Comparison of in situ microstructure measurements to different turbulence closure schemes in a 3-D numerical ocean circulation model

Andrea Costa[a,b], Andrea M. Doglioli[a], Patrick Marsaleix[c], Anne A. Petrenko[a]

[a]*Aix-Marseille Université, Université de Toulon, CNRS, IRD, Mediterranean Institute of Oceanography (MIO), Marseille, France Personnels UTLN*
[b]*current address: IBS Center for Climate Physics (ICCP), Pusan National University, Busan, Republic of Korea*
[c]*Laboratoire d'Aérologie UMR 5560, Université Paul Sabatier, CNRS, 31400 Toulouse*

Abstract

In situ measurements of kinetic energy dissipation rate ε and estimates of eddy viscosity K_Z from the Gulf of Lion (NW Mediterranean Sea) are used to assess the ability of $k - \varepsilon$ and $k - \ell$ closure schemes to predict microscale turbulence in a 3-D numerical ocean circulation model. Two different surface boundary conditions are considered in order to investigate their influence on each closure schemes' performance. The effect of two types of stability functions and optical schemes on the $k - \varepsilon$ scheme is also explored. Overall, the 3-D model predictions are much closer to the in situ data in the surface mixed layer as opposed to below it. Above the mixed layer depth, we identify one model's configuration that outperforms all the other ones. Such a configuration employs a $k - \varepsilon$ scheme with Canuto A stability functions, surface boundary conditions parameterizing wave breaking and an appropriate photosynthetically available radiation attenuation length. Below the mixed layer depth, reliability is limited by the model's resolution and the specification of a hard threshold on the minimum turbulent kinetic energy.

Keywords: Scalar vertical diffusivity, Kinetic energy dissipation rate, Turbulence closure schemes, Stability functions, Surface boundary conditions, Microstructure, Symphonie numerical model

Email address: `andrea.costa@pusan.ac.kr` (Andrea Costa)

1. Introduction

Turbulence is an essential mechanism for the transport of energy, salinity, and suspended and dissolved matter. Turbulent fluxes of such quantities are the result of correlated, small-scale fluctuations of the velocity field and of the transported quantity itself. The prevalent turbulence production mechanisms in coastal ocean are: mean shear, unstable stratification, Langmuir circulation (Farmer and Li, 1995) and breaking surface waves (Agrawal et al., 1992). For coastal ocean, mean shear is mainly generated by the action of winds and tides, but also by surface waves and baroclinic flows (e.g., Thorpe, 2005), including nonlinear internal waves (Toole and Schmitt, 1987). Unstable stratification results from surface processes such as surface cooling, evaporation or differential advection (e.g., Kantha and Clayson, 2000). Destruction of turbulence occurs by transformation into potential energy during stable stratification or viscous dissipation into heat (e.g., Kantha and Clayson, 2000). The complexity of these processes by themselves and of their interactions requires numerical models to cover a wide range of spatio-temporal scales and Reynolds number (e.g, Burchard et al., 2008). This is especially true in the upper ocean where all the above phenomena concur together to generate turbulence.

Upper ocean connects –through various turbulent mechanisms– the surface forcing from the atmosphere with the quiescent deeper ocean where heat and fresh water are sequestrated and released on longer time and global scales (Ferrari and Wunsh, 2009). Also, upper ocean turbulence plays an important role in biological phenomena by, for example, determining phytoplankton growth rate (Thomas and Gibson, 1990), influencing primary production (Flierl and Davis, 1993) and the onset of blooms (Taylor and Ferrari, 2011). The complexity of modelling such mechanism within ocean circulation numerical models gave rise to several approaches. In particular, many turbulence closure schemes have been proposed. The ones most frequently found in the ocean modelling community's literature are the $k - k\ell$ by Mellor and Yamada, 1982; the $k - \varepsilon$ by Rodi, 1987; the $k - k\omega$ by Wilcox, 1988; the $k - \ell$ by Gaspar et al., 1990 and the KPP by Large et al., 1994. Following recent numerical modelling literature (Ilicak et al., 2008; Reffray et al., 2015), in the present study, we consider the $k - \varepsilon$ and $k - \ell$ second moments closure (SMC) schemes. Note that other kinds of closure schemes such as the KPP (Large

et al., 1994) are not considered here being not as well suited as the other two schemes for a comparison with in situ data of kinetic energy dissipation rate ε.
Additional complexity is added to the modelling by the interplay of the SMC and the choice of boundary conditions. The choice of surface and bottom boundary conditions can also profit from a vast literature (e.g., Craig and Banner, 1994; Stacey and Pond, 1997; Estournel et al., 2001; Warner et al., 2005), aiming at modelling different forcing mechanisms. Furthermore, different stability functions can be chosen in order to include the effect of the parametrized non-local moments and pressure strain correlations in the dynamical equations (e.g., Galperin et al., 1988; Kantha and Clayson, 1994; Canuto et al., 2001). The choice of the optical scheme is particularly important considering the high number of studies coupling Symphonie to biochemical models as it can influences turbulent fluxes and nutrient availability. Thus, the in situ validation of the closure schemes, boundary conditions, stability functions, optical scheme and their interplay is fundamental for assessing the reliability of numerical models (Warner et al., 2005; Peters and Baumert, 2007; Arneborg et al., 2007; Ilicak et al., 2008).

The current study presents the comparison of kinetic energy dissipation rate ε measurements and vertical eddy viscosity K_Z estimates issued from a Self Contained Microstructure Profiler (SCAMP) with the predictions of a 3-D numerical ocean circulation model (Symphonie; Marsaleix et al., 2008) obtained with different model's setup. The aim is to gain some insights on which scheme and/or boundary conditions permit to have the representation of turbulence activity closer to the observations.
Microstructure measurements with the SCAMP profiler have already been used for turbulence estimates in lakes and ocean (e.g., Ruddick et al., 2000; Sharples and Moore, 2001; Burchard et al., 2002; Anis and Singhal, 2002; Sharples et al., 2003; Peters et al., 2009; Steinbuck et al., 2010; Steinbuck et al., 2011; Cuypers et al., 2012; Jurado et al., 2012; Bouffard and Boegman, 2013). The dataset we exploit is described in Section 2. It consists of measurements taken in a coastal environment in the Gulf of Lion (GoL).

The GoL is located in the northwestern Mediterranean Sea and is characterized by a large continental margin (Figure 1) and complex hydrodynamics (Millot, 1990). Its circulation is strongly influenced by the southwestward along-slope Northern Current. This density current flows in a cyclonic way

and constitutes a barrier between the coastal waters of the continental shelf
from the open northwestern Mediterranean Sea (Alberola and Millot, 1995;
Sammari et al., 1995; Petrenko, 2003). Cross-shore exchanges between the
GoL and offshore waters are regulated by wind induced dynamics (Estournel et al., 2003; Hauser et al., 2003; Petrenko et al., 2017) and by processes
associated with the Northern Current, such as intrusions into the continental shelf and barotropic and baroclinic instabilities (Conan and Millot, 1992;
Flexas et al., 1997; Petrenko et al., 2005; Barrier et al., 2016). The Gulf of
Lion is a suitable case study because of the high number of physical (Qiu
et al., 2010; Hu et al., 2011), sediment dispersion (Bourrin et al., 2011) and
biochemical (Pinazo et al., 2001; Herrmann et al., 2014) numerical studies
carried out there.

Symphonie has already been validated on a variety of different aspects
like current modelling and eddy generation (Rubio et al., 2009; Hu et al.,
2011; Kersalé et al., 2013), river plume dynamics (Reffray et al., 2004; Gatti
et al., 2006) and dense water formation (Dufau-Julliand et al., 2004; Estournel et al., 2016). But a study of the different SMC that the user can
implement in the Symphonie code has not yet been done. In particular, the
modeling of the near-surface physical and biogeochemical processes is sensitive to the choice of SMC and the computed K_Z values (Fraysse et al., 2014).
In general, we can regard all modeled large-scale circulation features in an
integrated fashion as they result from successive calculation steps and approximations. Hence, a major difficulty in validating numerical models –beside
the high number of variables at play– is the possible compensation of different errors between each other. This fact makes difficult to attribute a specific
amount of the total error on a certain quantity to a specific step in its calculation, in the present case the turbulence scheme. Here, our goal is to assess
the model predictions focusing on turbulence modelling in the most realistic
configuration we can achieve: 3-D dynamics with realistic forcing. Indeed,
mixing has a primary role in influencing the large-scale circulation motion
(Rhines, 1988; Ferrari, 2014). Nevertheless, the 3-D nature of the model
brings in play an augmented number of numerical issues, among which spurious numerical diffusion (Marsaleix et al., 2008; Marchesiello et al., 2009; Hu
et al., 2009) that are usually neglected in similar, but 1-D, studies (Gaspar
et al., 1990; Burchard et al., 2002; Reffray et al., 2015).

The manuscript is organized as follows. In Section 2 we describe the prop-

4

erties of the numerical model, the microstructure in situ data and how we carry out the comparison between them and the numerical data. In Section 3 we report the results of our analysis and we discuss them in Section 4. In Section 5 we summarize the conclusions of our study.

2. Materials and Methods

2.1. Numerical modelling

The numerical model Symphonie is a 3-D primitive equations, free surface, sigma coordinate ocean model, based on Boussinesq and hydrostatic approximations (Marsaleix et al., 2008; Marsaleix et al., 2009; Marsaleix et al., 2012). Components of current, temperature and salinity are computed on a C-grid (Arakawa and Lamb, 1977) using a classic finite difference method detailed in Marsaleix et al. (2006) and Marsaleix et al. (2008). This model has been extensively used in studies of the Mediterranean Sea, mostly at the scale of the continental shelves (Ulses, 2005; Estournel et al., 2003; Estournel et al., 2005), generally comparing satisfactorily with available in-situ observations of classical hydrological quantities. Symphonie has also been coupled to biochemical models for studies that demonstrated the impact of the turbulence level on determining the vertical flux of nutritive salts, the nutricline depth and –as a consequence– the results given by the biochemical models (Ulses et al., 2016; Herrmann et al., 2013). However, a study of the consequences of choosing a certain model's implementation of the Symphonie code has not yet been done. To fill this gap, we compare the model predictions of ε and K_Z with the values measured with the SCAMP profiler.

The model domain we use, shown in Figure 1, is that of Estournel et al. (2016). Note that all the measurements sites are far from the open boundaries. The horizontal resolution of the model grid is 1/110 degree (about 1 km). All the numerical experiments we perform cover the whole period in which in situ data are available: from 1 July 2010 to 13 March 2014, plus ten weeks of spin up.

In the vertical the model exploits a generalized sigma coordinate with 50 levels. Surface fluxes are computed using the bulk formulae by (Large and Yeager, 2004) and the 3-hours ECMWF by Estournel et al. (2016). The boundary condition for ε is deduced with a length scale reasoning from the value of the gradient Richardson number (i.e., the ratio between buoyant production of turbulence and shear production of turbulence); see Estournel

5

and Guedalia (1987), Michaud et al. (2012) and Appendix A for more details. In order to simulate the limiting effect of stable stratification, following Galperin et al. (1988), the minimum of ε is linked to the minimum turbulent kinetic energy value k_{min} through:

$$\varepsilon_{min} = k_{min}\frac{0.55^3 N}{0.53\sqrt{2}} \tag{1}$$

As a default in Symphonie, $k_{min} = 10^{-8}\ kg\ m^2/s^2$; this implies $\varepsilon_{min} = 10^{-12}\ m^2/s^3$. The choice of this threshold follows from underestimating the more standard value of $k_{min} = 10^{-6}\ kg\ m^2/s^2$ (Gaspar et al. 1990; Burchard et al. 2002), that is based on the estimate of the internal wave activity. Similar low-frequency buoyancy conditions are maintained for all numerical experiments using a nudging procedure on temperature and salinity toward the corresponding MERCATOR fields (product PSY2V4R4). The nudging time scale is 30 days, enabling the free development of higher frequencies (including those of the turbulence closure scheme), and, at the same time, ensuring that the different turbulence schemes are tested in similar general conditions of stratification.

We choose the $k - \varepsilon$ (Burchard and Bolding, 2001) and $k - \ell$ (Gaspar et al., 1990) closure schemes because, other than being the more exploited by Symphonie's users, they are also widely used in the wider scientific community. Reffray et al. (2015) showed that in a 1-D case the $k - \varepsilon$ scheme gives more reliable mixing estimate with respect to other schemes widely used in the literature: $k - kl$, $k - \omega$ and $k - \ell$. The $k - \ell$ scheme in Reffray et al. (2015) is based on Gaspar et al. (1990) but simplified for 1-D applications. We want to test if the original scheme by Gaspar et al. (1990) performs better than $k - \varepsilon$ in a 3-D case.

Moreover, our questioning the numerical results' sensitivity on the value of k_{min} follows from the study by Gaspar et al. (1990). Therein the authors encouraged (but not implemented) the use of a parametrization k_{min} rather than fixing a hard value. Herein we test if a good result can be achieved in a simpler way by specifying a different value of k_{min}.

A study by Burchard and Bolding (2001) showed that, in a 1-D study of temperature and mixed layer depth data of the well-known dataset OWS Papa (northern Pacific), the $k - \varepsilon$ closure scheme performs better when employing the stability functions proposed by Canuto et al. (2001) -commonly called

Canuto A- rather than the ones by Kantha and Clayson (1994), Rodi (1980) and Hossain (1980). On the other hand, Ilicak et al. (2008) showed that –in a 3-D study of the Red Sea outflow– the stability functions of both Canuto et al. (2001) and Kantha and Clayson (1994) perform similarly. We want to further investigate the different performance of these two stability functions in a 3-D case. Note also that this study –as opposed to Burchard and Bolding (2001) and Ilicak et al. (2008)– has a strong focus on microstructure measurements and not only on more standard quantities like mixed layer depth, temperature and salinity.

In the literature there are different formulations for K_Z. In particular, it can include the molecular diffusivity D_T (so that we always have $K_Z > D_T$; e.g., Burchard and Bolding 2001) or not (e.g., Han, 2014). Here we want to clarify the differences (if any) between the two approaches.

Not having the necessary computational power to explore all the possible combinations of these factors in a 3-D model (as done for example in a similar study in the 1-D case by Reffray et al. (2015), and in a 3-D case by Ilicak et al. (2008) but on a shorter time span), we restrict the study to a subset of combinations.

