Derivation and implementation of a non-gradient term to improve the oceanic convection representation within the k-e parameterization

Alexandre Legay[1], Bruno Deremble[1], and Hans Burchard[2]

[1]Institut des Géosciences de l’Environnement
[2]Leibniz Institute for Baltic Sea Research Warnemünde

February 16, 2024

Abstract

The representation of turbulent fluxes during oceanic convective events is important to capture the evolution of the oceanic mixed layer. To improve the accuracy of turbulent fluxes, we examine the possibility of adding a non-gradient component in their expression in addition to the usual downgradient part. To do so, we extend the $k-\varepsilon$ algebraic second-moment closure by relaxing the assumption on the equilibrium of the temperature variance $\overline{\theta’ˆ2}$. With this additional transport equation for the temperature variance, we obtain a $k - \varepsilon - \overline{\theta’ˆ2}$ model (the “$k \varepsilon t$” model) which includes a non-gradient term for the temperature flux. We validate this new model against Large Eddy Simulations (LES) in both wind-forced and buoyancy-driven regimes. In both cases, we find that the vertical profile of temperature is well captured by the $k \varepsilon t$ model. Particularly, for the buoyancy-driven regime, the non-gradient term increases the portion of the mixed layer that is stably stratified. This is an improvement since this portion is too small with the $k - \varepsilon$ parameterization. Finally, a comparison of the non-gradient term with the KPP non-local term gives insights for refining the KPP’s ad hoc shape polynomial.

Derivation and implementation of a non-gradient term to improve the oceanic convection representation within the $k - \varepsilon$ parameterization

Alexandre Legay[1], Bruno Deremble[1], Hans Burchard[2]

[1]Univ. Grenoble Alpes, CNRS, INRAE, IRD, Grenoble INP, IGE, Grenoble, France
[2]Leibniz Institute for Baltic Sea Research Warnemünde, Rostock, Germany

Key Points:

- Analytical derivation of a non-gradient term within the $k-\varepsilon$ model for improving the oceanic convection representation
- Comparison with large eddy simulations in both wind-forced and buoyancy-driven regimes confirms the improvement due to the non-gradient term
- The vertical profile of the non-gradient term is compared to the one of the KPP the non-local term

Corresponding author: Alexandre Legay, `alexandre.legay@univ-grenoble-alpes.fr`

Abstract

The representation of turbulent fluxes during oceanic convective events is important to capture the evolution of the oceanic mixed layer. To improve the accuracy of turbulent fluxes, we examine the possibility of adding a non-gradient component in their expression in addition to the usual downgradient part. To do so, we extend the $k-\varepsilon$ algebraic second-moment closure by relaxing the assumption on the equilibrium of the temperature variance $\overline{\theta'^2}$. With this additional transport equation for the temperature variance, we obtain a k–ε–$\overline{\theta'^2}$ model (the "$k\varepsilon t$" model) which includes a non-gradient term for the temperature flux. We validate this new model against Large Eddy Simulations (LES) in both wind-forced and buoyancy-driven regimes. In both cases, we find that the vertical profile of temperature is well captured by the $k\varepsilon t$ model. Particularly, for the buoyancy-driven regime, the non-gradient term increases the portion of the mixed layer which is stably stratified. This is an improvement since this portion is too small with the k–ε parameterization. Finally, comparison of the non-gradient term with the KPP non-local term gives insights for refining the KPP's ad hoc shape polynomial.

Plain Language Summary

In the ocean, vertical mixing of water occurs when cold air temperatures create dense cold water at the surface that tends to sink in the ocean or when a strong wind induces turbulence at the ocean surface. In numerical models, the classic approach to represent this vertical mixing is to consider that it is done entirely by diffusion. This means that the heat always goes from the warm water to the cold water, i.e. in the opposite direction of the gradient of the temperature. However, during intense events called "convection", some cold water parcels created at the top of the ocean can have enough thermal inertia and velocity to flow against the direction of the mean temperature gradient. This kind of phenomenon is often referred to coherent eddies or non-local turbulence. In this article, we perform an analytical derivation to give a mathematical expression of the impact of non-local mixing. We then compare our new model with more realistic three-dimensional models of convection and conclude that the new term derived here is important to reproduce the vertical profile of temperature in the ocean.

1 Introduction

In the realm of climate modeling, the oceanic mixed layer plays a critical role because it is responsible for regulating the oceanic heat uptake and carbon storage. Throughout much of the year, the mixed layer operates as a dynamic buffer, intimately interacting with the atmosphere. However, it is in late winter that the true importance of this layer becomes evident. In late winter, the mixed layer is deepest and direct contact is established with the deep ocean: it is during this period that the ocean effectively stores heat and CO2 (a mechanism sometimes pictured as Stommel's demon, see Luyten et al., 1983; Williams et al., 1995). Accurately representing the mixed layer is thus crucial because it directly affects our ability to make accurate predictions about future climate patterns (Treguier et al., 2023).

The depth of the mixed layer changes in response to various factors: it deepens when turbulent mixing is triggered by the mechanical effect of the wind and/or waves; or triggered by buoyancy effects: heat flux (cooling) and freshwater flux (evaporation, sea-ice formation). Conversely, the mixed layer becomes shallower typically during calm weather where there is less turbulence and restratifying mixed layer instabilities can develop, or when there is a stabilizing buoyancy flux due to warming (e.g. sunny condition) and/or freshwater input (e.g. precipitation, sea-ice melt, or river discharge). This restratification allows the surface layer to separate from the denser, deeper water (Stull, 1988). The explicit representation of the small-scale turbulence causing the mixing occurring in the mixed layer is of course impossible in climate models where the horizontal grid is often on the order of tens of kilometers. Instead, the ocean modeling community has developed parameterizations whose goal is to represent the mean effect of the turbulent fluctuations (Gaspar et al., 1990; Large et al., 1994; Burchard & Bolding, 2001; Umlauf & Burchard, 2003; Fox-Kemper et al., 2008; Reichl & Hallberg, 2018). The main purpose of a mixed layer parameterization is to propose a closure for the turbulent vertical fluxes $\overline{w'x'}$, where w' is the turbulent vertical velocity, x' the turbulent fluctuation of a property x (momentum, temperature, salinity, phytoplankton, etc...) and the overline denotes the ensemble averaging over small-scale fluctuations (see Stull, 1988). These turbulent fluxes, and all the other covariances $\overline{x'y'}$, are called the second-order moments. The traditional approach to close this problem consists of expressing these turbulent fluxes as a function of the vertical gradient of the mean property $X = \overline{x}$ (i.e. a downgradient parameterization), as shown here for the temperature

$$\overline{w'\theta'} = -K_t \partial_z \Theta \,, \tag{1}$$

with K_t being an eddy diffusivity coefficient. Among all the possibilities to compute K_t we would like to emphasize the Generic Length Scale (GLS) approach (Umlauf & Burchard, 2003) and more precisely the $k-\varepsilon$ closure (Burchard & Bolding, 2001). This closure consists in deriving two equations: one for the evolution of turbulent kinetic energy k, and one for dissipation ε. The downgradient formulation (1) also results from more complex algebraic second-moment closures even if it is not assumed a priori (Burchard & Baumert, 1995). The eddy diffusivity is obtained analytically and is a function of turbulent kinetic energy, dissipation, buoyancy frequency, and shear frequency. While this eddy diffusivity approach has been successfully applied in the oceanic and atmospheric modeling communities, it has also been quickly recognized that the shape of the temperature profile during a convective event is not well captured by this closure. In fact, Deardorff (1972) was among the first to realize that after a convective event, the stratification profile in the mixed layer is not neutral as one would expect for a perfectly well-mixed layer but is instead slightly stable. To illustrate this observation, we plot in Figure 1.a the typical shape of a normalized temperature profile in the mixed layer from a numerical model that explicitly resolves convection (see Mironov et al. (2000); details about the normalization are provided henceforth; we only wish to focus here on the shape of the temperature profile). This profile can be decomposed into two well-defined zones. Just below the air-sea interface, there is an unstable zone with cold water above warmer water ($\partial_z \Theta < 0$). Such layer is sometimes called the *thermal layer* (Lazier, 2001) and we define it here as the layer between the surface and the depth h_t at which $\partial_z \Theta = 0$. Below that depth h_t, we find the *convective layer*; a slightly stable layer that extends until the base of the mixed layer h_m. Both layers form the *mixed layer*. The position of h_t has been documented to be near $z = -0.4 h_m$ (see Zhou et al., 2018) such that more than half of the mixed layer is stably stratified. The presence of such stable stratification in the convective layer has been attributed to downward propagating plumes which remain coherent during their descent and deposit their negative buoyancy anomaly at their neutral level, thus creating a stable stratification (see Arakawa and Schubert (1974) or Emanuel (1991) for the atmospheric scenario).

Several options have emerged in the literature to reproduce this vertical temperature profile with a stable stratification. The atmospheric community has favored the

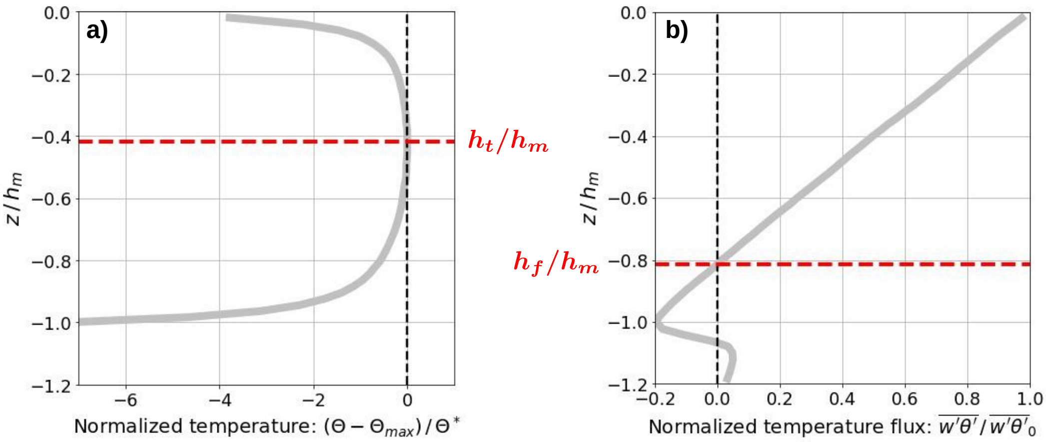

Figure 1. Normalized profile from LES data of Mironov et al. (2000) of a) the temperature and b) the vertical turbulent temperature flux. The depth is normalized by the mixed layer depth h_m, defined here as the minimum of the temperature flux. The temperature flux is normalized by its surface value $\overline{w'\theta'}|_0$. The temperature is normalized in $(\Theta - \Theta_{max})\,/\,\Theta^*$ with Θ_{max} the maximum of the temperature over the vertical and $\Theta^* = \overline{w'\theta'}|_0/w^*$ a scaling of the temperature, with $w^* = (\overline{w'\theta'}|_0\, h_m)^{1/3}$ a scaling of the velocity of the convective thermals (Willis & Deardorff, 1974; Marshall & Schott, 1999). Red dashed lines highlight the location h_t of the zero of the gradient $\partial_z\Theta$ and the location h_f of the zero of the temperature flux.

