. Melaku Canu, et al. (2003). "A finite element ecological model: A first application to the Venice Lagoon". In: *Environmental Modelling and Software* 18.2, pp. 131–145.
- Ünlülata, Ü. et al. (1990a). "On the Physical Oceanography of the Turkish Straits". In: *The Physical Oceanography of Sea Straits*. Ed. by L. J. Pratt. Dordrecht: Springer Netherlands, pp. 25–60. ISBN: 978-94-009-0677-8. DOI: 10.1007/978-94-009-0677-8\_2. URL: [https://doi.org/10.1007/978-94-009-0677-8\\_2](https://doi.org/10.1007/978-94-009-0677-8_2).
- (1990b). "On the Physical Oceanography of the Turkish Straits". In: *The Physical Oceanography of Sea Straits*. Ed. by L. J. Pratt. Dordrecht: Springer Netherlands, pp. 25–60. DOI: 10.1007/978-94-009-0677-8\_2. URL: [https://doi.org/10.1007/978-94-009-0677-8\\_2](https://doi.org/10.1007/978-94-009-0677-8_2).
- Verri, G. et al. (2018). "River runoff influences on the Central Mediterranean overturning circulation". In: *Climate Dynamics* 50.5-6, pp. 1675–1703.
- Vörösmarty, C.J., B. Fekete, and B.A. Tucker (1998). *River Discharge Database, Version 1.1 (RivDIS v1.0 supplement)*. Available through the Institute for the Study of Earth, Oceans, and Space / University of New Hampshire, Durham NH (USA).
- Walshaw, Chris and Mark Cross (2007). "JOSTLE: parallel multilevel graph-partitioning software—an overview". In: *Mesh partitioning techniques and domain decomposition techniques*, pp. 27–58.
- Wang, J. and M. Ikeda (1995). "Stability Analysis Of Finite Difference Schemes For Inertial Oscillations In Ocean General Circulation Models". In: 10, p. 9. DOI: 10.2495/CE950031.
- Wessel, Paul and Walter Smith (Apr. 1996). "A global, self-consistent, hierarchical, high-resolution shoreline database". In: *Journal of Geophysical Research* 101, pp. 8741–8743. DOI: 10.1029/96JB00104.
- YS, Husrevoglu (1993). "Modelling of the Dardanelles Strait lower-layer flow into the Marmara Sea". MA thesis. Institute of Marine Sciences, METU, Erdemli, Mersin, Turkey.
- Zavatarelli, M. and G.L. Mellor (1995). "A numerical study of the mediterranean sea circulation". In: *Journal of Physical Oceanography* 25.6, pp. 1384–1414.
- Zavatarelli, M. and N. Pinardi (2003). "The Adriatic Sea modelling system: a nested approach". In: *Annales Geophysicae* 21.1, pp. 345–364. DOI: 10.5194/angeo-21-345-2003. URL: <https://angeo.copernicus.org/articles/21/345/2003/>.

- Zavialov, I. et al. (2020). "Water exchange between the Sea of Azov and the Black Sea through the Kerch Strait". In: *Ocean Science* 16.1, pp. 15–30. DOI: 10.5194/os-16-15-2020. URL: <https://os.copernicus.org/articles/16/15/2020/>.
- Zhang, Y.J. et al. (2016). "Seamless cross-scale modeling with SCHISM". In: *Ocean Modelling* 102, pp. 64–81.
- Zhang, Yu et al. (2016). "Seasonal and interannual variability of the Arctic sea ice: A comparison between AO-FVCOM and observations". In: *Journal of Geophysical Research: Oceans* 121.11, pp. 8320–8350. DOI: <https://doi.org/10.1002/2016JC011841>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2016JC011841>. URL: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JC011841>.

# Acknowledgements

Several people collaborated to this work. The bulk of the SHYFEM parallelization has been conducted by CMCC ASC Division and, in particular, by Giorgio Micaletto, that I have to thank for the constant support.

The rest of the work has been carried at CMCC OPA Division where my advisors, Dr. Ivan Federico e Dr. Giorgia Verri have always supervised me and supported me, either from a purely scientific point of view or, not less importantly, motivational.

I acknowledge the importance of the overarching supervision of Prof. Nadia Pinardi with her scientific advice and her constant willigness to impart the knowledge.

I had benefit from my discussions with Dr. Mehmet Ilicak, especially in my stay at ITU, giving me deeper insights on the model numerics. He also has implemented part of the diagnostics output in SHYFEM.

I enjoyed a lot the company that I had in these 3 years at OPA Division in Lecce, I found all my colleagues kind and friendly. I include also the CMCC communication office in Lecce, no worries. I did not forget about you...

I want to thank my family. Unconditional and constant support not just throughout the last 3 years, throughout my life. Pure and simple.

There are my friends also. People that ask you "How is it going?". It looks so simple but it is so important. People that want to share their time with you. It is so important too. There's a friend, that could not give his support, but I'm sure he would have done if he had had the chance, if he had remained among us.