In particular, nine different numerical experiments employing different combinations of turbulent closure schemes, boundary conditions, stability functions, values of the minimum of turbulent kinetic energy and optical schemes are analyzed here (see Table 1 for a concise summary).

Five of these numerical experiments employ a closure scheme based on a $k-\varepsilon$ approach (Burchard and Bolding, 2001) —hereafter marked by the prefix **KE**. The other four numerical experiments employ a $k-\ell$ scheme based on Gaspar et al. (1990) —hereafter marked by the prefix **KL**. Details of these nine numerical experiments can be found in Appendix B. We test the effect of two possible surface boundary conditions. The first one (marked by a suffix **set**) supposes equilibrium between the production and dissipation terms in the dynamic equation for k (see Equation B.1). The second one (marked by a suffix **flu**), takes into account the effect of breaking waves of all scales, as suggested by Craig and Banner (1994); see Equation A.1 and Appendix A for details. The numerical experiments exploiting the **KE** scheme and using the stability by Canuto et al. (2001) –instead of the ones by Kantha and Clayson (1994)– are marked by the suffix **CAN**. Simulations with a higher minimum TKE –10^{-7} $kg\ m/s^2$ instead of 10^{-8} $kg\,m/s^2$– are marked with a suffix **MINk**.

All the numerical experiments with a higher threshold on k are such that $K_Z > D_T$. Therefore, in the other numerical experiments the total diffusivity could be smaller than the molecular one.

One numerical experiment (marked by the suffix **Opt**) investigates the effect of the attenuation length of the penetrative solar radiation Q_{sr}. Q_{sr} is parametrized as a two-band exponential scheme (Maraldi et al., 2013):

$$Q_{sr}(z) = Q_{sr}(0) \left[Re^{-z/l} + (1 - R)e^{-z/l_{PAR}} \right] \tag{2}$$

where the first right-hand term parameterizes the attenuation of red and near-infrared radiation (whose attenuation length is $l = 0.35$ cm); and the second right-hand term is the one of the visible and ultra-violet radiation; l_{PAR} is the photosynthetic available radiation diffuse attenuation length. $Q_{sr}(0)$ is the fraction of the available penetrative solar radiation at the surface assuming a constant albedo of 6.6% and R = 0.54 determines the fraction in each band of $Q_{sr}(0)$. Following Maraldi et al. (2013), the l_{PAR} default in Symphonie is set to 11 m. However, this value has to be considered as an annual climatological estimate of the photosynthetic available radiation. With this numerical experiment we test the effect of the seasonality of l_{PAR} by setting its value to 23 m coherently with the fact that most of the in situ measurements were acquired in September when we expect a low biological activity in the surface boundary layer of the GoL.

With this set of numerical experiments we can answer three principal questions: 1) which SMC between $k - \varepsilon$ and $k - \ell$ performs better with respect to our dataset? 2) what is the effect of the boundary conditions on the results of the numerical numerical experiments? and 3) what is the effect of the two stability functions?

2.2. SCAMP measurements

An in situ estimate of ε can be derived from high-resolution vertical profiles of temperature T. Batchelor (1959) derived the spectral shape of a conserved scalar field that is passively advected by an incompressible turbulent fluid with a molecular Prandtl number $Pr = \nu/D$ greater than 1 (seawater has $Pr = 7$ at 20°C), where ν and D are respectively the molecular viscosity and the molecular diffusivity of the scalar. In the present case the scalar is the temperature. Gibson and Schwarz (1963) derived the one-dimensional Batchelor spectrum $E(K)$ of temperature gradient as a function of the rate

8

of dissipation of temperature variance χ_T, the kinetic energy dissipation rate ε, the molecular diffusivity of temperature D_T and the circular wavenumber K:

$$E(K) = f(\chi_T, D_T, K_B) \tag{3}$$

where K_B is the inverse of the Batchelor length scale describing the length scales at which fluctuations in scalar concentration (temperature in this case) can still exist before being evened out by molecular diffusion. Therefore, (see Ruddick et al., 2000; Luketina and Imberger, 2000 and Steinbuck et al., 2009 for details) once measured the temperature vertical gradient at the millimeter scale and derived χ_T and K_B by fitting the Batchelor spectrum, a measure of ε follows from (Batchelor, 1959):

$$\varepsilon = \frac{\nu D_T^2}{K_B^4} \tag{4}$$

The temperature gradient profiles were measured with a SCAMP profiler. This instrument is equipped with a 100 Hz FP07 glass rod microthermistors (sensitivity of 0.001°C).
Our SCAMP was deployed in an upward configuration. After deployment, it sinks to a predetermined depth following an oblique trajectory and then rises up vertically at an approximately constant velocity $U = 10^{-1}\ m/s$. This type of trajectory permits the SCAMP to get away from the ship and be free from the influence of the ship's wake when rising up in an undisturbed water column. This allows to have reliable measurements of ε and K_Z near the sea surface (Anis, 2006).

The dataset (spanning the period 2010-2014) for this analysis consisted of 126 profiles of variable vertical extent —between 1 m below the surface and 100 m depth— collected in different sites in the Gulf of Lion (Figure 1). The profiles were collected during various oceanographic campaigns in the GoL conducted by the Mediterranean Institute of Oceanography (MIO - Marseille, France). Because the SCAMP measurements were opportunistically taken in cruises mainly not dedicated to turbulence measurements, repeated casts were often not possible. Data were collected mostly during summer and with meteorological conditions favorable to operations with a small boat, generally with wave heights less than half meter. Precipitation was always absent or negligible. Surface buoyancy flux was positive for 115

9

profiles indicating a gain of buoyancy by the ocean surface.

Temperature gradient spectra were computed from 128 points ($\approx$ 13 cm) windows without overlap. The choice of this segmentation resulted from a sensitivity analysis using different segmentation methods proposed in the literature (see Appendix C).

The vertical eddy viscosity coefficient K_Z can be derived on the basis of the turbulence intensity parameter $Re_b = \varepsilon/\nu N$, as follows. Re_b expresses the ratio of the destabilizing effect of turbulence to the stabilizing effect of stratification and viscosity. In different Re_b regimes, the mixing efficiency, expressing the portion of the energy produced by shear which is dissipated by viscosity, assumes different values and determines different vertical turbulent diffusivity of density K_ρ. Here we use a recent field-validated parametrization of K_ρ as function of Re_b proposed by Bouffard and Boegman (2013) based on a previous parametrization derived by Shih et al. (2005). At very low Re_b ($Re_b < 10^{2/3}/\sqrt{Pr} \approx 1.7$), the turbulent regime is regarded as diffusive and K_ρ is set equal to the temperature molecular diffusivity $D_T = 1.4 \times 10^{-7} m^2 s^{-1}$. At low Re_b ($10^{2/3}/\sqrt{Pr} < Re_b < (3 \ln \sqrt{Pr})^2 \approx 8.5$) turbulent mixing tends to be controlled by buoyancy effects with incomplete mixing favoring up-gradient fluxes reducing the mixing efficiency. Here K_ρ can be expressed as $K_\rho = 0.1 Pr^{-1/4} \nu Re_b^{3/2}$. For moderate Re_b ($(3 \ln \sqrt{Pr})^2 < Re_b < 400$) the mixing efficiency has the classical form derived by Osborn (1980): $K_\rho = 0.2\nu Re_b$. For high Re_b ($Re_b > 400$), turbulence is in an energetic regime, where $K_\rho = 4\nu\sqrt{Re_b}$. Therefore, for high Re_b the buoyancy flux $K_\rho N$ tends to vanish together with stratification: $K_\rho N = \sqrt{\nu\varepsilon} N \xrightarrow{N \to 0} 0$, as we expect in weakly stratified fluids (Osborn, 1980). As in the case of the Mediterranean Sea (e.g., Cuypers et al., 2012), when the density variations are dominated by those of temperature, the density vertical eddy diffusivity coefficient is assumed to be equal to the temperature vertical eddy diffusivity coefficient (Peters et al., 1988). Assuming then a turbulent Prandtl number $Pr_t = K_Z/K_\rho = 1$ (Hogg et al., 2001), it is possible to estimate the momentum eddy viscosity (or eddy viscosity) K_Z to be equal to K_ρ.

2.3. Comparison of numerical and in situ data

While the vertical resolution of the in situ profile of ε and K_Z is constant (13 cm), the vertical resolution of the model is variable and generally

coarser ($\approx 1 - 2\ m$) than the in situ one. In order to compare the numerical data to the measurements, the SCAMP data are grouped in windows centered on each sigma level. Then, the median of each window is calculated and compared to the numerical data at that sigma level. The choice of the median permits to give less importance to outliers and reduce the error due to the fact that the profiles were mainly single casts (Lozovatsky et al., 2005).

For both the in situ and numerical data, we define the surface mixed layer (MLD) as the depth at which the temperature is smaller than the surface value by 0.5°C (Anis, 2006; Jurado et al., 2012). Then we separate the data in the MLD from the ones below it. The data in the bottom layer are not considered for this analysis because of the low number of in situ profiles near the sea bottom.

Following Burchard et al. (2002), we define the standard deviation between the logarithmic numerical and in situ values of ε as:

$$\sigma_\varepsilon = \sqrt{\frac{1}{M} \sum_{i=1}^{M} \left[\log \left(\frac{\varepsilon_i^{mod}}{m^2 s^{-3}} \right) - \log \left(\frac{\varepsilon_i^{obs}}{m^2 s^{-3}} \right) \right]^2} \qquad (5)$$

where M is the number of points in a given profile, and ε^{mod} and ε^{obs} are the numerical and observed values respectively. Similar definitions hold for the decadal standard deviation of N (σ_N) and K_Z (σ_{K_Z}). To calculate these quantities, no division between data above and below the MLD is made because the number of in situ and numerical values in the MLD can differ for a given profile. Hereinafter, the mean values of σ_ε, σ_N and σ_{K_Z} on all the profiles are denoted σ_ε, σ_N and σ_{K_Z} for ease of reading, instead of $< \sigma_\varepsilon >$, $< \sigma_N >$ and $< \sigma_{K_Z} >$.

We also compare the probability density functions (PDFs) of the in situ and numerical values of kinetic energy dissipation rate ε, Brunt-Väisälä frequency N and eddy viscosity K_Z. In particular, the PDFs are calculated with a kernel density estimation (Bowman and Azzalini, 1997) on the base of the frequency distributions of values from all the in situ and numerical profiles.
In order to compare a distribution of numerical data $g(x)$ with the distribution of the in situ data $f(x)$, we compute the squared difference between the

empirical cumulative distribution functions (F_n and G_n) of the two distributions:

$$\Delta^2(S) = \int_0^\infty \left(F_n(x) - G_n(x + S) \right)^2 dx \tag{6}$$

We define the shift S between two distributions as the shift of $G_n(x)$ that permits to minimize Δ^2. A subscript will tell to which distribution the values of Δ^2 and S refer to: Δ_ε^2 and S_ε for the PDF of ε; Δ_N^2 and S_N for the PDF of N; $\Delta_{K_Z}^2$ and S_{K_Z} for the PDF of K_Z. An estimate of the error on Δ^2 values is obtained by re-sampling the in situ empirical distribution function (Nerini and Ghattas, 2007) in order to further account for uncertainty in the measurements.

3. Results

To quantify the agreement between the numerical and the in situ data we use two methods. One takes into account the deviation of the numerical data from the in situ profiles over the whole water column and expresses it by the decadal standard deviations (see Section 3.1 and Table 2). The other approach looks at the shape of the probability density functions of the numerical values of ε, N and K_Z and expresses it by the squared differences and shift values above and below the MLD (see Section 3.2 and Tables 3 and 4). We then also look at the in situ and numerical data in the (ε, N, K_Z) space (see Section 3.3).

3.1. In situ and numerical median profiles

Median profiles of dissipation ε for the numerical outputs and observations are shown in Figure 2 vs. nondimensional depth z/MLD. Note that, the average value of the in situ MLD value is 27.3 m. As reported in Table 2, all the numerical MLDs are systematically lower by $\approx 6 - 15\%$ compared to in situ MLDs. Overall, the KE numerical experiments predict a slightly deeper MLD, closer to the observed values.

The thick lines in Figure 2 represent the median dissipation profiles calculated using all profiles, where the shades represent the 95% bootstrap confidence interval. At the surface, all the numerical experiments agree with

12

the in situ data to within one order of magnitude with the in situ data. At this depth, KEflu is the numerical experiment with values the closest to the observations. Also note that, at the surface, the equilibrium boundary condition generates higher dissipation values than the flux boundary condition. As we descend the water column, beyond $z/MLD = 1$, the in situ data tend to become more variable due to the lower number of profiles reaching depths greater than the MLD. On the contrary, the numerical data tend to become much less variable. Moreover, there is an evident difference between the SMCs: when the lower threshold on k is applied, the KE numerical experiments always have lower levels of turbulence with respect to the KL numerical experiments. In this case, KEs' median values of ε are one order of magnitude lower than the KLs' ones. When a higher threshold on the kinetic energy k is applied, the difference is slightly reduced.

The in situ data within the mixed layer are characterized by higher mixing close to the surface and lower mixing near the MLD. This is reflected by a change of slope in the ε profiles at $z/MLD = 0.5$. KEflu, KLflu, KLset-MINk and KLfluMINk are the numerical experiments that better reflect this behavior. Note also that the in situ data show a clear increment of ε at $z/MLD = 1.1$ that is absent in the numerical data.

We do not show values for $z/MLD > 2$ because of the large bootstrap error bars due to the lower number of profiles in that depth range.

Looking at the column as a whole, the numerical experiments with a higher minimum value of kinetic energy (KEfluCANMINk, KEfluCANMINkOpt KLsetMINk and KLfluMINk) have the lower average error with respect to the measurements (see values of σ_ε in Table 2) because of the better agreement with in situ data below the MLD.

Median profiles of stratification N for the numerical outputs and observations are shown in Figure 3 vs. nondimensional depth z/MLD. Near the surface, the in situ data are considerably more variable than the numerical data. However, the numerical data are of the same order of magnitude as the measurements. As we go below the MLD, the median value of the in situ data increases in the depth range $1 < z/MLD < 1.8$. Overall, all the numerical experiments do not show such a marked increase in magnitude and are rather similar one to another (as we expect because all of them are nudged to the MERCATOR fields). However, KEfluCANMINkOpt seems to be the numerical experiment that better reproduce the in situ data's trend right below the MLD.