use of a mass flux parameterization which simulates the vertical movement of air parcels within convective clouds. It represents the ascent and descent of parcels, which transport heat, moisture, and other properties. These mass flux parameterizations have recently been introduced in ocean models (Giordani et al., 2020; Garanaik et al., 2024). Another, perhaps more ancient, approach taken by Large et al. (1994) was to add a positive non-gradient term $\mathit{\Gamma}$ in the parameterization of the flux in Equation (1): (see also Troen and Mahrt (1986); or Burchard and Petersen (1999) where the problem of missing non-gradient fluxes in downgradient parameterization is stated),

$$\overline{w'\theta'} = -K_t \partial_z \mathit{\Theta} + \mathit{\Gamma}. \tag{2}$$

$\mathit{\Gamma}$ being positive, it represents a positive turbulent temperature flux, i.e. a flux that follows the buoyancy effect (cold going down and hot going up). $\mathit{\Gamma}$ can thus be viewed as representing coherent structures (”non-local eddies”, ”coherent thermals”) that are subjected to the buoyancy force. Particularly, we see in equation 2 that $\mathit{\Gamma}$ allows to keep a positive turbulent temperature flux in situations of neutral ($\partial_z \mathit{\Theta} = 0$) or slightly stable ($\partial_z \mathit{\Theta} > 0$) temperature profiles. In other words, this means that, in stably-stratified conditions, coherent structures can be strong enough to counter the downgradient flux that acts in a counter-buoyancy direction. Note that this term was often written $\overline{w'\theta'} = -K_t(\partial_z \mathit{\Theta} - \gamma)$ with $\gamma = \mathit{\Gamma}/K_t$ (e.g. Deardorff, 1972; Large et al., 1994). In this formulation, γ corresponds to the maximal stable stratification where a positive turbulent temperature flux can be maintained even if the downgradient flux generates a counter-buoyancy effect. In Large et al. (1994), $\mathit{\Gamma}$ was defined with some constraints: to be zero at the surface and at the base of the mixed layer such that it is merely a redistribution of heat. The magnitude and the exact shape of this term were however chosen in a relatively ad hoc way to respect some empirical rules of convection.

The term $\mathit{\Gamma}$ was often referred in the literature as a ”non-local” term (Large et al., 1994; Ghannam et al., 2017) or a countergradient term (Deardorff, 1972; Troen & Mahrt, 1986; Gibbs et al., 2011). As we mentioned before, denomination ”non-local” refers to the fact that it is supposed to represent non-local eddies (coherent thermals). However Zhou et al. (2018) argued that this often-implied association of the non-gradient term to the non-local eddies is partially wrong. ”Non-local” can also indicate that the value of this term at a specific depth does not depend exclusively on properties evaluated at this depth. For example, in KPP, this term depends on the surface heat flux and on the total mixed layer thickness. The other used denomination, "countergradient", refers to the fact that, in the lower part of the mixed layer which is stable, this term acts with an opposite sign compared to the mean gradient. However, in the upper part of the mixed layer which is unstable, the denomination "countergradient" is very unsettling since this term acts as if it were a downgradient term. For these reasons, we will call this term "non-gradient", a more neutral denomination.

A key aspect of the addition of the non-gradient term is to relax the downgradient dependence and particularly the constraint that the depth at which $\overline{w'\theta'}$ vanishes is equal to the depth at which the gradient of the temperature profile vanishes (see Eq. 1). To better understand why this matters, we plot in Figure 1.b the vertical turbulent heat flux $\overline{w'\theta'}$ obtained in the same numerical model as presented before (Mironov et al., 2000). In this figure, we recover the traditional form of a linear decrease from the surface value (which corresponds to the magnitude of the surface flux) to a cancellation near the bottom of the mixed layer, which has been observed and described in several places (e.g. Large et al., 1994; Burchard & Bolding, 2001; Van Roekel et al., 2018). The exact depth at which the heat flux vanishes depends on the surface boundary conditions (wind and heat fluxes) but it has been documented to be close to $h_f = -0.8h_m$ at least in the free convection scenario (Garcia & Mellado, 2014). There is thus an obvious discrepancy between $h_t = -0.4h_m$ and $h_f = -0.8h_m$ such that Equation (1) cannot hold in most of the mixed layer and the addition of an extra term in the definition of the flux is physically relevant. Even if there is a consensus on the need to add a non-gradient component in the definition of the flux, the exact formulation of this flux remains a matter of debate. To develop a framework that is accurate, robust, and consistent with existing parameterizations, we have opted to focus on extending the $k - \varepsilon$ parameterization.

We first perform an analytical derivation of the non-gradient term. Since Deardorff (1972) and Cheng et al. (2020), we know that the non-gradient term is somehow related to the small-scale temperature variance $\overline{\theta'^2}$. We will therefore derive a second-moment closure that uses a full transport equation for the temperature variance $\overline{\theta'^2}$, in addition to the second-moment transport equations for k and ε, thus extending the $k-\varepsilon$ model to a $k - \varepsilon - \overline{\theta'^2}$ model (henceforth called the "$k\varepsilon t$" model). In this model, we get an analytical expression of a non-gradient term that shares several properties with the KPP non-local term: it is positive, and vanishes at the surface and at the bottom of the mixed layer. Last, we test the numerical implementation of $k\varepsilon t$ against Large Eddy Simulations (LES) and further compare its results to the predictions of a standard $k-\varepsilon$ model and KPP simulations.

2 Derivation and Implementation of the $k\varepsilon t$ Parameterization

This section introduces the second-order moments equations. We recall the hypotheses made in the GLS model to solve this system of equations. Then, we explain how we derive the $k\varepsilon t$ parameterization in the same formalism.

2.1 Formalism and Second-Order Moments Equations

The Reynolds Averaged Navier Stokes (RANS) equations used in ocean models are written for the mean velocities $\boldsymbol{U} = (U, V, W)$ and the mean temperature Θ. As in the original derivation of the $k-\varepsilon$ model, we consider here only one active tracer (temperature) that enters the equation of state. The RANS equations include the effect of turbulent fluctuations through the second-order moments $\overline{u_i'u_j'}$ and $\overline{u_i'\theta'}$. To close the system, we need to provide equations for these moments. We focus here on the procedure derived in Burchard and Bolding (2001). After adopting their closure assumptions for non-closed terms, and neglecting the rotational and viscous effects, the equations of second-order moments are

$$
\begin{aligned}
\partial_t \overline{u_i'u_j'} + \partial_l (U_l \overline{u_i'u_j'} + \overline{u_i'u_j'u_l'}) = & - c_1 \frac{\varepsilon}{k} (\overline{u_i'u_j'} - \frac{2}{3}\delta_{ij} k) \\
& + P_{ij} - c_2 (P_{ij} - \frac{2}{3}\delta_{ij} P) \\
& + B_{ij} - c_3 (B_{ij} - \frac{2}{3}\delta_{ij} B) \\
& - c_4 k S_{ij} \\
& - c_5 Z_{ij} \\
& - \frac{2}{3}\delta_{ij}\varepsilon,
\end{aligned}
\tag{3}
$$

$$
\begin{aligned}
\partial_t \overline{u_i'\theta'} + \partial_j (U_j \overline{u_i'\theta'} + \overline{u_i'u_j'\theta'}) = & - c_{1T} \frac{\varepsilon}{k} \overline{u_i'\theta'} \\
& - (1 - c_{2T})\, \overline{u_j'\theta'}\, \partial_j U_i - \overline{u_i'u_j'}\, \partial_j \Theta \\
& + (1 - c_{3T})\, \beta_i\, \overline{\theta'^2} \\
& + c_{4T}\, \overline{u_j'\theta'}\, V_{ij},
\end{aligned}
\tag{4}
$$

$$\partial_t \overline{\theta'^2} + \partial_j (U_j \overline{\theta'^2} + \overline{u'_j \theta'^2}) = -2\, \overline{u'_j \theta'}\, \partial_j \Theta - 2 \frac{1}{c_T} \frac{\varepsilon}{k}\, \overline{\theta'^2}, \tag{5}$$

with

- $P_{ij} = -\partial_l U_i\, \overline{u'_l u'_j} - \partial_l U_j\, \overline{u'_l u'_i}$: Production/destruction of $\overline{u'_i u'_j}$ by the shear
- $B_{ij} = \beta_i\, \overline{u'_j \theta'} + \beta_j\, \overline{u'_i \theta'}$: Production of $\overline{u'_i u'_j}$ by the buoyancy
- $S_{ij} = \frac{1}{2}(\partial_i U_j + \partial_j U_i)$: Shear tensor
- $V_{ij} = \frac{1}{2}(\partial_i U_j - \partial_j U_i)$: Vorticity tensor
- $Z_{ij} = V_{il}(\overline{u'_l u'_j} - \frac{2}{3}\delta_{lj} k) + V_{jl}(\overline{u'_l u'_i} - \frac{2}{3}\delta_{li} k)$: Symmetric tensor associated to the vorticity
- $k = \frac{1}{2}(\overline{u'^2} + \overline{v'^2} + \overline{w'^2})$: Turbulent Kinetic Energy (TKE)
- $P = \frac{1}{2} P_{ii}$: Production of TKE by the shear
- $B = \frac{1}{2} B_{ii}$: Production/destruction of TKE by the buoyancy
- ε: Dissipation of TKE

Further definitions are δ_{ij} the Kronecker delta, $\boldsymbol{\beta} = (0, 0, \alpha g)$, α the thermal expansion coefficient and g the gravitational acceleration. In the equations, the Einstein summation convention is adopted.

Coefficients c_1, c_2, c_3, c_4, c_5 are empirical coefficients for the parameterization of the pressure-velocity correlation tensor $\Pi_{ij} = \overline{u'_i \partial_j p} + \overline{u'_j \partial_i p}$, coefficients $c_{1T}, c_{2T}, c_{3T}, c_{4T}$ are for the parameterization of the pressure-temperature correlations $\Pi_i^\theta = \overline{\theta' \partial_i p}$, and c_T for the parameterization of the temperature variance dissipation. Further details about these parameterizations can be found in Canuto et al. (2001). We report the values of these coefficients in Table 1. These values are the ones of Canuto et al. (2001) model A, converted into the notations used here (it is the same as the values reported in Table 1 of Burchard and Bolding (2001) except for minor typos on c_3 and c_4 that have been identified. Exact formulations of these coefficients are given in Appendix B).

We are now going to explain the classic procedure used in the GLS models for solving the system, where the new model differs and what are the consequences.

Table 1. Values of the coefficients appearing in the second-order moment equations

c_1	c_2	c_3	c_4	c_5	c_{1T}	c_{2T}	c_{3T}	c_{4T}	c_T
2.5	0.984	0.5	0.512	0.416	5.95	0.6	0.33	0.4	1.44

2.2 GLS Procedure

The GLS procedure is as follows. Firstly, we consider the boundary layer approximation where the vertical scale is much less than the horizontal scale. Horizontal gradients are then neglected in comparison to the vertical gradients. A direct consequence is the simplification of the continuity equation in $\partial_z W = 0$. The resulting expressions of the tensors P_{ij}, B_{ij}, S_{ij}, V_{ij} and Z_{ij} are given in Appendix C.