13

Looking at the column as a whole, KLsetMINk has the lowest average error with respect to the measurements (see values of σ_N in Table 2).

The measurements of ε and N make it possible to calculate the turbulence intensity parameter Re_b. The probability density functions of the observed Re_b are reported in Figure S1 in the Extra Materials. In particular, in our dataset, 30% of the in situ Re_b values fell in the diffusive regime, 15% in the buoyancy-controlled regime, 29% in the transitional regime and 26% in the energetic regime. If only the data above the MLD are considered these numbers respectively become 17%, 10%, 31% and 42%; or 36%, 16%, 28% and 20%, if only the data below the MLD are considered.

Median profiles of eddy diffusivity K_Z for the numerical outputs and observations are shown in Figure 4 vs. nondimensional depth z/MLD. In the surface layer the numerical values are often more variable than the in situ ones. The opposite is observed below the MLD.
Among the four basic model's configurations (KEset, KEflu, KLset and KLflu), near the surface, KEflu shows the best agreement with the observations among the simulations with the lower threshold on k. KEflu is not as good as KEset, but it is significantly improved by using the Canuto A stability functions. KLset and KLflu perform similarly but have an error in opposite directions. Near the surface, KLset is higher than KLflu. When a higher threshold on k is applied, KLsetMINk is closer to the observations than KLfluMINk.
Below the MLD, the numerical experiments appear to be grouped on the base of the threshold on k: no differences between different SMC and boundary conditions appear obvious. Noticeably, below the MLD, the numerical experiments with a higher threshold on the kinetic energy k are more compatible with the in situ data than the other numerical experiments. The median value of K_Z below the MLD appears to have a seasonal behavior. In particular, it is equal to $5.90 \times 10^{-6} m^2/s^2$ in spring, 3.02×10^{-6} in summer, $3.27 \times 10^{-6} m^2/s^2$ in autumn and $5.57 \times 10^{-7} m^2/s^2$ in winter.
Overall, the numerical experiments that have the lower average error with respect to the measurements are the four numerical experiments with a high threshold on k, as the values of σ_{K_Z} in Table 2 indicate.

3.2. In situ and numerical probability density functions

We now focus on the comparison of the in situ and numerical PDFs of ε, N and K_Z values.

465

Above the MLD, the in situ dissipation rate data (in black in Figures 5a-b), show a clear bimodal behavior with one peak at $\varepsilon \approx 5 \times 10^{-9} m^2/s^3$ and one at $\varepsilon \approx 5 \times 10^{-6} m^2/s^3$. The numerical distributions show this feature in a less marked way and are generally biased low with respect to the in situ data's PDF. As values of S_ε and Δ_ε^2 in Table 3 show, above the MLD, the distribution more aligned to the in situ data one are KEfluMINk and KEflu-MINkOpt, while the distribution with the shape more resembling the in situ data one is KEfluMINkOpt.

Below the MLD (Figures 5c-d), all the numerical PDFs have mean values similar to the ones of the corresponding PDF in the mixed layer, but they all are much more narrow than the in situ data. KL numerical experiments have higher mean values than KE ones and the numerical experiments with a threshold on k have higher mean values than the corresponding numerical experiments without threshold. As we see in Table 4, below the MLD, KLsetMINk and KLfluMINk have the same shift with respect to the in situ distribution, while KEfluCANMINkOpt and KLfluCANMINkOpt have the lower value of Δ_ε^2. But we note that the differences between the values of Δ_ε^2 below the MLD are due to small differences in the low right-end tails that are unlikely to be significant.

485

The probability density functions of the in situ stratification data (in black) and of the numerical data are shown in Figure 6. Both above and below the MLD, there is no drastic difference between the numerical distributions in both shape and peak. The similar behavior of all the numerical experiments can be due to the low frequency nudging to the MERCATOR fields. Overall, above the MLD (Figures 6a-b), the numerical experiments overestimate the stratification. Note that this is not evident by just looking at the median profiles in Figure 3. As values of Δ_N^2 in Table 3 indicate, the distribution more resembling the in situ one is KEfluCANMINkOpt.

Below the MLD (Figures 6c-d) the peaks of the numerical PDFs are less aligned than above the MLD. In this layer, the numerical distributions show a peak at $N \approx 5 \times 10^{-3.3} s^{-1}$ that is not present in the measurements. As values of Δ_N^2 in Table 4 show, KEfluCANMINk is the distribution with the shape more resembling the in situ one above the MLD, while KEfluCAN is

15

the more resembling below the MLD. However, we note that, when the secondary peak is removed, the numerical experiments with the better PDF of N are KEfluCANMINk and KEfluCANMINkOpt.

The probability density functions of the in situ eddy viscosity data (in black) and of the numerical data are shown in Figure 7. Above the MLD (Figure 7a-b), the in situ values have three peaks with a higher peak at $K_Z \approx 5 \times 10^{-4} m^2/s^2$. None of the numerical experiments reproduces this behavior. As values of σ_{K_Z} in Table 2 suggest, above the MLD the distribution most aligned to the in situ one is KLsetMINk. Besides, $\Delta^2_{K_Z}$ values in Table 3 indicate that KEfluCANMINkOpt is the distribution with the shape most similar to the observations above the MLD. The effect of raising k_{min} is similar on the two SMC families: it shifts the distribution towards higher values and modifies the peaks' height. However, KLsetMINk and KLfluMINk show a prominent first peak and thus perform less well than KEfluCANMINk and KEfluCANMINkOpt.
Below the MLD (Figures 7c-d), similarly to what seen for ε, the numerical distributions are clearly different from the in situ one. $\Delta^2_{K_Z}$ values in Table 4 indicate that KEfluCAN, KEfluCANMINk, KEfluCANMINkOpt and KLfluMINk are the distributions with the shape more resembling the in situ one. But, as we already noted for the distributions of ε below the MLD (Figure 5), in this layer there is not a real difference between the numerical experiments.

3.3. (ε,N,K_Z) space

All the measurements of ε, N and K_Z and the corresponding numerical data are shown in Figure 8. The point clouds reveal that the in situ values are much less disperse than the model ones (not evident from the picture; see Supplementary Data). In Figure 8a-b we can see six main planes corresponding to: KEfluCAN, KEset and KEflu, KLset, KLflu and some in situ data at low K_Z. Indeed, these last points lie on a plane with constant $k = k_{min}$ as prescribed by the parametrization we used (see Methods and Discussion). Data belonging to the numerical experiments with a higher threshold on k (KEfluCANMINk, KEfluCANMINkOpt, KLsetMINk and KLfluMINk) are first found at $K_Z \approx 2 \times 10^{-6}$ (Figure 8b). The KE numerical experiments with $k_{min} = 10^{-8} m^2/s^2$ have planes with an average value of ε that is lower than the average value of ε of the KL numerical experiments with $k_{min} = 10^{-8} m^2/s^2$. These last ones have lower average values of

16

ε than the numerical experiments with $k_{min} = 10^{-7} m^2/s^2$. This closely mirrors what we already noted in Figure 5c-d. But it turns out that the points lying on the planes in Figure 8 do not belong exclusively to depths greater than the MLD (the differentiation between above and below the MLD is not shown in Figure 8).

4. Discussion

By considering the median profiles, the values of logarithmic standard deviation and the shape of the probability density functions, we can progress towards the identification of an optimal Symphonie's configuration with respect to our dataset.

From visual inspection of the median profiles it is apparent that the default threshold on k (used for KEset, KEflu, KLset, KLflu) underestimates the turbulent activity (5a-b). By using Equation 1 and a value of $\varepsilon \approx 10^{-4} m^2/s^3$ that we can infer from the in situ data below the MLD (Figure 2), we estimated the value of a more appropriate threshold ($k \approx 10^{-7} m^2/s^2$) that was used for the numerical experiments with the higher threshold (KEfluMINk, KEfluMINkOpt, KLsetMINk and KLfluMINk). As Figures 2 and 6 and the values of σ_ε and σ_{K_Z} in Table 2 indicate, this new threshold permitted to get better estimates of both ε and K_Z.
The decadal standard deviation values are suited for whole-column comparisons between numerical experiments and in situ data, when it is interesting to have the general behavior of the numerical experiment (e.g., Burchard et al., 2002). However, we are also interested to details of the model's configuration that likely have a major effect only in the MLD and cannot be appreciated by this metric.
The comparison of the shape of the numerical and in situ probability density functions is more appropriate to examine the model's behavior above the MLD. In particular, it permits to investigate its response to the different configuration choices on the full range of values of ε, N and K_Z. For example, Figure 5a-b shows that the effect of raising the threshold value of k is not just simply a shift, as is suggested by the σ_ε values and the median profiles (Figure 2). In fact, the numerical experiments with a higher threshold (KEfluCANMINk, KEfluCANMINkOpt, KLsetMINk, KLfluMINk) are not merely shifted towards higher values but also show significant differences in

17

the peaks' heights. This is due to two facts: (i) with higher values of kinetic energy, it is generally more probable to observe higher values of dissipation, and (ii) the highest values in each of the ε PDF' peaks cannot be arbitrarily large because the left-hand peak mainly contains values below the MLD and vice-versa for the right-hand peak (data not shown).

Regarding the average value of the MLD estimated by the different numerical experiments, we note that it is systematically lower than the observed one. Generally, the numerical MLD coincides with the depth at which the kinetic energy reaches its minimum value (Figure S2 in the Extra Materials). Possible reasons for the mismatch could be some missing or not well parameterized surface turbulent phenomena (e.g., Langmuir turbulence), as well as the uncertainty on the drag coefficients.

The analysis of ε data above the MLD allows examination of the effect of boundary conditions, stability functions and optical scheme. The effect of the surface boundary conditions is readily seen by visually inspecting the ε median profiles (Figure 2). The KEset, KEflu, KLset and KLflu profiles highlight that, at the surface, the equilibrium boundary condition determines ε values lower than both the observations and the values of the corresponding numerical experiments employing a flux boundary condition. Note that the values of ε close to the surface that one obtains using a given boundary condition also depend on the employed SMC. In fact, the $k - \varepsilon$ and the $k - \ell$ schemes have different values of the constants in the surface boundary condition for the kinetic energy (see Appendix A).
As a consequence, an influence of the boundary conditions is also found in the PDFs of ε (Figure 5). In fact, the value of ε in the upper tail of the PDF is different for the schemes employing the equilibrium boundary condition and the ones employing the flux boundary condition as one would expect as the higher values near the surface are predominantly affected by the surface boundary conditions rather than by the threshold on k. Indeed, the right peak at higher ε values is mainly due to values from the depth range $0 \leq z/MLD \leq 0.5$.
Regarding the average values of the MLD, we find that all the numerical experiments predict a shallower mixed layer depth. A reason for this behavior could be the uncertainty in the drag coefficient and/or to the presence of some not well-parameterized mixing phenomenon.
Looking at the profiles of turbulent production terms P and B in Equations

18

B.1 and B.1 (Figures S3 and S4 in the Extra Materials), we find that turbulence is mainly due to shear. In particular, we find that shear is enhanced at the mixed layer depth. This indicates that the increased levels of turbulence right below the MLD that we observe in situ (Figure 2) are due to internal waves s or intermittent KelvinHelmholtz instability within the stratified layer (Woods, 1968; Grant and Belcher, 2011). The fact that the model does not reproduce this enhanced turbulence is due to the too coarse resolution that does not permit resolving such processes.

Secondly, the effect of the stability functions is made clear by both the median profiles and the PDFs of ε. In Figure 2 it appears that the Canuto A stability functions improve the agreement with respect to the observations (cf. KEflu vs. KEfluCAN), especially near the surface. The same information is given by the PDFs (Figure 5; Table 3): KEfluCAN's PDF has just two peaks as the in situ PDF and, thus, it is in better agreement with the observations. This result did not change when we raised the threshold on the kinetic energy.

Thirdly, the effect of the optical scheme is best seen by looking at the ε PDF (Figure 5). Indeed, we see that the modified optical scheme enables KEfluCANMINkOpt to reproduce remarkably well the in situ distribution. We attribute this significant improvement to the better estimate of the temperature and salinity profiles (see Figures S5 and S6 in the Extra Materials) that, in turn, improves the stratification. Indeed, the profile of N (Figure 6) of KEfluCANMINkOpt is slightly better than the one of KEfluCANMINk. Even if this is not captured by the values of S_N (Table 3), we note that, right below the MLD, KEfluCANMINkOpt better reproduces the peak that we observe in the in situ stratification. However, it is not straightforward to say how this can influence the dynamics above the MLD.(indeed the KEfluCANMINkOpt's N PDFs are totally comparable with the ones of KEfluCANMINk). The only obvious difference that we observe above the MLD between the KEfluCANMINk's and KEfluCANMINkOpt's stratifications is that the median profile of the buoyancy turbulent production term B of KEfluCANMINkOpt is significantly higher than the one of KEfluCANMINk (Figure S4 in the Extra Materials).

The analysis of the PDFs of N indicates that all the numerical experiments overestimate the stratification above in the surface layer (Figures

19

6a-b). On the other hand the stratification is well reproduced by all the numerical experiments below the MLD. The only absent feature in the numerical data of N is a bump of high values right below the MLD where the entrainment takes place and where we also observe higher levels of turbulence (see Figure 2). Overall, below the MLD the numerical PDFs of ε and K_Z markedly differ from the in situ ones: they are more peaked than the in situ PDFs. A possible cause for this behavior could be the too coarse resolution of the model that does not permit to adequately resolve turbulent processes below the MLD or to the different natures of mixing mechanisms.

Also the analysis of the K_Z probability density functions leads to the conclusion that KEfluCANMINkOpt is the better configuration. Even though the values of $\Delta^2_{K_Z}$ in Table 3 indicate that there is not a clear difference between KEfluCANMINk and KEfluCANMINkOpt, we note that KEfluCANMINkOpt K_Z PDF has three peaks as the in situ PDF whereas KEfluCANMINk's PDF has only two. In general, the $k-\varepsilon$ scheme performs significantly better in reproducing the in situ PDF of K_Z. This is likely due to the advantage of $k-\varepsilon$ in estimating ε with respect to $k-\ell$.

The inclusion of the molecular diffusivity D_T in the formulation of K_Z did not cause significant differences, probably because the turbulence level is too high to allow to appreciate the difference. Here we did not show a numerical experiment in which we change only this aspect of the model's implementation. However, one example is given in the Extra Materials (Figure S7).

The analysis of the data below the MLD shows a sharp disagreement between the numerical and the in situ values of ε and K_Z. However, we can observe three main features: i) the typical value of numerical ε depends on the SMC that one employs; ii) the model values of ε and K_Z are much less variable than the measurements; iii) setting a hard threshold on the kinetic energy limits the reliability of the modeling in this layer.
The first point is made clear by Figure 2. In fact, we can see that when the same threshold on k is applied, the ε values obtained with the $k-\varepsilon$ scheme are one order of magnitude smaller than the ones obtained with the $k-\ell$ scheme. This is partially explained by the fact that the two SMCs use different values for the constants in Equations B.1 and B.2 that relate ε to the kinetic energy. Indeed, $k-\varepsilon$ uses a lower value ($c_0 = 0.55$) than what

20

$k - \ell$ does $c_\varepsilon = 0.88$. However, this fact alone cannot explain the one-order-of-magnitude difference that we observe below the MLD.

The second point is illustrated by Figures 5c-d and 6c-d. Therein, the numerical PDFs in the lower layer are much less spread out than the in situ ones. A possible cause for this could be the resolution of the model that does not permit to adequately resolve the turbulent processes in this layer.

Also, a possible cause for the sharp discrepancy between numerical and in situ PDFs below the MLD could be the use of a hard threshold for k. In fact, Figure 8 suggests that this choice leads to a non-physical result, with many low-energy points on a single plane in the space (K_Z, N, ε). This behavior results directly from Equations B.6 and B.7 implying $\varepsilon = Pr^{-1} c_k c_\varepsilon k^2 / K_Z$ for the KL scheme and from Equations B.3 and B.4 implying $\varepsilon = \sqrt{2} c_0^3 S_M k^2 / K_Z$ for the KE scheme (where c_k, c_ε and c_0 are constants). This is exactly the behavior we observe in Figure 8b, with the planes following the line $\log \varepsilon = -\log K_Z + const$. Thus, the points on the planes in Figure 8 are values with a minimal kinetic energy. In situ data do not show this behavior, suggesting that specifying a hard value for k_{min} is likely not the best choice. A more suitable approach would probably be a parametrization of k_{min} on the base of the different turbulent processes at play. For example, Gaspar et al. (1990) hypothesized the use of a parametrization on the internal wave field activity when dealing with data in a depth range similar to the one in our analysis. Note here that such a solution is likely to depend on the amount of numerical mixing and to the processes that are resolved by each model.

5. Conclusions

Vertical mixing in the surface layer of the ocean plays an essential role in both physical and biochemical phenomena. Therefore, it must be correctly estimated by numerical models. Different turbulence closure schemes have different performances in predicting mixing. Moreover, the problem is made more complicated by the interplay of the SMC with other aspects of ocean dynamical numerical modelling like boundary conditions, stability functions and optical scheme. The influence of the minimum value of kinetic energy that is allowed has also been investigated.

We performed nine numerical experiments and compared their estimates

21

of the turbulent quantities ε, N and K_Z to in situ microstructure measurements in the Gulf of Lion. In particular, two SMCs were considered: a $k - \varepsilon$ scheme proposed by Burchard and Bolding (2001) (herein KE); and a $k - \ell$ based on Gaspar et al. (1990) (herein KL). We considered two surface boundary conditions: one supposing equilibrium between the production and dissipation terms in the dynamic equation for k (herein set); and one taking into account the effect of breaking waves of all scales based on Craig and Banner (1994) (herein flu). We considered the stability functions proposed by Kantha and Clayson (1994) and Canuto et al. (2001). In addition, we also considered two different attenuation lengths for the photosynthetically available radiation in the model's optical scheme. The combinations of these factors that we explored are resumed in Table 1.

A recent study by Reffray et al. (2015) showed that, in a 1-D situation, the KE closure scheme gives better mixing estimates than other widely used SMCs. However, that study did not compare KE to a mixing length scheme based on Gaspar et al. (1990). Our study fills this gap. Moreover, our study exploits a 3-D model that, by definition, considers more terms in the dynamical equations than what 1-D models do. Additionally, it compares the modeled turbulence activity directly to microstructure turbulence measurements rather than with derived quantities like the mixed layer depth. The two SMCs do not show relevant differences when estimating K_Z both above and below the MLD. The fact that the two SMC perform similarly in estimating K_Z when KL has an advantage in estimating ε, is not explained by differences in estimating N and need further investigation in the future. Previous studies, in 1-D and 3-D numerical simulations, gave conflicting indications on the effect of the stability functions proposed by Kantha and Clayson (1994) and Canuto et al. (2001) on the KE scheme. We find that the effect of the Canuto A stability functions on the KE is to improve the performance of the closure scheme when the PDFs of ε and K_Z are considered.

The value of the kinetic energy threshold plays a pivotal role in approaching the observations below the MLD (see Figures 2 and 4). Nonetheless, we found a non physical behavior of the numerical experiments for low kinetic energy levels. This supports the idea that the minimum of kinetic energy should rather be parametrized as function of different turbulent mechanisms rather than being a hard threshold (as already hypothesized by Gaspar et al.,

22

1990). Such a parameterization should also depend on the SMC that is employed. Indeed, we found that the two SMC predict different turbulence levels below the MLD even when they employ the same threshold on the turbulent kinetic energy.

765

We found that the attenuation length of the photosynthetically available radiation plays an important role in determining the stratification and, as a consequence, the performance of the model in predicting ε and K_Z. This result highlights the importance of biological activity in influencing physical 770 processes.

Our study shows that the comparison with in situ microstructure data can effectively help in setting up the implementation of a SMC for an ocean numerical model. In the future, we should expect more utilization of mi-775 crostructure data as new automated instrumentation becomes available. In particular, microstructure and wave height probes mounted on autonomous platforms, such as drifting profilers or gliders, will permit new and bigger datasets, especially during adverse meteorological situations. This kind of data will provide essential data to developers of future ocean-wave coupled 780 models, as these are expected to significantly advance the ocean modelling state of the art by reducing the need of general surface boundary conditions.

for constructive criticism of our manuscript and two anonymous reviewers for their insightful comments on the paper, as these comments led to a significant improvement of the work.

A. Boundary conditions

A.1. Surface boundary conditions for k
A.1.1. Equilibrium

The k surface boundary conditions can be obtained hypothesizing equilibrium between production and dissipation. The surface energy production due to wind shear can be expressed as:

$$K_Z \left(\frac{\partial u}{\partial x}, \frac{\partial v}{\partial y} \right) = \frac{(\tau_x, \tau_y)}{\rho_0}$$

where $\tau = \sqrt{\tau_x^2 + \tau_y^2}$ is the wind stress and ρ_0 is the reference density of sea water.

By inserting this last equation in the expression for the shear production term (Equation B.1) and using the expression for ε proposed by Galperin et al. (1988) (Equation B.1), we obtain the surface boundary condition for k:

$$k_s = \frac{\tau/\rho_0}{\sqrt{2^{1/2} c_0^3 S_M}}$$

Herein, S_M is obtained by supposing a null stratification (i.e., $G_h = 0$ in Equation 24 in Galperin et al. 1988). Under this hypothesis, S_M becomes constant. In the general case, the expression for S_M are given by Kantha and Clayson (1994) and Canuto et al. (2001). The value of the stability coefficient c_0 for the $k - \varepsilon$ scheme is issued from Warner et al. (2005). The value of c_0 for the $k - \ell$ scheme is issued from Gaspar et al. (1990).

A.1.2. Flux

Alternatively, the boundary conditions can be specified as surface flux conditions, namely:

$$K_Z \frac{\partial k}{\partial z} = F \tag{A.1}$$

24

where the surface flux can be computed as $F = 100(\tau/\rho_0)^{3/2}$ (Craig and Banner, 1994) or directly prescribed from the 'wave to ocean' turbulence flux computed by a wave model when available.

A.2. Bottom and surface boundary conditions for ε

The ε surface and bottom conditions are computed on the first level under the surface and above. Let z_1 denotes the distance between this level and the considered boundary. Boundary conditions for ε are obtained from k and Equation B.2, using the latter with some appropriate hypothesis for l_B a boundary length scale value. A simple formulation (Warner et al., 2005) is eventually given by $l_B = 0.4(z_1 + z_0)$, where z_0 is a length scale representing the roughness of the bottom boundaries. Unfortunately, this formulation potentially leads to unrealistic high values when the underlying hypothesis of neutral stratification is no longer valid (a situation that is more likely to occur in deep zone where the vertical grid resolution near the bottom is generally coarse). One way to solve this problem is to introduce a dependency on the gradient Richardson number (Estournel and Guedalia, 1987; Michaud et al., 2012):

$$l_B = \begin{cases} 0.4(z_1 + z_0) & \text{if} \quad Ri < 0 \\[2em] 0.4(z_1 + z_0)(1 - 5Ri) & \text{if} \quad 0 \leq Ri \leq 0.16 \\[2em] 0.4(z_1 + z_0)(1 + 41Ri)^{-0.8} & \text{if} \quad Ri > 0.16 \end{cases}$$

B. Closure schemes

B.1. $k - \varepsilon$ scheme

Following Burchard and Bolding (2001), the equations describing the dynamics for k and ε are:

$$\frac{dk}{dt} = \frac{\partial}{\partial z}\left(K_Z \frac{\partial k}{\partial z}\right) + P + B - \varepsilon \tag{B.1}$$

$$\frac{\partial \varepsilon}{\partial t} = \frac{\partial}{\partial z}\left(\frac{K_Z}{\sigma} \frac{\partial k}{\partial z}\right) + \frac{\varepsilon}{k}(c_1 P + c_3 B - c_2 \varepsilon) \tag{B.2}$$

25

where

$$P = K_Z \left\{ \left(\frac{\partial u}{\partial z} \right)^2 + \left(\frac{\partial v}{\partial z} \right)^2 \right\} \tag{B.3}$$

and

$$B = \frac{g}{\rho_0} K_T \frac{\partial \rho}{\partial z} \tag{B.4}$$

are the production terms due to shear and buoyancy respectively. Parameters values issued from Warner et al. (2005).

The eddy viscosity K_Z and the temperature eddy diffusivity K_T used in Equations B.1 and B.1 are given by:

$$K_Z = \sqrt{2k} l S_M, \quad K_T = \sqrt{2k} l S_H \tag{B.5}$$

The turbulent length l is related to k, TKE and ε according to

$$l = c_0^3 k^{3/2} \varepsilon^{-1} \tag{B.6}$$

S_M and S_H , the quasi-equilibrium stability functions defined by Kantha and Clayson (1994), depend on the Richardson flux number.

B.2. $k - \ell$ scheme

We used the $k - \ell$ closure scheme proposed by Gaspar et al. (1990). Therein the authors assume ε to be (following Kolmogorov, 1942):

$$\varepsilon = c_\varepsilon k^{3/2} / l_\varepsilon \tag{B.7}$$

where k is the TKE, $k = \sqrt{u'^2 + u'^2 + w'^2}$ and $c_\varepsilon = 0.7$ (following Bougeault and Lacarrere, 1989).

The eddy momentum diffusivity is related to TKE according to:

$$K_Z = c_k l_k \sqrt{k}$$

where c_k has to be determined.

The dissipation and mixing length scales l_k and l_ε are the ones proposed by Bougeault and Lacarrere (1989):

$$l_\varepsilon = (l_u l_d)^{1/2}$$

$$l_k = \min(l_u, l_k)$$

that are determined through

$$\frac{g}{\rho_0} \int_z^{z+l_u} [\overline{\rho}(z) - \overline{\rho}(z')] dz' = k(z)$$

$$\frac{g}{\rho_0} \int_z^{z-l_d} [\overline{\rho}(z) - \overline{\rho}(z')] dz' = k(z)$$

where ρ is the water density.

As highlighted by Gaspar et al. (1990), these length scales have a straightforward physical interpretation: they are the distances traveled upward/downward by a fluid particle by converting all of its kinetic energy into potential energy.

Hypothesizing the turbulence to be stationary and homogeneous Gaspar et al. (1990) show that, in stably stratified regions, the model parameterizations yields:

$$K_Z = \sqrt{2} Pr^{-1} c_k k N^{-1} \tag{B.8}$$

$$\varepsilon = \frac{1}{\sqrt{2}} c_\varepsilon k N \tag{B.9}$$

from which it follows that $c_k = 0.1$.

C. Segmentation and quality fit

In order to apply the Batchelor's theory to the temperature gradient one has to divide the data into segments in which turbulence can be considered homogeneous, viz. the turbulent motion can be regarded as a random motion which average properties are independent of position in the fluid. Given the fact that the first aim is to fit the data to a theoretical Batchelor's spectrum,

27

arguably the best segmentation method is the one giving the fits with the best quality.

To determine the fit quality we use a Maximum Likelihood Estimation technique. Following Ruddick et al. (2000), we calculated the goodness of a fit by maximizing the joint probability of the measured spectrum with respect to the theoretical Batchelor spectrum. Hence, the higher the joint probability, the higher the fit goodness. In the following we call this joint probability JP (corresponding to $C11$ in Ruddick et al., 2000).

To make the fit goodness criteria more rigorous, following Sanchez et al. (2011) we also require that: i) the mean absolute deviation (MAD) of the ratio between the in situ and theoretical spectra within the fitting domain be lower than 1.1; ii) the signal to noise ratio (SNR) be lower than 1.3; iii) the likelihood ratio (LHR) –which quantifies if the measured spectrum fits the Batchelor's spectra or a power-law spectrum better– be lower than 2.

We tested which of the following segmentation methods permitted to have the better fits in our dataset: a segmentation based on an eight order AR model by Imberger and Ivey (1991) (II91); a constant segmentation of 1024 data points in which the values of temperature gradient variance in 7 sub-segments are in the same order of magnitude proposed by Sanchez et al. (2011) (S11), a constant segmentation of 1024 data points with no overlap suggested by Cuypers et al. (2012) (C12); a constant segmentation of 128 data points with 50% overlap suggested by Moniz et al. (2012) (M12); a constant 512 data points segmentation with no overlap; and a constant 1024 data points segmentation with 30% overlap. In practice, we fitted the segmented data to the Batchelor's spectrum following the fitting procedure of Steinbuck et al. (2009) and we compared the results searching for the distribution of JP values with the greater proportion of high values. The results are depicted in Figure 9.