Secondly, we consider that the moments $\overline{u_i'\theta'}$ and $\overline{\theta'^2}$ are in local equilibrium, meaning that the sum of the time variations, the advective transports and the turbulent transports of these moments is zero (i.e. the left-hand sides of equations (4) and (5) are zero). Concerning the moments $\overline{u_i'u_j'}$, the trick is to not make this assumption directly for $\overline{u_i'u_j'}$ but rather to the anisotropic part of these moments $\overline{u_i'u_j'}-2/3\,\delta_{ij}k$ to keep the time variation and the transports of the TKE to be non-zero. These assumptions correspond to the level $2\frac{1}{2}$ in the hierarchy of models proposed by Mellor and Yamada (1982). This hierarchy has been derived with scaling arguments based on the level of anisotropy of every term. The scaling at level 3 results naturally in neglecting transports and time variations for $\overline{u_i'u_j'} - 2/3\,\delta_{ij}k$ and $\overline{u_i'\theta'}$. However, neglecting these terms for the $\overline{\theta'^2}$ equation is not justified by the scaling process and is much more an ad hoc practical hypothesis that results in obtaining this so-called level $2\frac{1}{2}$ in which the system of equations is now algebraic. Indeed, we obtain the following set of equations

$$0 = -c_1\frac{\varepsilon}{k}(\overline{u_i'u_j'}-\frac{2}{3}\delta_{ij}k)+(1-c_2)(P_{ij}-\frac{2}{3}\delta_{ij}P)+(1-c_3)(B_{ij}-\frac{2}{3}\delta_{ij}B)-c_4kS_{ij}-c_5Z_{ij}, \quad (6)$$

$$0 = -c_{1T}\frac{\varepsilon}{k}\overline{u_i'\theta'} - (1-c_{2T})\,\overline{u_j'\theta'}\,\partial_j U_i - \overline{u_i'u_j'}\,\partial_j\Theta + (1-c_{3T})\,\beta_i\,\overline{\theta'^2} + c_{4T}\,\overline{u_j'\theta'}\,V_{ij}, \quad (7)$$

$$0 = -2\,\overline{u_j'\theta'}\,\partial_j\Theta - \frac{2}{c_T}\frac{\varepsilon}{k}\,\overline{\theta'^2}, \quad (8)$$

and, if we assume that k and ε are known, we have a linear system of 10 equations with 10 unknowns : ($\overline{u'^2}$, $\overline{v'^2}$, $\overline{w'^2}$, $\overline{u'v'}$, $\overline{u'w'}$, $\overline{v'w'}$, $\overline{u'\theta'}$, $\overline{v'\theta'}$, $\overline{w'\theta'}$, $\overline{\theta'^2}$). For clarity, these 10 equations are written explicitly in Appendix D. We solved this system thanks to the symbolic calculus software Mathematica and we confirmed the expressions obtained by Burchard and Bolding (2001):

$$\overline{u'w'} = -c_\mu \frac{k^2}{\varepsilon} \partial_z U, \tag{9}$$

$$\overline{v'w'} = -c_\mu \frac{k^2}{\varepsilon} \partial_z V, \tag{10}$$

$$\overline{w'\theta'} = -c'_\mu \frac{k^2}{\varepsilon} \partial_z \Theta, \tag{11}$$

which reflect downgradient fluxes with an eddy viscosity $K_m = c_\mu \frac{k^2}{\varepsilon}$ and an eddy diffusivity $K_t = c'_\mu \frac{k^2}{\varepsilon}$. The dimensionless functions c_μ and c'_μ are the so-called "stability functions" and can be expressed in the following forms

$$c_\mu = \frac{n_0 + n_1\alpha_N + n_2\alpha_M}{d_0 + d_1\alpha_N + d_2\alpha_M + d_3\alpha_N\alpha_M + d_4\alpha_N^2 + d_5\alpha_M^2}, \tag{12}$$

$$c'_\mu = \frac{n_{0T} + n_{1T}\alpha_N + n_{2T}\alpha_M}{d_0 + d_1\alpha_N + d_2\alpha_M + d_3\alpha_N\alpha_M + d_4\alpha_N^2 + d_5\alpha_M^2}, \tag{13}$$

with $\alpha_N = \frac{k^2}{\varepsilon^2}N^2$, $\alpha_M = \frac{k^2}{\varepsilon^2}M^2$, $N^2 = -g/\rho_0\,\partial_z\rho$ the (squared) buoyancy frequency, and $M^2 = (\partial_z U)^2 + (\partial_z V)^2$ the (squared) shear frequency. Coefficients n_i, n_{iT} and d_i depend on the coefficients c_i and c_{iT}. Their full expressions are given in Appendix E. Taking the values of the c_i and c_{iT} given in Table 1, the stability functions are approximately as follows

$$c_\mu = \frac{0.1067 + 0.01732\alpha_N - 0.0001205\alpha_M}{1 + 0.2398\alpha_N + 0.02872\alpha_M + 0.005154\alpha_N\alpha_M + 0.006930\alpha_N^2 - 0.00003372\alpha_M^2}, \tag{14}$$

$$c'_\mu = \frac{0.1120 + 0.003766\alpha_N + 0.0008871\alpha_M}{1 + 0.2398\alpha_N + 0.02872\alpha_M + 0.005154\alpha_N\alpha_M + 0.006930\alpha_N^2 - 0.00003372\alpha_M^2}. \tag{15}$$

To compute the fluxes in Eqs. (9) - (11), we still need to know k and ε. In a GLS model, we solve two prognostic equations, one for k and one for another variable that can be linked to ε. The choice of this second equation is the main difference between the different GLS models ($k{-}\varepsilon$: Hanjalić and Launder (1972); Rodi (1987), $k{-}kl$: Mellor and Yamada (1982), $k-\omega$: Wilcox (1988), $k-\tau$: Zeierman and Wolfshtein (1986); Thangam et al. (1992)). In this paper, we focus on the $k{-}\varepsilon$ model which solves directly the equation for ε. The TKE equation and the ε equation are as follows

$$D_t k = P + G - \varepsilon + \mathscr{D}_k, \tag{16}$$

$$D_t \varepsilon = \frac{\varepsilon}{k}(c_{\varepsilon 1} P + c_{\varepsilon 3} G - c_{\varepsilon 2}\varepsilon) + \mathscr{D}_\varepsilon, \tag{17}$$

with

- $D_t(\cdot) = [\partial_t + U\partial_x + V\partial_y](\cdot)$: Total derivative
- $\mathscr{D}_k = \partial_z(\frac{K_m}{\sigma_k}\partial_z k)$ and $\mathscr{D}_\varepsilon = \partial_z(\frac{K_m}{\sigma_\varepsilon}\partial_z \varepsilon)$: Diffusion terms
- σ_k and σ_ε: Schmidt numbers for TKE and dissipation
- $P \equiv (-\overline{u'w'}\,\partial_z U - \overline{v'w'}\,\partial_z V) = c_\mu \alpha_M \varepsilon$: Production of TKE by the shear
- $G \equiv \beta_3\,\overline{w'\theta'} = -c'_\mu \alpha_N \varepsilon + c'^{*}_\mu \alpha_T \varepsilon$: Production/destruction of TKE by the buoyancy
- $c_{\varepsilon 1}$, $c_{\varepsilon 2}$ and $c_{\varepsilon 3}$: Empirical coefficients

The TKE equation (16) was obtained by taking the trace of the Reynolds stress equations (3). With the boundary layer approximation which neglects the horizontal gradient in comparison to the vertical ones, taking this trace gives $D_t k + \frac{1}{2}(\partial_z \overline{w' u'_i u'_i}) = P + G - \varepsilon$. We then consider downgradient formulations for the third-order moments $\overline{w' u'_i u'_i}$ and finally results in Equation (16). We want to highlight that the diffusion term thus comes from the divergence of the third-order moments.

An exact equation for ε can be derived but, in practice, this equation needs drastic assumptions to be closed. We used in Equation (17) the classic assumptions of scaling the sources and sinks of ε with the ones of the TKE through empirical coefficients $c_{\varepsilon 1}$, $c_{\varepsilon 2}$ and $c_{\varepsilon 3}$ (see Burchard & Bolding, 2001).

Values $\sigma_k = 1$, $c_{\varepsilon 1} = 1.44$ and $c_{\varepsilon 2} = 1.92$ are frequently used in the literature (Rodi, 1987). Value $\sigma_\varepsilon = 1.20$ is found according to (14) of Umlauf and Burchard (2003). Finally, for $c_{\varepsilon 3}$, it is often considered two different values in order to keep $c_{\varepsilon 3} G$ always as a source term of ε (Rodi, 1987; Burchard & Bolding, 2001; Umlauf & Burchard, 2003; Warner et al., 2005; Reffray et al., 2015). A positive value $c_{\varepsilon 3}^{+}$ is used when G is positive (stable stratification) and a negative value $c_{\varepsilon 3}^{-}$ is used when G is negative (unstable stratification). However, Umlauf and Burchard (2005) argued that this is not necessary and that better results (particularly for the heat flux profile) are obtained with considering always a negative value. We do this choice and the value $c_{\varepsilon 3} = -0.65$ is obtained according to (26) of Umlauf et al. (2003) (by considering a steady state Richardson number equal to 0.25).

2.3 Procedure for the $k\varepsilon t$ Parameterization

The new procedure differs from the GLS one by considering that the temperature variance $\overline{\theta'^2}$ is not at equilibrium anymore. Relaxing this assumption takes us from the level $2\frac{1}{2}$ to the level 3 in the hierarchy of Mellor and Yamada (1982). Beyond this mathematical justification, the idea of keeping the non-equilibrium $\overline{\theta'^2}$ equation originated from the fact that the $\overline{\theta'^2}$ dependence appears only in the $\overline{w'\theta'}$ equation (see Eqs. (3) and (4)). Thus, a physical change in the shape of the $\overline{\theta'^2}$ profile will directly impact $\overline{w'\theta'}$. Because we now have an equation for the temperature variance, we are left with (6) and (7) that form a system of 9 equations with 9 unknowns: ($\overline{u'^2}$, $\overline{v'^2}$, $\overline{w'^2}$, $\overline{u'v'}$, $\overline{u'w'}$, $\overline{v'w'}$, $\overline{u'\theta'}$, $\overline{v'\theta'}$, $\overline{w'\theta'}$). For clarity, these 9 equations are written explicitly in Appendix F. We solve this system thanks to Mathematica and we obtain the following expressions:

$$\overline{u'w'} = -c_\mu \frac{k^2}{\varepsilon} \partial_z U, \tag{18}$$

$$\overline{v'w'} = -c_\mu \frac{k^2}{\varepsilon} \partial_z V, \tag{19}$$

$$\overline{w'\theta'} = -c'_\mu \frac{k^2}{\varepsilon} \partial_z \Theta + c'^{*}_\mu \frac{k}{\varepsilon} \beta_3 \overline{\theta'^2}. \tag{20}$$