In Figure 9 the histogram of the JP values is plotted with the cumulative distribution of the latter. Note that each histogram is normalized to the total number of segments obtained with the corresponding segmentation method. As it is evident by eye, the M12 is the segmentation method that gives the highest percentage of good profiles, as highlighted by the higher histogram bars on the right end of the distribution with respect to the other cases. Moreover this is the method that has the vertical line lying at the

28

highest value, meaning that a higher percentage of segments have a better fit. Furthermore cases b), e) and d) show that progressively reducing the segmentation window gives better results. Case f) also shows, when compared to c), that adding overlap, though permitting to have a higher resolution, does not augment the percentage of good fits. These evidences suggest to apply a 128 points ($\approx$ 12.5 cm) segmentation with no overlap. Anyway, not needing such a high resolution in order to compare the in situ data with the numerical data, we avoided the overlap. The distribution of JP values of this segmentation method did not significantly differ from the one of M12 (data not shown).

References

Agrawal, Y., Terray, E., Donelan, M., Hwang, P., III, A.W., Drennan, W., Kahma, K., Krtaigorodskii, S., 1992. Enhanced dissipation of kinetic energy beneath surface waves. Nature 359, 219–220.

Alberola, C., Millot, C., 1995. On the seasonal and mesoscale variabilities of the Northern Current during the PRIMO-0 experiment in the western Mediterranean Sea. Oceanol. Acta 2, 163–192.

Anis, A., 2006. Similarity relationships in the unstable acquatic surface layer. J. Marine Syst. 33. L19609.

Anis, A., Singhal, G., 2002. Mixing in the surface boundary layer of a tropical freshwater reservoir. J. Marine Syst. 33, 225–243.

Arakawa, A., Lamb, V., 1977. Computational design of the basic dynamical processes of the UCLA general circulation model. Meth. Comp. Phys. 17, 173–275.

Arneborg, L., Fiekas, V., Ulmaf, L., Burchard, H., 2007. Gravity current dynamics and entrainment - a process study based on observation in the Arkona basin. J. Phys. Oceanogr. 37, 2094–2113.

Barrier, N., Petrenko, A., Ourmieres, Y., 2016. Strong intrusions of the Northern Mediterranean Current on the eastern Gulf of Lion: insights from in-situ observations and high resolution numerical modelling. Ocean Model. 66, 313–327.

Batchelor, G., 1959. The spectrum of small-scale variation of convected quantities like temperature in turbulent fluid. part 1. general discussion and the case of small conductivity. J. Fluid Mech. 5, 113–133.

Bouffard, D., Boegman, L., 2013. A diapycnal diffusivity model for stratified environmental flows. Dynam. Atmos. Oceans 61-62, 14–34.

Bougeault, P., Lacarrere, P., 1989. Parameterization of orography-induced turbulence in a meso-beta scale model. Mon. Weather Rev. 117, 1872–1890.

Bourrin, F., Friend, P., Amos, C., Manca, E., C. Ulses and, A.P., de Madron, X.D., Thompson, C., 2011. Sediment dispersal from a typical Mediterranean flood: The Têt River, Gulf of Lion. Cont. Shelf Res. 28, 1895–1910.

Bowman, A., Azzalini, A., 1997. Applied Smoothing Techniques for Data Analysis. New York: Oxford University Press Inc.

Burchard, H., Bolding, K., 2001. Comparative analysis of four second-moment turbulence closure models for the oceanic mixed layer. J. Phys. Oceanogr. 31, 1943–1968.

Burchard, H., Bolding, K., Rippeth, T., Stips, A., Simpson, J., Sundermann, J., 2002. Microstructure of turbulence in the northern north sea: a comparative study of observations and model simulations. J. Sea Res. 47, 223–238.

Burchard, H., Craig, P., Gemmrich, J., van Haren, H., Mathieu, P.P., Meier, H.M., Smith, W.M.N., Prandke, H., T.P. Rippeth and, E.S., Smyth, W., Welsh, D., Wijesekera, H., 2008. Observational and numerical modeling methods for quantifying coastal ocean turbulence and mixing. Progr. Oceanogr. 76, 399–442.

Canuto, V., Howard, A., Cheng, Y., Dubovikov, M., 2001. Ocean turbulence. Part I: One-point closure modelmomentum and heat vertical diffusivities. J. Phys. Oceanogr. 31, 1413–1426.

Conan, P., Millot, C., 1992. Variability of the Northern Current off Marseille, western Mediterranean Sea, from February to June 1992. Oceanol. Acta 2, 193–205.

30

Craig, P., Banner, M., 1994. Modelling wave-enhanced turbulence in the ocean surface layer. J. Phys. Oceanogr. 24, 2546–2559.

Cuypers, Y., Bouruet-Aubertot, P., Marec, C., Fuda, J.L., 2012. Characterization of turbulence and validation of fine-scale parametrization in the Mediterranean Sea during BOUM experiment. Biogeosciences 9, 3131–3149.

Dufau-Julliand, C., Marsaleix, P., Petrenko, A., Dekeyser, I., 2004. Three-dimensional modeling of the Gulf of Lions hydrodynamics (northwest Mediterranean) during January 1999 (MOOGLI3 Experiment) and late winter 1999: Western Mediterranean Intermediate Waters (WIWs) formation and its cascading over theshelf break. J. Geophys. Res. 190.

Estournel, C., Broche, P., Marsaleix, P., Devenon, J., Auclair, F., Vehil, R., 2001. The Rhone river plume in unsteady conditions : numerical and experimental results. Estuar. Coast. Shelf S. 53, 25–38.

Estournel, C., Guedalia, D., 1987. A new parameterization of eddy diffusivities for nocturnal boundary layer modelling. Bound. Lay. Meteorol. 39, 191–203.

Estournel, C., de Madron, X.D., Marsaleix, P., Auclair, F., Julliand, C., Vehil, R., 2003. Observation and modeling of the winter coastal oceanic circlation in the Gulf of Lion under wind conditions influenced by the continental orography (FETCH experiment). J. Geophys. Res.-Oceans 108, 25–38.

Estournel, C., Testor, P., Damien, P., amd P. Marsaleix, F.D., Conan, P., Kessouri, F., de Madron, X.D., Coppola, L., Lellouche, J.M., Belamari, S., Mortier, L., Ulses, C., Bouin, M., Prieur, L., 2016. High resolution modeling of dense water formation in the north-western Mediterranean during winter 2012-2013: Processes and budget. J. Geophys. Res. 121.

Estournel, C., Zervakis, V., Marsaleix, P., Papadopoulos, A., Auclair, F., Perovolitis, L., Tragou, E., 2005. Dense water formation and cascading in the Gulf of Thermalikos (North Aegean), from observaions and modelling. Cont. Shelf Res. 25, 2366–2386.

Farmer, D., Li, M., 1995. Patterns of bubble clouds organized by Langmuir circulation. J. Phys. Oceanogr. 25, 1426–1440.

31

Ferrari, R., 2014. Oceanography: What goes down must come up. Nature 513, 179–180.

Ferrari, R., Wunsh, C., 2009. Ocean circulation kinetic energy-reservoirs, sources and sinks. Ann. Rev. Fluid Mech. 41, 253–282.

Flexas, M.M., de Madron, X.D., Garcia, M.A., Canals, M., Arnau, P., 1997. Flow variability in the Gulf of Lions during the MATER HFF experiment (March-May 1997). J. Mar. Sys. 33, 197–214.

Flierl, G., Davis, C., 1993. Biological effects of Gulf Stream meandering. J. Marine Res. 51, 529–560.

Fraysse, M., Pairaud, I., Ross, O., Faure, V., Pinazo, C., 2014. Intrusion of Rhone river diluted water into the Bay of Marseille: generation processes and impacts on ecosystem functioning. J. Geophys. Res. Oceans 119.

Galperin, B., Kantha, L., Hassid, S., Rosati, A., 1988. A quasi-equilibrium turbulent energy model for geophysical flows. J. Atmos. Sci 55, 45–62.

Gaspar, P., Gregoris, Y., Lefevre, J., 1990. A simple eddy kinetic energy model for the simulation of oceanic vertical mixing: Tests at station Papa and long-term upper ocean study site. J. Geophys. Res. 95, 16179–16193.

Gatti, J., Petrenko, A., Leredde, Y., Devenon, J.L., Ulses, C., 2006. The Rhone river dilution zone present in the northeastern shelf of the Gulf of Lion in December 2003. Continental Shelf Res. 26, 1794–1805.

Gibson, C., Schwarz, W., 1963. The universal equilibrium spectra of turbulent velocity and scalar fields. J. Fluid Mech. 16, 365–384.

Grant, A., Belcher, S., 2011. Wind-driven mixing below the oceanic mixed layer. J. Phys. Oceanogr. 41, 1556–1575.

Han, L., 2014. A two-time-level split-explicit ocean circulation model (MAS-NUM) and its validation. Acta Oceanol. Sin. 33, 11–35.

Hauser, D., Branger, H., Bouffies-Cloché, S., Despiau, S., Drennan, W.M., Dupuis, H., Durand, P., de Madron, X.D., Estournel, C., Eymard, L., Flamant, C., Graber, H.C., Gurin, C., Kahma, K., Lachaud, G., Lefvre, J.M., Pelon, J., Pettersson, H., Piguet, B., Queffeulou, P., Tailliez, D.,

Tournadre, J., Weill, A., 2003. The FETCH experiment: an overview. J. Geophys. Res. Oceans 108.

Herrmann, M., Diaz, F., Estournel, C., Marsaleix, P., Ulses, C., 2013. Impact of atmospheric and oceanic interannual variability on the Northwestern Mediterranean Sea pelagic planktonic ecosystem and associated carbon cycle. J. Geophys. Res. 118, 1–22.

Herrmann, M., Estournel, C., Adloff, F., Diaz, F., 2014. Impact of climate change on the northwestern Mediterranean Sea pelagic planktonic ecosystem and associated carbon cycle. J. Geophys. Res. Oceans 119, 5815–5836.

Hogg, A., Ivey, G., Winters, K., 2001. Hydraulics and mixing in controlled exchange flows. J. Geophys. Res. 106, 959–972.

Hossain, M.S., 1980. Mathematische Modellierung von turbulenten Auftriebs-sstromungen. Ph.D. dissertation. University of Karlsruhe, Karlsruhe, Germany.

Hu, Z., Doglioli, A., Petrenko, A., Marsaleix, P., Dekeyser, I., 2009. Numerical simulations of eddies in the Gulf of Lion. Ocean Modell. 28/4, 203–208.

Hu, Z., Petrenko, A., Doglioli, A., Dekeyser, I., 2011. Numerical study of eddy generation in the western part of the Gulf of Lion. J. Geophys. Res. Oceans 116.

Ilicak, L., Özgökmen, T., Peters, H., Baumert, H., Iskandarini, M., 2008. Performance of two-equation turbulence closures in three-dimensional simulations of the Red Sea overflow. Ocean Modell. 24, 122–139.

Imberger, J., Ivey, G., 1991. On the nature of turbulence in a stratified fluid. Part I: The energetics of mixing. J. Phys. Oceanogr. 21, 650–658.

Jurado, E., van der Woerd, H., Dijkstra, H., 2012. Microstructure measurements along a quasi-meridional transect in the northeastern Atlantic Ocean. J. Geophys. Res. 117, C04016.

Kantha, L., Clayson, C., 1994. An improved mixed layer model for geophysical applications. J. Geophys. Res. 99, 25235–25266.

33

Kantha, L., Clayson, C., 2000. Processes en geophysical fluid flows. International Geophysics Series. Vol. 67 .

Kersalé, M., Petrenko, Doglioli, A., Dekeyser, I., Nencioli, F., 2013. Physical characteristics and dynamics of the coastal Latex09 Eddy derived from in situ data and numerical modeling. J. Geophys. Res. 118, 1–11.

Kolmogorov, A., 1942. The equation of turbulent motion in an incompressible fluid. Izv. Akad. Nauk SSSR ser Fiz. 6, 56–58.

Large, W., McWilliams, J., Doney, S., 1994. Oceanic vertical mixing: a review and a model with a non-local boundary layer parametrization. Rev. Geophys. 32, 363–403.

Large, W., Yeager, S., 2004. Diurnal to decadal global forcing for ocean and sea-ice models: The data set and flux climatologies. NCAR Tech. Note NCAR/TN-4601STR doi:10.5065/D6KK98Q6.

Lozovatsky, I., Figueroa, M., Roget, E., Fernando, H., Shapovalov, S., 2005. Observation and scaling of the upper mixed layer in the North Atlantic. J. Geophys. Res. 110.

Luketina, D., Imberger, J., 2000. Determining turbulent kinetic energy dissipation from Batchelor curve fitting. JAOT 18, 100–113.

Maraldi, C., Chanut, J., Levier, B., Ayoub, N., Mey, P.D., Reffray, G., Lyard, F., Cailleau, S., Drevillon, M., Fanjul, E., Sotillo, M., Marasalaix, P., the Mercator Research, Team, D., 2013. NEMO on the shelf: assessment of the Iberia-Biscay-Ireland configuration. Ocean Sci. 9, 745–771.

Marchesiello, P., Debreu, L., Couvelard, X., 2009. Spurious diapycnal mixing in terrain-following coordinate models: The problem and a solution. Ocean Modell. 26, 156–169.

Marsaleix, P., Auclair, F., Duhaut, T., Estournel, C., Nguyen, C., Ulses, C., 2012. Alternatives to the Robert-Asselin filter. Ocean Modell. 41, 53–66.

Marsaleix, P., Auclair, F., Estournel, C., 2006. Considerations on open boundary conditions for regional and coastal ocean models. Ocean Technol. 23, 1604–1613.

Marsaleix, P., Auclair, F., Estournel, C., 2009. Low-order pressure gradient schemes in sigma coordinate models: The seamount test revisited. Ocean Model. 30, 169–177.

Marsaleix, P., Auclair, F., Floor, J., Herrmann, M., Estournel, C., Pairaud, I., Ulses, C., 2008. Energy conservation issues in sigma-coordinate free-surface ocean models. Ocean Model. 20, 61–89.

Mellor, G., Yamada, T., 1982. Development of a turbulence closure model for geophysical fluid problems. Geophys. Res. Phys. 20, 851–855.

Michaud, H., Marsaleix, P., Leredde, Y., Estournel, C., Bourrin, F., Lyard, F., Mayet, C., Ardhuin, F., 2012. Three-dimensional modelling of wave-induced current from the surf zone to the inner shelf. Ocean Sci. 8, 657–681.

Millot, C., 1990. The Gulf of Lions' hydrodynamics. Cont. Shelf Res. 10, 885–894.

Moniz, R., Fong, D., Woodson, C., Willis, S., Stacey, M., Monismith, S., 2012. Scale-dependent dispersion within the stratified interior on the shelf of northern Monterey Bay. J. Phys. Oceanogr 44, 1049–1064.

Nerini, D., Ghattas, B., 2007. Classifying densities with functional regression trees : an application in oceanology. Comput. Stat. Data An. 101, 4287–4297.

Osborn, T., 1980. Estimates of the local rate of vertical diffusion from dissipation measurements. J. Phys. Oceanogr 10, 83–89.

Peters, H., Baumert, H., 2007. Validating a turbulence closure against estuarine microstructure measurements. Ocean Model. 19, 183–203.

Peters, H., Baumert, H., Jacob, J., 2009. Observations of turbulent mixing in a phytoplankton thin layer: Implications for formation, maintenance, and breakdown. Limnol. Oceanogr. 54, 1353–1368.

Peters, H., Gregg, M., Toole, J., 1988. On the parametrization of equatorial turbulence. J. Geophys. Res. 93, 19–1218.

Petrenko, A., Doglioli, A., Nencioli, F., Kersalé, M., Hu, Z., d'Ovidio, F., 2017. A review of the LATEX project: mesoscale to submesoscale processes in a coastal environment. Ocean Dynamics , 1–21.

Petrenko, A.A., 2003. Variability of circulation features in the Gulf of Lion NW Mediterranean Sea. Importance of inertial currents. Oceanol. Acta 26, 323–338.

Petrenko, A.A., Leredde, Y., Marsaleix, P., 2005. Circulation in a stratified and wind-forced Gulf of Lions, NW Mediterranean Sea: in situ and modeling data). Cont. Shelf Res. 25, 7–27.

Pinazo, C., Marsaleix, P., Millet, B., Estournel, C., Kondrachoff, V., Véhil, R., 2001. Phytoplankton variability in summer in the northwestern Mediterranean: Modelling of the wind and freshwater impacts. J. Coastal Res. 17.

Qiu, Z., Doglioli, A., Hu, Z., Marsaleix, P., Carlotti, F., 2010. The influence of hydrodynamic processes on zooplankton transport and distributions in the North Western Mediterranean: Estimates from a Lagrangian model. Ecol. Model. 221, 2816–2827.

Reffray, G., Bourdalle-Badie, R., Calone, C., 2015. Modelling turbulent vertical mixing sensitivity using a 1-D version of NEMO. Geosci. Model Dev. 8, 69–86.

Reffray, G., Fraunie, F., Marsaleix, P., 2004. Secondary flows induced by wind forcing in the Rhône region of freshwater influence. Ocean Dynamics 54, 179–196.

Rhines, P., 1988. Mixing and large-scale ocean dynamics. Elsevier Oceanography Series 46, 263–284.

Rodi, W., 1980. Turbulence models and their application in hydraulics. International Association for Hydraulic Research, Delft, Netherlands.

Rodi, W., 1987. Examples of calculation methods for flow and mixing in stratified fluids. J. Geophys. Res. 92, 5305–5328.

Rubio, A., Barnier, B., Jorda, G., Espino, M., Marsaleix, P., 2009. Origin and dynamics of mesoscale eddies in the Catalan Sea (NW Mediterranean): Insight from a numerical model study. J. Geophys. Res. 114.

Ruddick, B., Anis, A., Thompson, K., 2000. Maximum likelihood spectral fitting: The Batchelor spectrum. J. Atmos. Oceanic Technol. 17, 1541–2012.

36

Sammari, C., Millot, C., Prieur, L., 1995. Aspects of the seasonal and mesoscale variabilities of the Northern Current inferred from the PROLIG-2 and PROS-6 experiments. Deep-Sea Res. 6, 893–917.

Sanchez, X., Roget, E., Planella, J., Forcat, F., 2011. Small-scale spectrum of a scalar field in water: The Batchelor and Kraichnan models. J. Phys. Oceanogr 41, 2155–2167.

Sharples, J., Coates, M., Sherwood, J., 2003. Quantifying turbulent mixing and oxygen fluxes in a mediterranean-type, microtidal estuary. Ocean Dynam. 53, 126–136.

Sharples, J., Moore, M., 2001. Internal tide dissipation, and vertical nitrate flux at the shelf edge of NE New Zealand. J. Geophys. Res. 106, 14069–14081.

Shih, L., Koseff, J., Ivey, G., Ferziger, J., 2005. Parameterization of turbulent fluxes and scales using homogeneous sheared stably stratified turbulence simulations. J. Fluid Mech. 525, 193–21.

Stacey, M., Pond, S., 1997. On the MellorYamada turbulence closure scheme: The surface boundary condition for q^2. J. Phys. Oceanogr. 27, 2081–2086.

Steinbuck, J., Stacey, M., Monismith, S., 2009. An evaluation of χ_t estimation techniques: Implications for Batchelor fitting and ϵ. J. Atmos. Oceanic Technol. 26, 1652–1662.

Steinbuck, J.V., Genin, A., Monismith, S.G., Koseff, J.R., Holzman, R., Labiosa, R.G., 2010. Turbulent mixing in fine-scale phytoplankton layers: Observations and inferences of layer dynamics. Cont. Shelf Res. 30, 442–455.

Steinbuck, J.V., Koseff, J., Genin, A., Stacey, M., Monismith, S., 2011. Horizontal dispersion of ocean tracers in internal wave shear. J. Geophys. Res. 116.

Taylor, J., Ferrari, R., 2011. Shutdown of turbulent convection as a new criterion for the onset of spring phytoplankton blooms. Limnol. Oceanogr. 56, 2293–2307.

Thomas, W., Gibson, C., 1990. Effects of small-scale turbulence on microalgae. J. Appl. Phycology 2, 71–77.

Thorpe, S., 2005. The turbulent ocean. Cambridge University Press .

Toole, J., Schmitt, R., 1987. Small-scale structures in the north-west Atlantic sub-tropical front. Nature 327, 47–49.

Ulses, C., 2005. Dynamique océanique et transport de la matière particulaire dans le Golfe de Lion: Crue, tempête et période hivernale. PhD thesis, Universitè Paul Sebatier Toulouse .

Ulses, C., Auger, P.A., Soetaert, K., Marsaleix, P., Diaz, F., Coppola, L., Herrmann, M., Kessouri, F., Estournel, C., 2016. Budget of organic carbon in the North-Western Mediterranean Open Sea over the period 2004-2008 using 3-D coupled physical-biogeochemical modeling. J. Geophys. Res. 121.

Warner, J., Sherwood, C., Arango, H., Signell, R.P., 2005. Performance of four turbulence closure models implemented using a generic length scale method. Ocean Model. 8, 81–113.

Wilcox, D., 1988. Reassessment of the scale-determining equation for advanced turbulence models. AIAA Journal 26, 1299–1310.

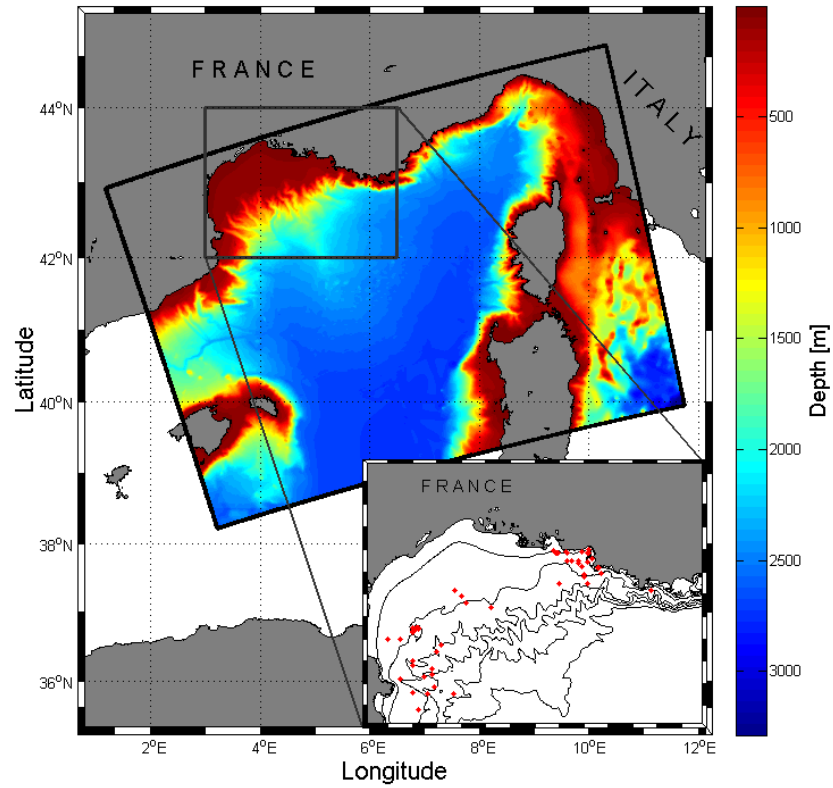Woods, J., 1968. Wave-induced shear instability in the summer thermocline. J. Fluid Mech. 32, 791–800.