The momentum fluxes are still downgradient with an eddy viscosity $K_m = c_\mu \frac{k^2}{\varepsilon}$ whereas the temperature flux now has a "non-gradient" contribution $\Gamma_{k\varepsilon t} = c'^{*}_\mu \frac{k}{\varepsilon} \beta_3 \overline{\theta'^2}$

related to the temperature variance in addition to the downgradient part with eddy diffusivity $K_t = c'_\mu \frac{k^2}{\varepsilon}$. The stability functions c_μ, c'_μ and c'^*_μ can be expressed in the following forms

$$c_\mu = \frac{n_0 + n_1\alpha_N + n_2\alpha_M + n_3\alpha_T}{d_0 + d_1\alpha_N + d_2\alpha_M + d_3\alpha_N\alpha_M + d_4\alpha_N^2 + d_5\alpha_M^2}, \tag{21}$$

$$c'_\mu = \frac{n_{0T} + n_{1T}\alpha_N + n_{2T}\alpha_M}{d_0 + d_1\alpha_N + d_2\alpha_M + d_3\alpha_N\alpha_M + d_4\alpha_N^2 + d_5\alpha_M^2}, \tag{22}$$

$$c'^*_\mu = \frac{n^*_{0T} + n^*_{1T}\alpha_N + n^*_{2T}\alpha_M}{d_0 + d_1\alpha_N + d_2\alpha_M + d_3\alpha_N\alpha_M + d_4\alpha_N^2 + d_5\alpha_M^2}, \tag{23}$$

with $\alpha_N = \frac{k^2}{\varepsilon^2}N^2$, $\alpha_M = \frac{k^2}{\varepsilon^2}M^2$, and $\alpha_T = \frac{k}{\varepsilon^2}\beta_3^2\overline{\theta'^2}$. Coefficients n_i, n_{iT} and d_i depends on the coefficients c_i and c_{iT}. Their full expressions are given in Appendix G. Taking the values of the c_i and c_{iT} given in Table 1, the stability functions are approximately as follows

$$c_\mu = \frac{0.1067 + 0.0001072\alpha_N - 0.0001205\alpha_M + 0.004673\alpha_T}{1 + 0.07843\alpha_N + 0.02872\alpha_M + 0.0003389\alpha_N\alpha_M + 0.001506\alpha_N^2 - 0.00003372\alpha_M^2}, \tag{24}$$

$$c'_\mu = \frac{0.1120 + 0.003766\alpha_N + 0.0008871\alpha_M}{1 + 0.07843\alpha_N + 0.02872\alpha_M + 0.0003389\alpha_N\alpha_M + 0.001506\alpha_N^2 - 0.00003372\alpha_M^2}, \tag{25}$$

$$c'^*_\mu = \frac{0.1120 + 0.003766\alpha_N + 0.003344\alpha_M}{1 + 0.07843\alpha_N + 0.02872\alpha_M + 0.0003389\alpha_N\alpha_M + 0.001506\alpha_N^2 - 0.00003372\alpha_M^2}. \tag{26}$$

As in the GLS procedure, the TKE and ε equations (equations (16) and (17)) are solved prognostically. The only difference in these equations is about the $c_{\varepsilon 3}$ coefficient which is now calculated to be $c_{\varepsilon 3} = -1.83$ according to (26) of Umlauf et al. (2003) (by considering a steady state Richardson number equal to 0.25).

Beyond this minor change, one key difference is that the temperature variance is now also solved prognostically through:

$$D_t\overline{\theta'^2} = -2\,\overline{w'\theta'}\,\partial_z\Theta - \frac{2}{c_T}\frac{\varepsilon}{k}\,\overline{\theta'^2} + \mathscr{D}_{\overline{\theta'^2}}, \tag{27}$$

with $\mathscr{D}_{\overline{\theta'^2}} = \partial_z(\frac{K_m}{\sigma_{\overline{\theta'^2}}}\partial_z\overline{\theta'^2})$ the diffusion and $\sigma_{\overline{\theta'^2}}$ the Schmidt number for the temperature variance. As for the TKE equation, the diffusion term $\mathscr{D}_{\overline{\theta'^2}}$ results from the closure of the third-order moment $\overline{w'\theta'\theta'}$ by a downgradient formulation. We did not find any estimations of the Schmidt number $\sigma_{\overline{\theta'^2}}$ in the literature and, as a first guess, we took $\sigma_{\overline{\theta'^2}} = \sigma_k = 1$, meaning that the temperature variance is diffused with the same intensity as TKE.

We add several comments about the non-gradient term $\Gamma_{k\varepsilon t} = c'^{*}_{\mu}\frac{k}{\varepsilon}\beta_3\,\overline{\theta'^2}$ we obtained for the temperature flux. Firstly, we recall that, by writing $\overline{w'\theta'} = -K_m(\partial_z\Theta - \gamma_{k\varepsilon t})$, we highlight that $\gamma_{k\varepsilon t} = \frac{c'^{*}_{\mu}}{c'_{\mu}}\frac{1}{k}\beta_3\overline{\theta'^2}$ gives the stable stratification towards which $\partial_z\Theta$ tends to relax. Secondly, the form of $\Gamma_{k\varepsilon t}$ can be compared to the one found by Deardorff (1972). By reasoning with the $\overline{w'\theta'}$ equation, Deardorff (1972) found a non-gradient term $\Gamma_{\mathrm{Deardoff}} \propto l/k^{1/2}\,\overline{\theta'^2}$ with l a mixing length introduced for the parameterization of the pressure-temperature correlation. If we consider the classic scaling $l \propto k^{3/2}/\varepsilon$ (see for example Rodi, 1987; Umlauf & Burchard, 2003, 2005), we obtain $\Gamma_{\mathrm{Deardoff}} \propto k/\varepsilon\,\overline{\theta'^2}$. The non-gradient expressions of $\Gamma_{k\varepsilon t}$ and $\Gamma_{\mathrm{Deardoff}}$ thus both exhibit the same dependence on the turbulence time scale k/ε and on the temperature variance $\overline{\theta'^2}$. This is fundamentally different from $\Gamma_{\mathrm{KPP}} \propto G\,\overline{w'\theta'}|_{z=0}$ which is written explicitly as a redistribution of the surface temperature flux $\overline{w'\theta'}|_{z=0}$ according to an empirical shape function G that is a third-order polynomial of the dimensionless vertical coordinate z/h with h the mixed layer depth.

Finally, we point out that, just as we retained the non-equilibrium equation of $\overline{\theta'^2}$ to obtain a non-gradient term for $\overline{w'\theta'}$, it would be tempting to retain the non-equilibrium equation of $\overline{w'^2}$ to obtain non-gradient terms for the velocity fluxes $\overline{u'w'}$ and $\overline{v'w'}$. We solved this problem and, astonishingly, the velocity fluxes $\overline{u'w'}$ and $\overline{v'w'}$ in this context are still downgradient. Results of this $k-\varepsilon-\overline{\theta'^2}-\overline{w'^2}$ model are detailed in Appendix H.

2.4 1D Models Simulations

We implemented the $k\varepsilon t$ parameterization, with the formalism described in section 2.3, in the 1D code presented in Fearon et al. (2020). This code is a standalone 1D vertical version of the Coastal and Regional Ocean COmmunity model (CROCO, `https://www.croco-ocean.org/`) and allows to run simulations with KPP, TKE, and several GLS schemes (note that we also re-implemented the $k-\varepsilon$ model with the formalism presented in section 2.2, that is equivalent to using the Canuto et al. (2001) stability functions).

The temperature variance equation (27) is discretized using a backward Euler scheme in time. To preserve the positivity of $\overline{\theta'^2}$, the Patankar trick is used (Patankar, 1980; Burchard, 2002; Lemarié et al., 2021). Boundary conditions for the temperature variance are zero at the bottom of the domain (Dirichlet condition), while at the surface a homogeneous Neumann condition is used (no flux of temperature variance).

For every test case, we performed the simulations using the $k-\varepsilon$ model, the $k\varepsilon t$ model, and the KPP model. The changes induced by the $k\varepsilon t$ model, particularly the influence of the non-gradient term, will be analyzed by comparing with the $k-\varepsilon$ model. Concerning the KPP scheme, the simulations were done with and without its non-gradient term. The goal is to compare this term and its effect to the non-gradient term obtained in the $k\varepsilon t$ parameterization. The version of KPP used here is the original one described in Large et al. (1994).

2.5 LES Simulation

In order to validate the $k\varepsilon t$ model, we performed some LES simulations. Practically, we use the Basilisk code (`http://basilisk.fr`, Popinet, 2020) to solve the three-dimensional Boussinesq equations in a small oceanic patch near the air-sea interface. We intend to explicitly compute the turbulent fluxes and the mean vertical profiles of temperature for buoyancy-driven convection and wind-driven convection. We can then compare these fluxes with the parameterization. The size of the domain is $L_x = L_y = 1200$ m (periodic in the horizontal direction), and $L_z = 600$ m. The grid resolution is isotropic (2.3 m) with $512 \times 512 \times 256$ cells. All variables are discretized at the cell center and are advected using the Bell-Collela-Glaz method. There is no explicit viscosity and no explicit diffusivity: both these terms are handled implicitly by the advection scheme. The surface forcing (wind and heat flux) is applied at the upper grid cell with a relaxation term. The bottom boundary condition is free slip for the velocity and inhomogeneous Neumann for the temperature (set to the initial stratification). The model is initialized with zero velocity and prescribed stratification for temperature (see next paragraph) to which we add a small random perturbation of magnitude 10^{-3} °C. We use an adaptive time step adjusted with a CFL condition of 0.6. Averages are computed in a post-processing step: the overbar is interpreted here as a horizontal average and primes are deviations from this horizontal average.

2.6 The Two Test Cases: Cooling-Dominant and Wind-Dominant

Two simulation setups were defined in order to capture the different convective regimes highlighted in Legay et al. (2024). The first configuration is a cooling-dominant simulation forced by a surface net heat flux of $Q_0 = -320\,\mathrm{W\,m^{-2}}$ and a wind stress of $\tau_x = 0.64\,\mathrm{N\,m^{-2}}$; it is initialized with a surface temperature of 293 K and a constant stratification of 3.9 K/1000 m. The second one is a wind-dominant simulation forced by a surface net heat flux of $Q_0 = -8\,\mathrm{W\,m^{-2}}$ and a wind stress of $\tau_x = 0.41\,\mathrm{N\,m^{-2}}$; it is initialized with a surface temperature of 293 K and a constant stratification of 1.2 K/1000 m. Rotation is included with a Coriolis frequency of $f = 10^{-4}\ \mathrm{s^{-1}}$, this corresponds to a latitude of 44 °N. The two cases are simulated with 10 days of constant forcing conditions. For the 1D simulations, the domain is discretized on the same vertical grid as the 3D model (uniformly spaced vertical grid of 256 points), and the time step is 360 s.

2.7 Nondimensionalization

In order to compare the shape of the different profiles, variables are made dimensionless. For the depth, we found that using the depth of the maximum temperature variance $z[\max(\overline{\theta'^2})]$ as a proxy of the mixed layer depth h_m is the best choice for two main reasons. Firstly, the temperature variance is well converged with a maximum that is prominent, easy to identify, and located at the same depth as the classic definition of the minimum of $\overline{w'\theta'}$ (see Figure 2). Second, this definition holds for wind-dominant simulations whereas in this case, the temperature flux profile can be far from the idealized version presented in Figure 1. We mention that while this method works in most cases, there are some conditions where $\overline{\theta'^2}$ is maximum at the surface. In this case, we simply considered the second maximum strictly below the surface. We tested other definitions of h_m such as the minimum of the temperature flux $\overline{w'\theta'}$ or other definitions of the mixed layer depth h_m, but they appeared to be less robust definitions (subject to noisy variations).

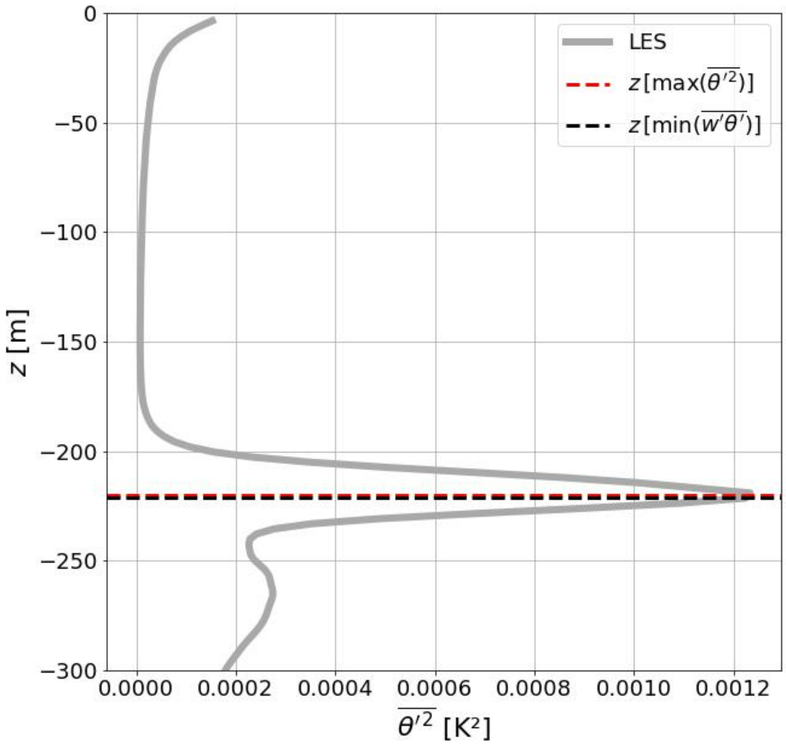

Figure 2. Temperature variance profile of the LES simulation at the end of the simulation for the cooling-dominant case. Dashed lines indicate two different proxies of the mixed layer depth: the maximum of the temperature variance and the minimum of the temperature flux. In this case, these two proxies are localized at the same depth.

The other nondimensionalizations consist in normalizing the temperature flux by its surface value $\overline{w'\theta'}|_0 = Q_0/(\rho_0 c_p)$, with $\rho_0 = 1027\,\mathrm{kg\,m^{-3}}$ the reference density and $c_p = 4000\,\mathrm{J\,kg^{-1}\,K^{-1}}$ the specific heat capacity; and normalizing the temperature in $(\Theta - \Theta_{max})\,/\,\Theta^*$ with Θ_{max} the maximum of the temperature over the vertical and $\Theta^* = \overline{w'\theta'}|_0/w^*$ a scaling of the temperature, with $w^* = (-B_0 h_m)^{1/3}$ a scaling of the velocity of the convective thermals (Willis & Deardorff, 1974; Marshall & Schott, 1999), $B_0 = g\alpha Q_0/(\rho_0 c_p)$ the surface buoyancy flux, g the gravitational acceleration and α the thermal expansion coefficient taken equal to $2.6 \times 10^{-4}\,\mathrm{K^{-1}}$.

3 Results and Discussion

3.1 Cooling-Dominant Case

Figure 3 presents the dimensionless temperature flux profile of the $k-\varepsilon$ and the $k\varepsilon t$ simulations at the end of the 10 days of simulations for the cooling-dominant case. The $k\varepsilon t$ flux is further decomposed into its downgradient $(-K_t \partial_z \Theta)$ and its non-gradient $(\Gamma_{k\varepsilon t})$ components (see Eq. 20). It is remarkable that even if the expression of the total flux changed drastically between the two parameterizations, the $k\varepsilon t$ profile is very similar to the $k-\varepsilon$ one that exhibits the classic pattern expected for a cooling-dominant simulation: a linear decrease from the surface to the bottom of the mixed layer where it reaches a minimum which is approximately -0.2 times the surface flux. The non-gradient flux is positive (by definition), and it is zero at the surface and at the bottom of the mixed layer; hence, it does not add or remove any heat but rather redistributes heat among the mixed layer. This term is responsible for warming the upper part of the mixed layer and cooling the lower part of the mixed layer (the temperature equation is of the form $D_t\Theta = ... - \partial_z \overline{w'\theta'}$ and it is then the sign of $-\partial_z \Gamma_{k\varepsilon t}$ that is important to distinguish between cooling and warming). This is qualitatively the effect we expect from a coherent thermal: thermals grow by entraining cold water near the surface, resulting in a warming of the upper part of the mixed layer, and then detrain in the environment which results in a cooling of the bottom part of the mixed layer.

Figure 4 presents the dimensionless temperature profile of the $k-\varepsilon$, the $k\varepsilon t$, the KPP, and the LES simulations at the end of the 10 days of simulations. Dashed lines highlight the location h_t, the depth at which $\partial_z\Theta = 0$ for each case. The overall comparison with the LES is better with $k\varepsilon t$ scheme than with $k-\varepsilon$: while the $k-\varepsilon$ model

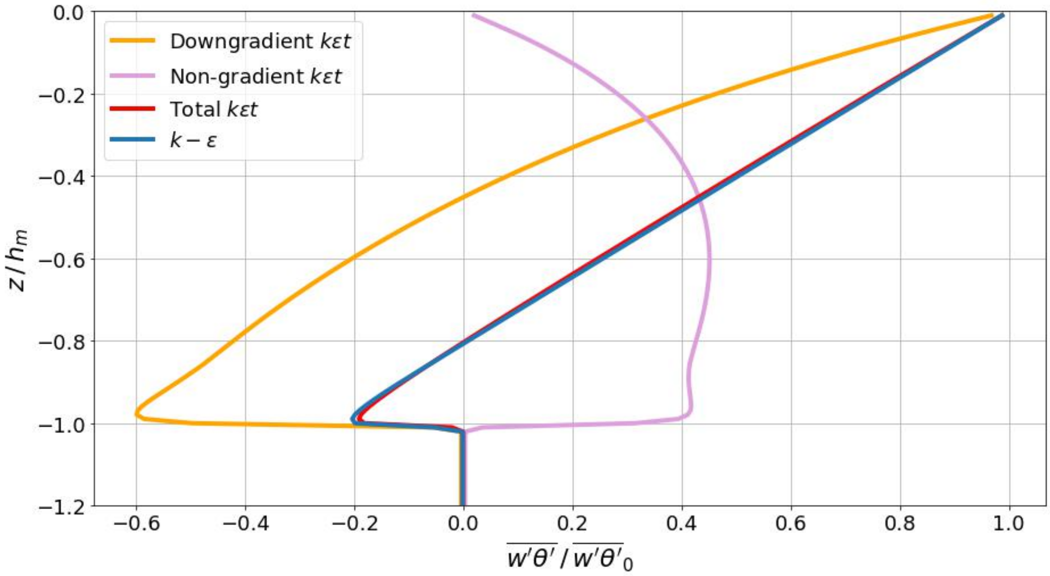

Figure 3. Dimensionless temperature flux profiles of the $k-\varepsilon$ and the $k\varepsilon t$ simulations at the end of the 10 days of simulations for the cooling-dominant case. The $k\varepsilon t$ flux is further decomposed into its downgradient ($-K_t\partial_z\Theta$) and its non-gradient ($\Gamma_{k\varepsilon t}$) components (see Eq. 20).

predicts $h_t = h_f = -0.8h_m$ (by definition of a pure downgradient flux, see Fig. 3), this co-location constraint is relaxed in the $k\varepsilon t$ simulation, for which $h_t = -0.44h_m$, which is closer to the LES ($h_t = -0.41h_m$). The KPP scheme predicts $h_t = -0.2h_m$, whereas the KPP simulation without the non-local term Γ_{KPP} gives $h_t = -0.93h_m$. Therefore, Γ_{KPP} has the same expected effect to raise h_t up as the non-gradient term of $k\varepsilon t$ but, none of the two KPP simulations (with or without Γ_{KPP}) give a satisfactory h_t in comparison to the LES.

Figure 5 shows the temporal evolution of h_t/h_m for the 10 days of the simulation. The evolution of this quantity in the LES, although a bit noisy, shows h_t/h_m between -0.4 and -0.6 at the end of the simulation. The $k-\varepsilon$ values decrease then stabilize around -0.8. The KPP simulation quickly stabilizes near -0.2 whereas KPP without the non-local term gives a continuous decrease of h_t/h_m with values reaching -0.93 at the end of the 10 days. The $k\varepsilon t$ curve, among all schemes, exhibits the closest values to those of the LES. However, it results in a continuous increase during 10 days. This behavior can be modified by considering a different value of $\sigma_{\overline{\theta'^2}}$. Thus, another simulation of the $k\varepsilon t$ model with $\sigma_{\overline{\theta'^2}} = 10$ (a temperature variance that diffuses 10 times less

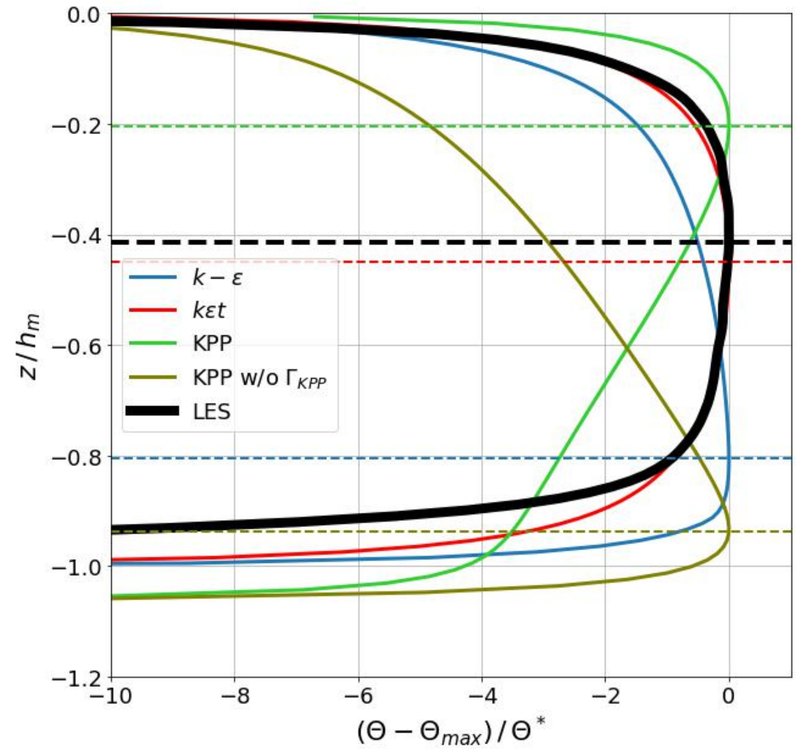

Figure 4. Dimensionless temperature profiles of the $k - \varepsilon$, the $k\varepsilon t$, the KPP, and the LES simulations at the end of the 10 days of simulations for the cooling-dominant case. The KPP model was run with and without its non-gradient term Γ_{KPP}. Dashed lines highlight the location h_t of the zero of the gradient $\partial_z \Theta$.