Figure 1: Numerical model domain. The color code represents the water depth. The Gulf of Lion is magnified in the smaller box where the measurements sites are represented by red dots. Note that many profiles were taken at the same location over time. The black lines in the smaller box represent the 0, 50, 100, 500, 1000 and 1500 m isobaths.
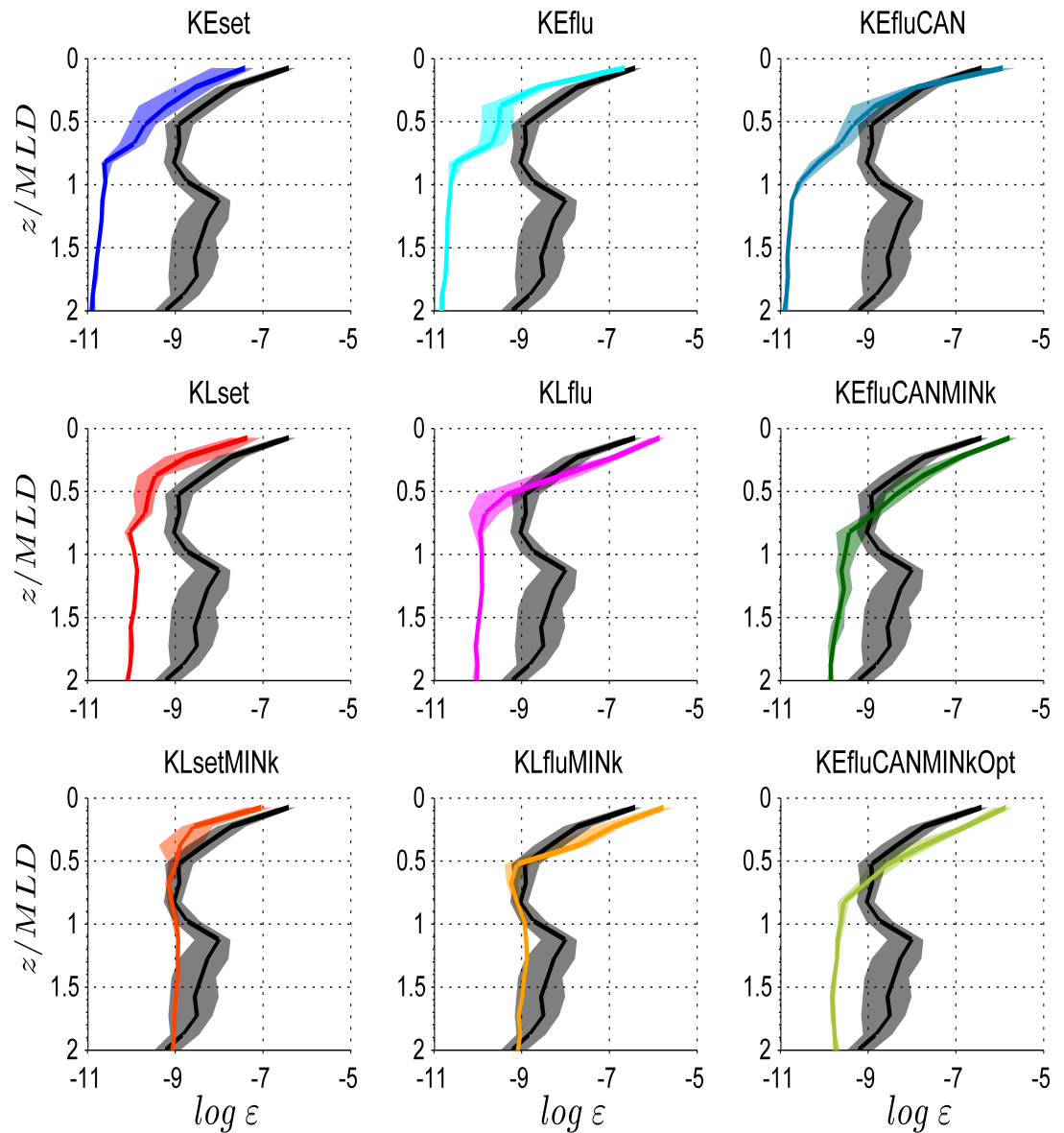
Figure 2: The thick lines represent median values of ε estimated from in situ data (in black) and numerical experiments. The shades indicate 95% bootstrap confidence intervals. Water depth z is non-dimensionalised with respect to the mixed layer depth MLD for each profile.
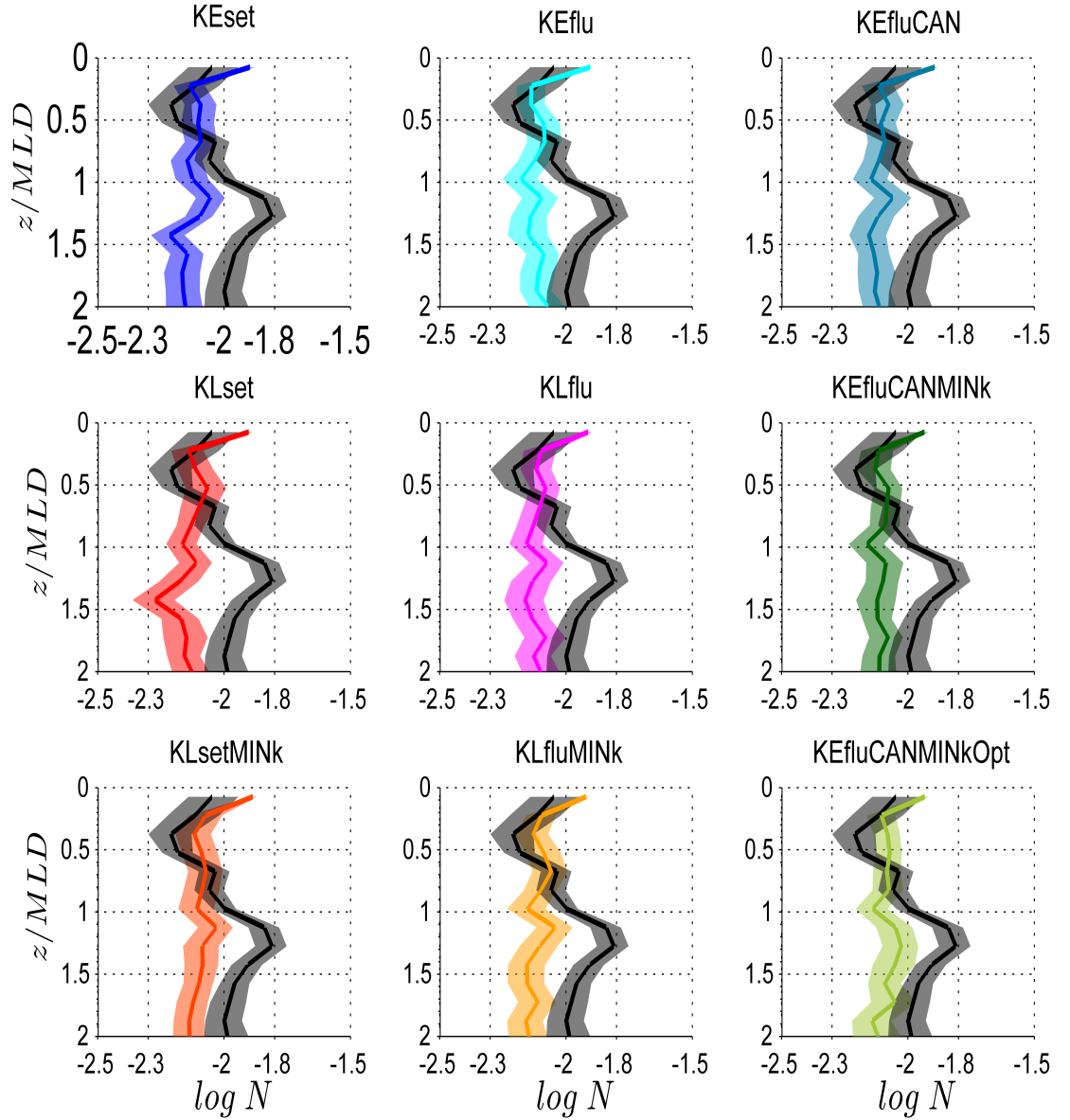
Figure 3: The thick lines represent median values of N estimated from in situ data (in black) and numerical experiments. The shades indicate 95% bootstrap confidence intervals. Water depth z is non-dimensionalised with respect to the mixed layer depth MLD for each profile.
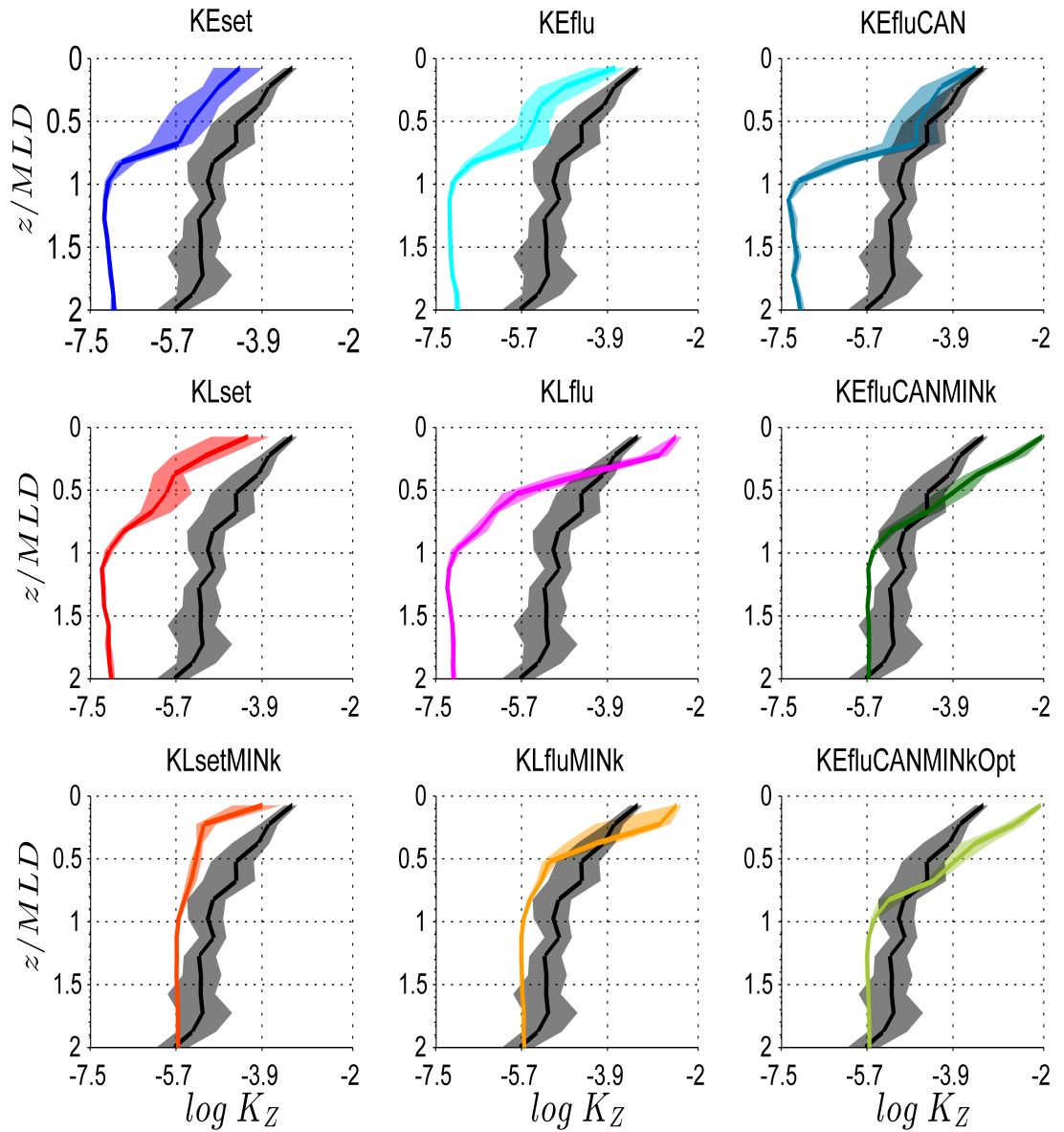
Figure 4: The thick lines represent median values of K_Z estimated from in situ data (in black) and numerical experiments. The shades indicate 95% bootstrap confidence intervals. Water depth z is non-dimensionalised with respect to the mixed layer depth MLD for each profile.
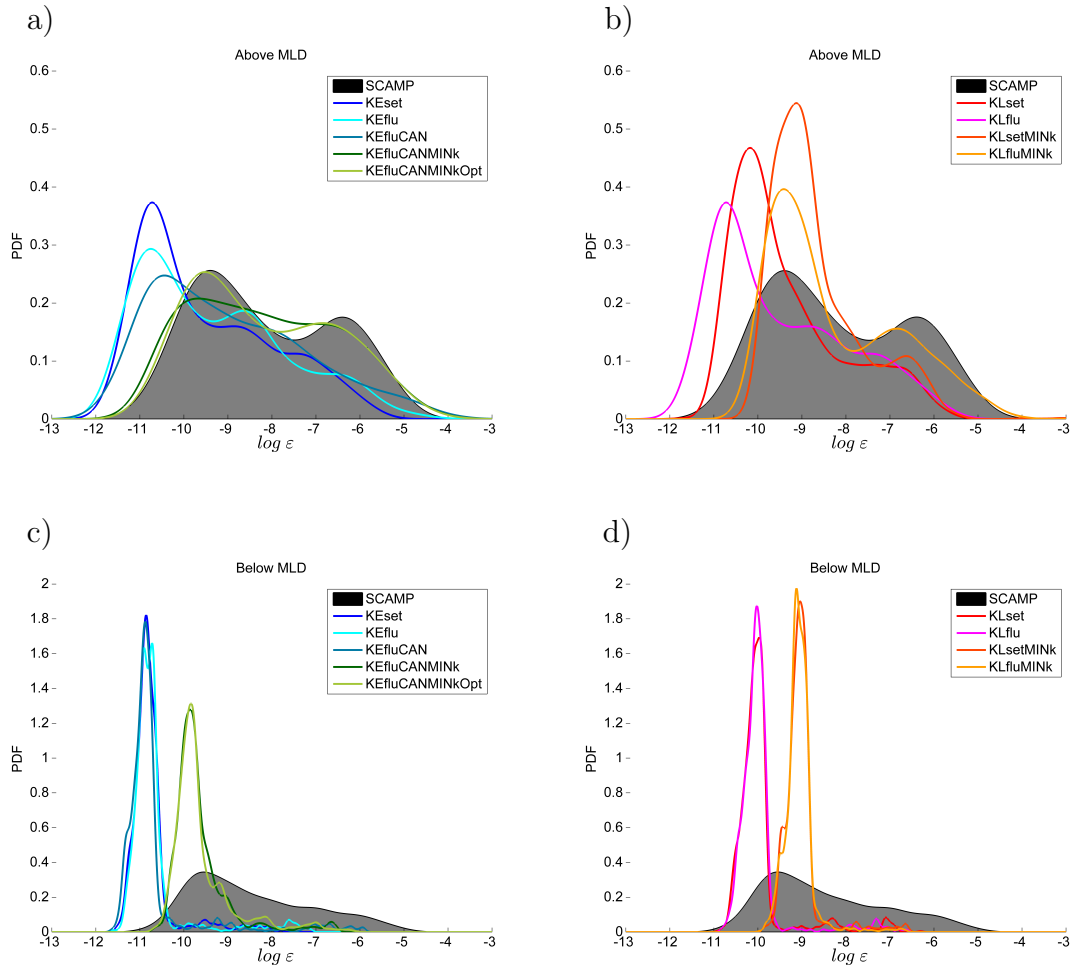
Figure 5: a) and b): ε probability density functions issued from all the in situ (black shade) and numerical data (thick lines) in the surface mixed layer for the KE closure scheme and the KL closure scheme respectively. c) and d): the same but below the surface mixed layer.