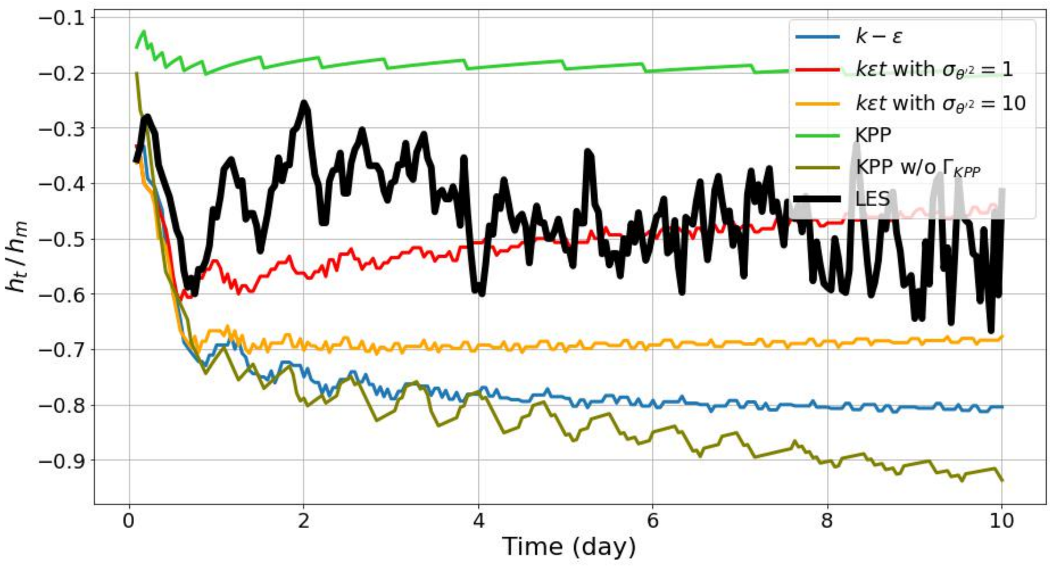

Figure 5. Temporal evolution of h_t/h_m in the cooling-dominant case for the $k - \varepsilon$, the $k\varepsilon t$, the KPP and the LES simulations. The $k\varepsilon t$ simulation was run with two different values of the Schmidt number for the temperature variance: $\sigma_{\overline{\theta'^2}} = 1$ and $\sigma_{\overline{\theta'^2}} = 10$. The KPP model was run with and without its non-gradient term $\mathit{\Gamma}_{\mathrm{KPP}}$.

than the velocities) gives h_t/h_m that stabilizes around -0.7. This preliminary test highlights the need to adjust all parameters of this closure with advanced Bayesian methods such as the ones used in Souza et al. (2020) and Wagner et al. (2023). This calibration procedure would require an ensemble of LES simulations in order to not overfit the parameters to the two LES used here and this task is beyond the scope of this study.

Figure 6 shows a comparison between the non-gradient term of $k\varepsilon t$ (run with two different values $\sigma_{\overline{\theta'^2}} = 1$ and $\sigma_{\overline{\theta'^2}} = 10$) and the non-local term of KPP at the end of the 10 days of simulation. These profiles share the property of vanishing at the surface and at the bottom of the mixed layer, they therefore both act as a redistribution of heat in the mixed layer. The KPP term appears to have a single-mode shape. In fact, $\mathit{\Gamma}_{\mathrm{KPP}}$ can be written as $\mathit{\Gamma}_{\mathrm{KPP}}(z) = C_s G(z) \overline{w'\theta'}|_0$ with C_s a constant (see for example Equation (20) of Van Roekel et al., 2018). The vertical dependence is entirely contained in G which is a third-order polynomial. Hence, $\mathit{\Gamma}_{\mathrm{KPP}}$ can only have a single positive mode. Instead, $\mathit{\Gamma}_{k\varepsilon t}$ presents a bi-modal shape for both $\sigma_{\overline{\theta'^2}} = 1$ and $\sigma_{\overline{\theta'^2}} = 10$. For $\sigma_{\overline{\theta'^2}} = 1$, the two modes are close one to the other but, for $\sigma_{\overline{\theta'^2}} = 10$, the non-gradient term presents two clear distinct modes. In the latter case, the simple qualitative way of see-

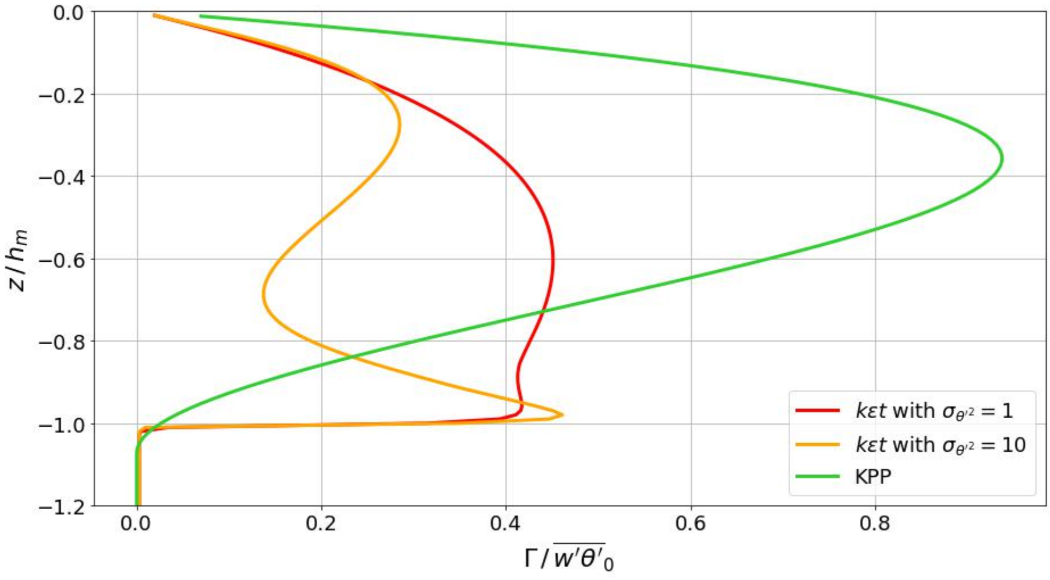

Figure 6. Dimensionless profiles of the non-gradient term of $k\varepsilon t$ (run with two different values $\sigma_{\overline{\theta'^2}} = 1$ and $\sigma_{\overline{\theta'^2}} = 10$) and KPP at the end of the 10 days of simulation for the cooling-dominant case.

ing the non-gradient term as the effect of a thermal is no longer relevant. This point is supported by Zhou et al. (2018) who proved that the often-implied association of the gradient and non-gradient term terms to the local and non-local eddies is partially wrong. Analyses of the contribution of the different factors of $\mathit{\Gamma}_{k\varepsilon t} = c'^*_\mu \frac{k}{\varepsilon}\beta_3 \overline{\theta'^2}$ (not shown) indicated that the mode close to the mixed layer bottom is mainly due to a maximum of $\overline{\theta'^2}$ whereas the mode closest to the surface is a result of a complex interaction of all the terms in the expression of the non-gradient term. Knowing that $\mathit{\Gamma}_{k\varepsilon t}$ presents a bi-modal shape could be of interest for adapting the KPP non-gradient term. For example, it would be possible to consider $\mathit{\Gamma}_{\mathrm{KPP}}$ as a sum of two polynomials rather than one for trying to catch this bi-modal feature.

3.2 Wind-Dominant Case

Figure 7 presents the dimensionless temperature flux profile of the $k-\varepsilon$ and the $k\varepsilon t$ simulations at the end of the 10 days of simulations for the wind-dominant case. For the shape of the flux, we get similar conclusions as in the cooling-dominant case: we ob-

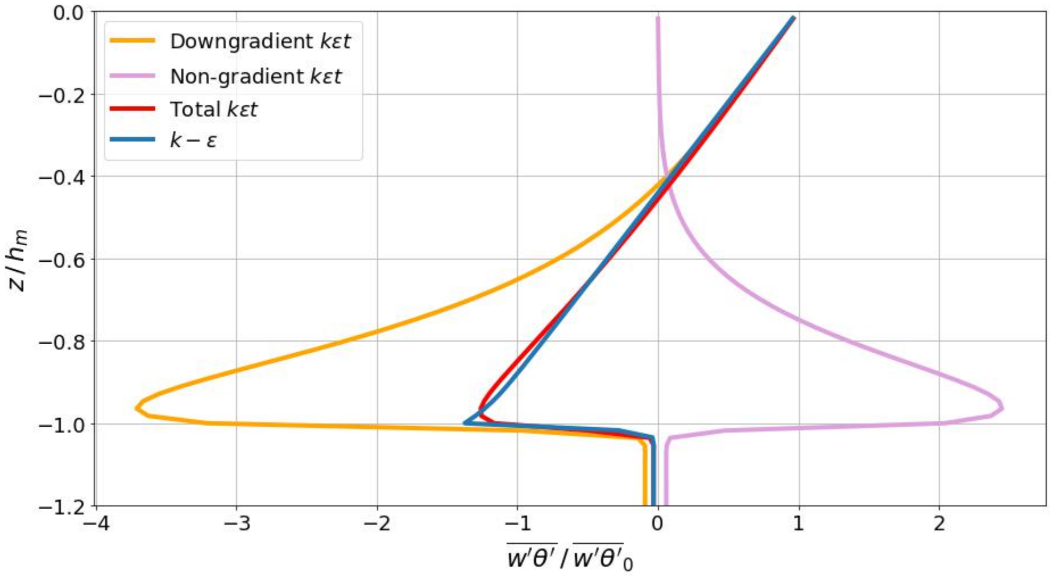

Figure 7. Dimensionless temperature flux profiles of the $k - \varepsilon$ and the $k\varepsilon t$ simulations at the end of the 10 days of simulations for the wind-dominant case. The $k\varepsilon t$ flux is further decomposed into its downgradient and its non-gradient components.

tain a remarkable agreement between the $k\varepsilon t$ total profile and the $k{-}\varepsilon$ profile even if the expression of the total flux changed between the two parameterizations.

Figure 8 presents the dimensionless temperature profile of the $k{-}\varepsilon$, the $k\varepsilon t$, the KPP, and the LES simulations at the end of the 10 days of simulations. Here again, dashed lines highlight the location of h_t in all cases. The effect of the non-gradient term of $k\varepsilon t$ of raising h_t is negligible here, and this is fine since $k{-}\varepsilon$ correctly predicts the LES profile. Instead, the difference between KPP and KPP without $\mathit{\Gamma}_{\mathrm{KPP}}$ is substantial. KPP without $\mathit{\Gamma}_{\mathrm{KPP}}$ gives a good profile while the full KPP results in a profile that presents a high value of h_t. The fact that $k{-}\varepsilon$ and KPP without $\mathit{\Gamma}_{\mathrm{KPP}}$ are already satisfactory suggests that non-gradient effects are less important in this wind-dominant case than in the cooling-dominant case. If we adopt the disputed view of associating non-gradient effects to non-local eddies, this suggests that the deepening is here dominated by local eddies driven by shear while the deepening in the cooling-dominant case is driven by non-local thermals.