Figure 6: a) and b): N probability density functions issued from all the in situ (black shade) and numerical data (thick lines) in the surface mixed layer for the KE closure scheme and the KL closure scheme respectively. c) and d): the same but below the surface mixed layer.
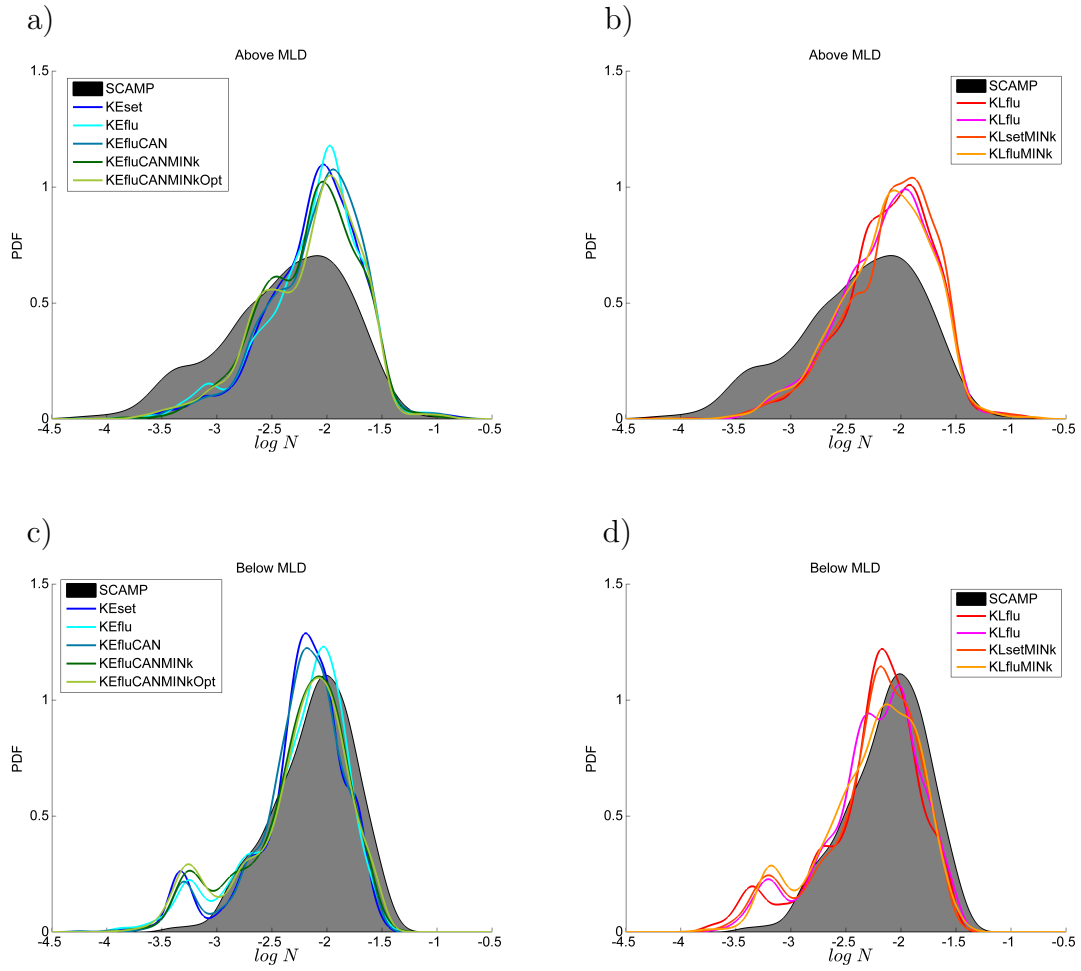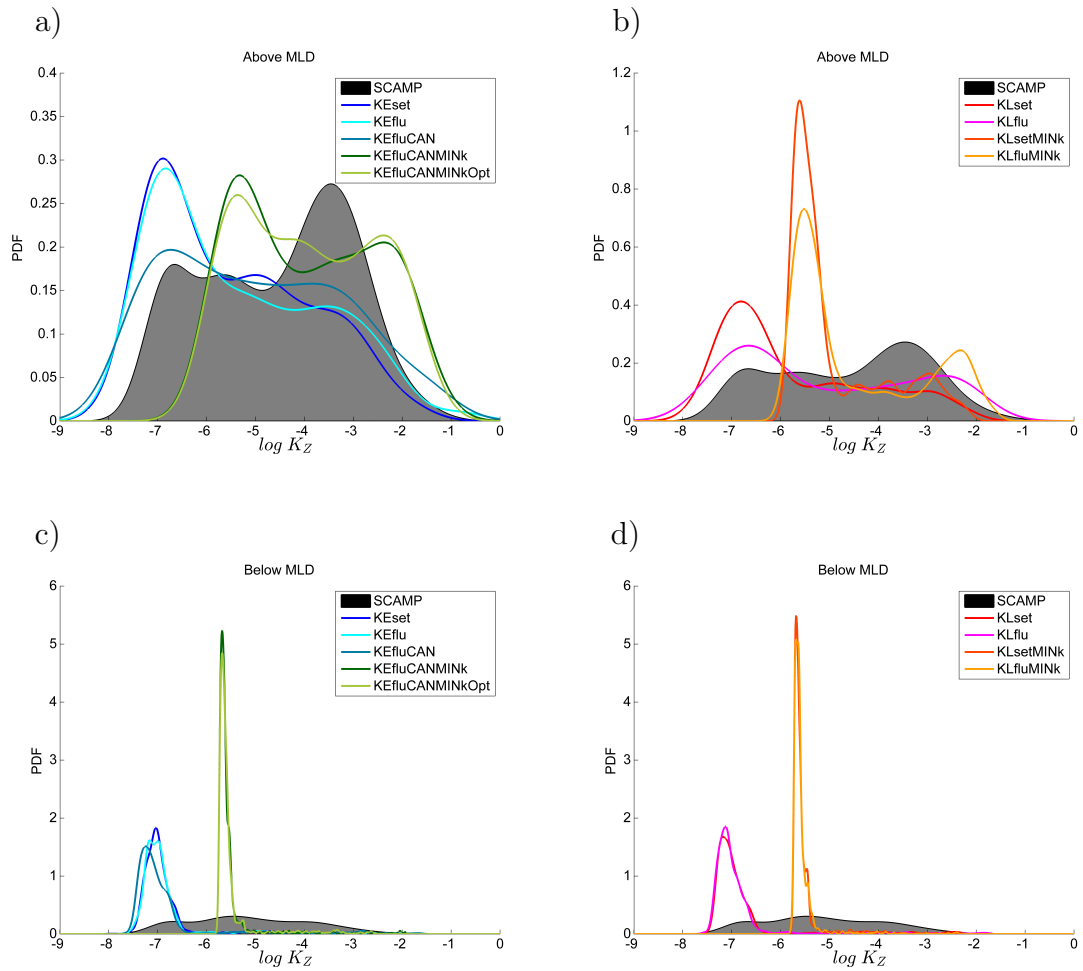
Figure 7: a) and b): K_Z probability density functions issued from all the in situ (black shade) and numerical data (thick lines) in the surface mixed layer for the KE closure scheme and the KL closure scheme respectively. c) and d): the same but below the surface mixed layer. Note the difference in the range of the y axes in a) and b).
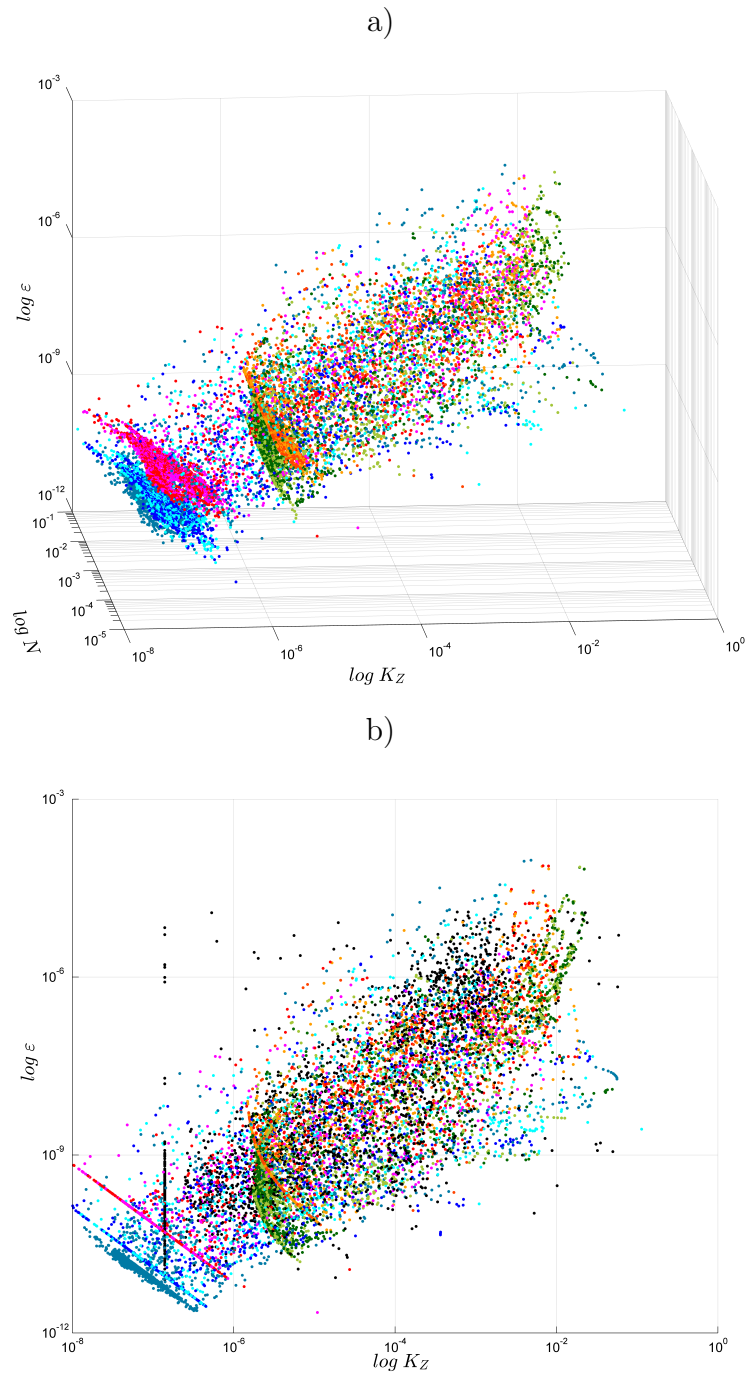
a)

b)

Figure 8: a) All in situ (in black) and numerical data of K_Z, N and ε. b) All in situ (in black) and numerical data of K_Z and ε. The color code is the same as in the other figures.

46
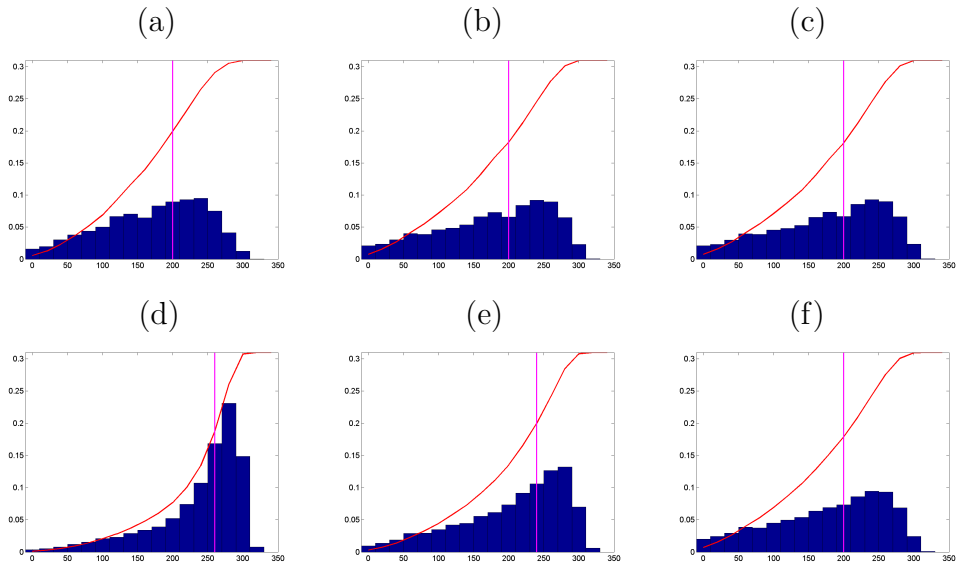
Figure 9: a) JP values distribution obtained with the II91 segmentation method, b) same as a) for S11, c) same as a) for C12, d) same as a) for M12. e) same as a) for a constant 512 points segmentation. f) same as a) for a 1024 points segmentation with 30% overlap. The red curve is the cumulative distribution of JP values for each case and the vertical line marks the point where the cumulative distribution equals 0.6. The database for this analysis comprehended 126 profiles with an average depth of 50m in various meteorological conditions in the Gulf of Lion.

	SMC	Boundary Conditions	k_{min} $[m^2/s^2]$	Stability Functions	l_{PAR}
KEset	$k - \varepsilon$	Equilibrium	10^{-8}	Kantha and Clayson (1994)	11m
KEflu	$k - \varepsilon$	Flux	10^{-8}	Kantha and Clayson (1994)	11m
KEfluCAN	$k - \varepsilon$	Flux	10^{-8}	Canuto et al. (2001)	11m
KEfluCANMINk	$k - \varepsilon$	Flux	10^{-7}	Canuto et al. (2001)	11m
KEfluCANMINkOpt	$k - \varepsilon$	Equilibrium	10^{-7}	Canuto et al. (2001)	23m
KLset	$k - \ell$	Equilibrium	10^{-8}	none	11m
KLflu	$k - \ell$	Flux	10^{-8}	none	11m
KLsetMINk	$k - \ell$	Equilibrium	10^{-7}	none	11m
KLfluMINk	$k - \ell$	Flux	10^{-7}	none	11m

Table 1: Set of numerical experiments performed for this study.

	MLD ratio	σ_ε	σ_N	σ_{K_Z}
KEset	0.91	2.48	1.16	2.19
KEflu	0.91	2.43	1.15	2.16
KEfluCAN	0.85	2.42	1.14	2.11
KEfluCANMINk	0.94	1.88	1.16	1.63
KEfluCANMINkOpt	0.89	1.82	1.16	1.56
KLset	0.88	2.07	1.14	2.20
KLflu	0.87	2.10	1.14	2.21
KLsetMINk	0.85	1.59	1.12	1.46
KLfluMINk	0.86	1.67	1.14	1.54

Table 2: MLD ratio indicates the average ratio between the numerical prediction of the MLD and the in situ value. σ_ε indicates the mean value of the decadal standard deviation of the ε profiles. σ_N indicates the mean value of the decadal standard deviation of the N profiles. σ_{K_Z} indicates the mean value of the decadal standard deviation of the K_Z profiles.

Above MLD	S_ε	Δ^2_ε	S_N	Δ^2_N	S_{K_Z}	$\Delta^2_{K_Z}$
KEset	1.31	0.010±0.002	-0.47	0.018±0.001	0.96	0.032±0.003
KEflu	1.21	0.008±0.001	-0.47	0.015±0.001	0.83	0.039±0.004
KEfluCAN	0.87	0.006±0.001	-0.50	0.016±0.001	0.46	0.021±0.003
KEfluCANMINk	0.62	0.0031±0.0009	-0.45	0.013±0.001	0.49	0.023±0.002
KEfluCANMINkOpt	1.07	0.0008±0.0003	-0.47	0.011±0.001	1.19	0.022±0.003
KLset	0.03	0.012±0.002	-0.41	0.014±0.001	-0.66	0.053±0.005
KLflu	0.34	0.041±0.005	-0.52	0.017±0.001	0.31	0.060±0.005
KLsetMINk	0.07	0.082±0.006	-0.43	0.017±0.001	-0.15	0.139±0.009
KLfluMINk	0.03	0.014±0.003	-0.41	0.014±0.001	-0.68	0.022±0.003

Table 3: S_ε indicates the shift of the numerical distribution of ε in the surface layer respect to the in situ one. Δ^2_ε is the squared difference between the numerical and in situ distribution of ε values in the surface layer. Similar definitions apply for different subscript variables. Errors are calculated with a re-sampling procedure.

Below MLD	S_ε	Δ^2_ε	S_N	Δ^2_N	S_{K_Z}	$\Delta^2_{K_Z}$
KEset	1.97	0.180±0.006	0.22	0.004±0.0005	0.96	0.032±0.004
KEflu	1.93	0.163±0.008	0.19	0.005±0.0006	0.83	0.039±0.004
KEfluCAN	1.98	0.143±0.008	0.22	0.002±0.0003	0.46	0.021±0.004
KEfluCANMINk	0.91	0.136±0.007	0.22	0.006±0.0006	-0.66	0.023±0.003
KEfluCANMINkOpt	0.88	0.11±0.006	0.22	0.008±0.0009	-0.68	0.022±0.002
KLset	1.21	0.186±0.008	0.22	0.003±0.0005	0.50	0.053±0.004
KLflu	1.2	0.168±0.007	0.23	0.004±0.0005	1.19	0.060±0.005
KLsetMINk	0.26	0.203±0.006	0.22	0.004±0.0006	0.31	0.080±0.007
KLfluMINk	0.29	0.22±0.004	0.24	0.008±0.0009	-0.15	0.022±0.002

Table 4: S_ε indicates the shift of the numerical distribution of ε below the surface layer respect to the in situ one. Δ^2_ε is the squared difference between the numerical and in situ distribution of ε values below the surface layer. Similar definitions apply for different subscript variables. Errors are calculated with a re-sampling procedure.