Figure 9 shows the temporal evolution of h_t/h_m for all models. The LES evolution consists of a continuous decrease until near -0.45 at the end of the simulation (with no

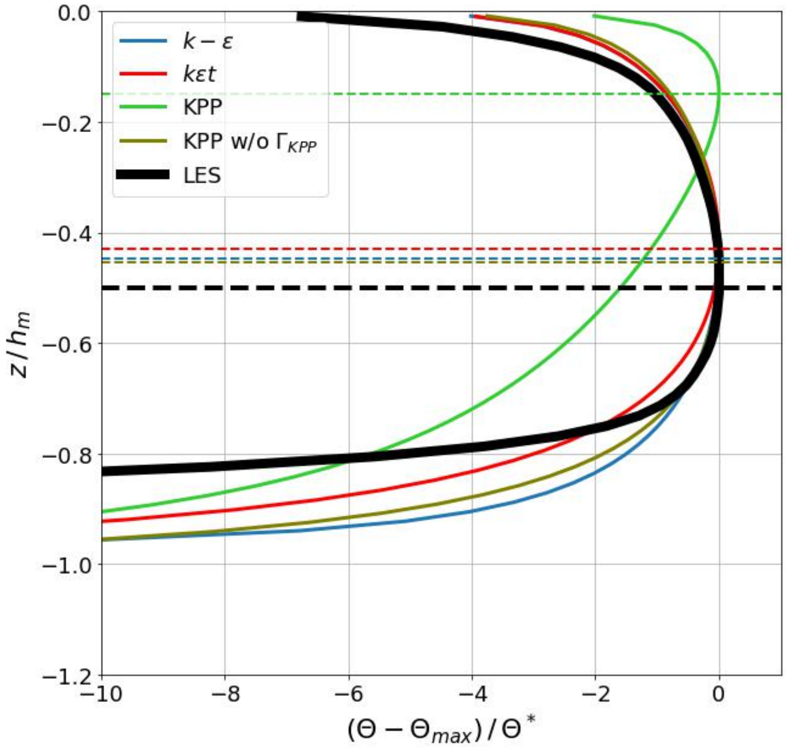

Figure 8. Dimensionless temperature profiles of the $k - \varepsilon$, the $k\varepsilon t$, the KPP, and the LES simulations at the end of the 10 days of simulations for the wind-dominant case. The KPP model was run with and without its non-gradient term $\mathit{\Gamma}_{\mathrm{KPP}}$. Dashed lines highlight the location h_t of the zero of the gradient $\partial_z \mathit{\Theta}$.

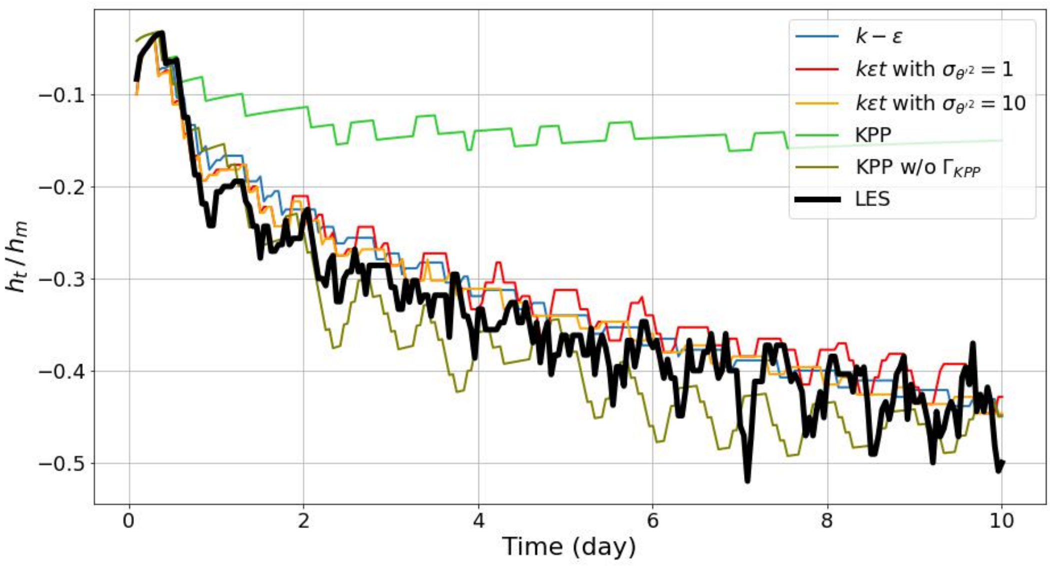

Figure 9. Temporal evolution of h_t/h_m in the wind-dominant case for the $k - \varepsilon$, the $k\varepsilon t$, the KPP and the LES simulations. The $k\varepsilon t$ simulation was run with two different values of the Schmidt number for the temperature variance: $\sigma_{\overline{\theta'^2}} = 1$ and $\sigma_{\overline{\theta'^2}} = 10$. The KPP model was run with and without its non-gradient term $\mathit{\Gamma}_{\mathrm{KPP}}$.

clear convergence). This evolution is reproduced by $k{-}\varepsilon$, $k\varepsilon t$ and KPP without $\mathit{\Gamma}_{\mathrm{KPP}}$. Instead, the comparison of the full KPP with the LES is not in favor of KPP, since h_t/h_m stabilizes around -0.15 in this case. The LES evolution presents inertial oscillations of h_t/h_m at the inertial period $T_f = 2\pi/f = 17\,\mathrm{h}\,30\,\mathrm{min}$. This is captured by $k\varepsilon t$ and KPP without $\mathit{\Gamma}_{\mathrm{KPP}}$ but not by the full KPP and $k{-}\varepsilon$. Changing the value of $\sigma_{\overline{\theta'^2}}$ gives here almost no effect, contrary to what was highlighted in the cooling-dominant case. This supports the fact that non-gradient effects are probably negligible in wind-dominant regimes.

Figure 10 shows a comparison between the non-gradient term of $k\varepsilon t$ (run with two different values $\sigma_{\overline{\theta'^2}} = 1$ and $\sigma_{\overline{\theta'^2}} = 10$) and KPP at the end of the 10 days of simulation. The KPP shape is very similar to the one of the cooling-dominant case (Figure 6). On the opposite, the $k\varepsilon t$ profiles changed drastically in comparison to the cooling-dominant case. Indeed, these profiles present now a single-mode shape. The mode closest to the surface disappeared because $\overline{\theta'^2}$ is here equal to zero at the surface. This behavior can again inspire the construction of the KPP non-gradient term. If, as suggested, it is constructed by the sum of two polynomials, the polynomial with its maximum close

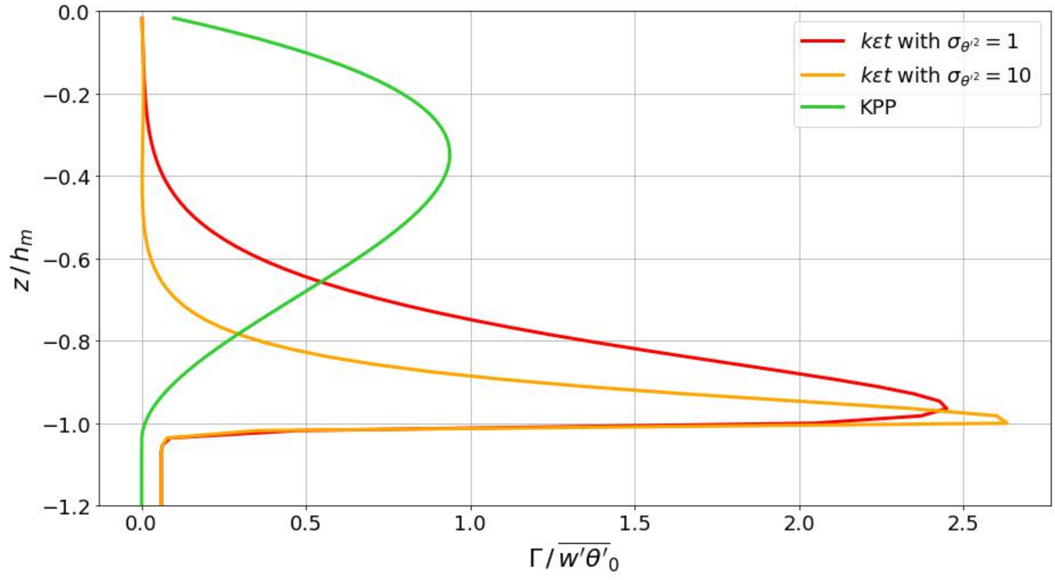

Figure 10. Dimensionless profiles of the non-gradient term of $k\varepsilon t$ (run with two different values $\sigma_{\overline{\theta'^2}} = 1$ and $\sigma_{\overline{\theta'^2}} = 10$) and KPP at the end of the 10 days of simulation for the wind-dominant case.

to the surface must vanish in the wind-dominant case whereas the second polynomial with its maximum near the mixed layer bottom must be present both in the cooling-dominant and wind-dominant conditions.

4 Conclusion

The primary motivation behind this research was the need to improve the representation of oceanic convection processes in ocean models. Indeed, most parameterizations adopt a downgradient approach for which the mixing of tracers and momentum primarily occurs in the direction of their gradients, such as from regions of high temperature to low temperature. However, in convective situations, this simplistic assumption falls short, and turbulent fluxes cannot be solely explained or formulated as a downgradient process (Zhou et al., 2018). While this property was originally recognized for atmospheric convection (Hourdin et al., 2002), oceanographers were also aware of this aspect of convection when they introduced a non-local term in KPP (Large et al., 1994). In this context, the non-local term represented the influence and transport of tracers across different spatial locations within a convective system, even when these locations are not immediately adjacent. In simpler terms, the non-local term accounted for the long-range mixing of properties that occur in convective events. While there is an ongoing effort to capture this property with a mass flux parameterization (Giordani et al., 2020), our approach has been to seek an analytical formulation of this non-local (or non-gradient) term within the GLS (Generic Length Scale) models and we can retrospectively comment on that choice based on the results obtained in this article. One key argument for this approach is the desire for consistency and integration within the modeling framework. By deriving the non-gradient term analytically within the GLS framework, we aim to ensure that all components of the parameterization align seamlessly. This approach avoids potential mismatches or inconsistencies that may arise when adding external components to existing parameterizations. Another crucial argument is the need to deepen our physical understanding of oceanic convection processes. By analytically deriving the non-gradient term within the GLS framework, we gain insights into the underlying physics and dynamics governing this term. This understanding can lead to more robust and physically grounded parameterization, improving our ability to capture convective processes accurately. Last, our approach offers flexibility for optimization and adaptation. As the GLS framework provides a versatile platform for parameterization, we can adapt and refine the derived non-gradient term to suit specific oceanic conditions or scenarios. This adaptability is valuable for tailoring the parameterization to different modeling and research needs.

In order to assess the validity of our approach we have compared our scheme with Large eddy simulations (LES). The main metric that we analyzed was the depth of the thermal layer h_t which corresponds to the depth at which $\partial_z \Theta = 0$. We have verified that the effect of the non-gradient term is to raise h_t such that a significant part of the mixed layer is stably stratified (at least in the thermally driven convection), an aspect that was not well reproduced by the $k - \varepsilon$ model. We also noted that since the main effect of this term is to redistribute heat, the addition of the non-gradient term does not have a profound impact on the evolution of the depth of the mixed layer. We have also conducted extensive comparisons with KPP. With these comparisons, we have unveiled common aspects between our newly derived non-gradient term and certain aspects of the KPP non-local term. This comparison suggests that our work has the potential to serve as a source of inspiration for enhancing and fine-tuning the KPP parameterization. Particularly, it could be used to modify the definition of the ad hoc polynomial that shape the diffusivity and the non-local term in KPP.

Extending the derivation to include salinity will allow us to more comprehensively capture the behavior of oceanic convection. We are currently working on this approach: the main challenge is that the non-gradient term for salinity involves coupled equations with temperature, making the analytical derivation significantly more complex. Solving these coupled equations analytically is mathematically challenging and may require additional hypotheses. Another issue is also that the computational demands of implementing a coupled temperature-salinity non-gradient term within ocean models may increase. This can affect model efficiency and require adjustments in computational resources. Despite these difficulties, the extension of the non-gradient term derivation to salinity promises a more comprehensive and accurate representation of oceanic convection. In the near future, our research plans also entail a systematic re-evaluation of all GLS parameters and $k\varepsilon t$ parameters. To achieve this, we will employ an ensemble of LES simulations, with a resolution high enough to capture the energetic eddies in entrainment layers, in conjunction with Bayesian methods (Wagner et al., 2023). Bayesian methods offer a data-driven approach to parameter estimation, allowing us to incorporate real-world observations and LES data into the parameterization process.

Appendix A Open Research

All the codes and the data used for the study are available through the GitHub repository `https://github.com/legaya/James2024-ket/` or the following DOI: `https://doi.org/10.5281/zenodo.10562734`. These archives contain the two Jupyter Notebooks used for performing the 1D simulations and all the analyses, the 1D model described in section 2.4 as Fortran Modules, the Fortran codes needed for generating these modules, the files needed to perform the LES simulations, and the LES results as netCDF files.

Appendix B Coefficients in the Second-Order Moment Equations

Coefficients $c_1, c_2, c_3, c_4, c_5, c_{1T}, c_{2T}, c_{3T}, c_{4T}, c_T$ used in Eqs (3) - (5) are linked to the coefficients introduced by Canuto et al. (2001) through the following formulas:

$$\begin{aligned} c_1 &= 1/\lambda, \qquad c_2 = \alpha_1, \qquad c_3 = 1-\beta_5, \qquad c_4 = 4/3\,\alpha_1 - 4/5, \qquad c_5 = \alpha_1 - \alpha_2, \\ c_{1T} &= \lambda_5/2, \qquad c_{2T} = 3/4\,\alpha_3, \qquad c_{3T} = \gamma_1, \qquad c_{4T} = \alpha_3/2, \qquad c_T = 2\,\lambda_8/(1-\gamma_1). \end{aligned} \tag{B1}$$

Appendix C Expressions of the Main Tensors under the Boundary Layer Approximation

After applying the boundary layer approximation, the tensors $P_{ij}, B_{ij}, S_{ij}, V_{ij}, Z_{ij}$ used in Eqs (3) - (5) simplify to

$$P_{ij} = \begin{pmatrix} -2\,\partial_z U\,\overline{u'w'} & -\partial_z U\,\overline{v'w'} - \partial_z V\,\overline{u'w'} & -\partial_z U\,\overline{w'^2} \\ -\partial_z U\,\overline{v'w'} - \partial_z V\,\overline{u'w'} & -2\,\partial_z V\,\overline{v'w'} & -\partial_z V\,\overline{w'^2} \\ -\partial_z U\,\overline{w'^2} & -\partial_z V\,\overline{w'^2} & 0 \end{pmatrix} \tag{C1}$$

$$B_{ij} = \begin{pmatrix} 0 & 0 & \beta_3\,\overline{u'\theta'} \\ 0 & 0 & \beta_3\,\overline{v'\theta'} \\ \beta_3\,\overline{u'\theta'} & \beta_3\,\overline{v'\theta'} & 2\,\beta_3\,\overline{w'\theta'} \end{pmatrix} \tag{C2}$$

$$S_{ij} = \frac{1}{2}\begin{pmatrix} 0 & 0 & \partial_z U \\ 0 & 0 & \partial_z V \\ \partial_z U & \partial_z V & 0 \end{pmatrix} \tag{C3}$$

$$V_{ij} = \frac{1}{2}\begin{pmatrix} 0 & 0 & \partial_z U \\ 0 & 0 & \partial_z V \\ -\partial_z U & -\partial_z V & 0 \end{pmatrix} \tag{C4}$$

$$Z_{ij} = \begin{pmatrix} \overline{u'w'}\,\partial_z U & \frac{1}{2}\,\overline{v'w'}\,\partial_z U + \frac{1}{2}\,\overline{u'w'}\,\partial_z V & \frac{1}{2}\,\partial_z U\,(\overline{w'^2 - u'^2}) - \frac{1}{2}\,\partial_z V\,\overline{u'v'} \\ \frac{1}{2}\,\overline{v'w'}\,\partial_z U + \frac{1}{2}\,\overline{u'w'}\,\partial_z V & \overline{v'w'}\,\partial_z V & \frac{1}{2}\,\partial_z V\,(\overline{w'^2 - v'^2}) - \frac{1}{2}\,\partial_z U\,\overline{u'v'} \\ \frac{1}{2}\,\partial_z U\,(\overline{w'^2 - u'^2}) - \frac{1}{2}\,\partial_z V\,\overline{u'v'} & \frac{1}{2}\,\partial_z V\,(\overline{w'^2 - v'^2}) - \frac{1}{2}\,\partial_z U\,\overline{u'v'} & -\overline{u'w'}\,\partial_z U - \overline{v'w'}\,\partial_z V \end{pmatrix} \tag{C5}$$

Appendix D The Algebraic System of 10 Equations of the GLS Formalism

For clarity, we give here the explicit writing of the 10 equations presented in Eqs (6) - (8) and that are the basis of the GLS formalism:

$$
\begin{aligned}
0 &= -c_1 \frac{\varepsilon}{k}(\overline{u'^2} - \frac{2}{3}k) + (1-c_2)(-\frac{4}{3}\,\overline{u'w'}\,\partial_z U + \frac{2}{3}\,\overline{v'w'}\,\partial_z V) - \frac{2}{3}(1-c_3)\beta_3\,\overline{w'\theta'} - c_5\overline{u'w'}\,\partial_z U \\
0 &= -c_1 \frac{\varepsilon}{k}(\overline{v'^2} - \frac{2}{3}k) + (1-c_2)(-\frac{4}{3}\,\overline{v'w'}\,\partial_z V + \frac{2}{3}\,\overline{u'w'}\,\partial_z U) - \frac{2}{3}(1-c_3)\beta_3\,\overline{w'\theta'} - c_5\overline{v'w'}\,\partial_z V \\
0 &= -c_1 \frac{\varepsilon}{k}(\overline{w'^2} - \frac{2}{3}k) + (\frac{2}{3} - \frac{2}{3}c_2 + c_5)(\overline{u'w'}\,\partial_z U + \overline{v'w'}\,\partial_z V) + \frac{4}{3}(1-c_3)\beta_3\,\overline{w'\theta'} \\
0 &= -c_1 \frac{\varepsilon}{k}\,\overline{u'v'} - (1-c_2)(\overline{v'w'}\,\partial_z U + \overline{u'w'}\,\partial_z V) - \frac{1}{2}c_5(\overline{v'w'}\,\partial_z U + \overline{u'w'}\,\partial_z V) \\
0 &= -c_1 \frac{\varepsilon}{k}\,\overline{u'w'} - (1-c_2)\overline{w'^2}\,\partial_z U + (1-c_3)\beta_3\,\overline{u'\theta'} - \frac{1}{2}c_4 k\,\partial_z U - \frac{1}{2}c_5(\overline{w'^2}\,\partial_z U - \overline{u'^2}\,\partial_z U - \overline{u'v'}\,\partial_z V) \\
0 &= -c_1 \frac{\varepsilon}{k}\,\overline{v'w'} - (1-c_2)\overline{w'^2}\,\partial_z V + (1-c_3)\beta_3\,\overline{v'\theta'} - \frac{1}{2}c_4 k\,\partial_z V - \frac{1}{2}c_5(\overline{w'^2}\,\partial_z V - \overline{v'^2}\,\partial_z V - \overline{u'v'}\,\partial_z U) \\
0 &= -c_{1T} \frac{\varepsilon}{k}\,\overline{u'\theta'} - (1 - c_{2T} - \frac{1}{2}c_{4T})\,\overline{w'\theta'}\,\partial_z U - \overline{u'w'}\,\partial_z \Theta \\
0 &= -c_{1T} \frac{\varepsilon}{k}\,\overline{v'\theta'} - (1 - c_{2T} - \frac{1}{2}c_{4T})\,\overline{w'\theta'}\,\partial_z V - \overline{v'w'}\,\partial_z \Theta \\
0 &= -c_{1T} \frac{\varepsilon}{k}\,\overline{w'\theta'} - \overline{w'^2}\,\partial_z \Theta + (1-c_{3T})\beta_3\,\overline{\theta'^2} - \frac{1}{2}c_{4T}(\overline{u'\theta'}\,\partial_z U + \overline{v'\theta'}\,\partial_z V) \\
0 &= -2\,\overline{w'\theta'}\,\partial_z \Theta - \frac{2}{c_T}\frac{\varepsilon}{k}\,\overline{\theta'^2}
\end{aligned}
\tag{D1}
$$

Appendix E Coefficients of the Stability Functions for the GLS Formalism

Coefficients $n_0, n_1, n_2, n_{0T}, n_{1T}, n_{2T}, d_0, d_1, d_2, d_3, d_4, d_5$ of the GLS stability functions (Eqs. (12) and (13)) have the following definitions:

$$
n_0 = \frac{4 - 4c_2 + 3c_4}{6c_1},
$$

$$
n_1 = \frac{c_1 c_{1T} c_T (1 - c_{3T})(4 - 4c_2 + 3c_4) - 2c_1(1-c_3)(2 - 2c_{2T} - c_{4T}) + 4c_{1T}(1-c_3)(c_4 - c_5)}{6c_1^2 c_{1T}^2},
$$

$$
n_2 = \frac{-c_{4T}(4 - 4c_2 + 3c_4)(2 - 2c_{2T} - c_{4T})}{24 c_1 c_{1T}^2},
$$

$$
n_{0T} = \frac{2}{3c_{1T}}, \qquad n_{1T} = \frac{2(1-c_3)}{3c_1 c_{1T}^2},
$$

$$
n_{2T} = \frac{c_1 c_{4T}(4 - 4c_2 + 3c_4) + 8c_5 c_{1T}(1 - c_2 + c_5) - 2c_4 c_{1T}(2 - 2c_2 + 3c_5)}{12 c_1^2 c_{1T}^2},
$$

$$d_0 = 1, \qquad d_1 = \frac{7 - 7c_3 + 3c_1 c_T(1 - c_{3T})}{3c_1 c_{1T}},$$

$$d_2 = \frac{3c_5^2 + 6c_5(1 - c_2) + 2(1 - c_2)^2}{3c_1^2} - \frac{c_{4T}(2 - 2c_{2T} - c_{4T})}{4c_{1T}^2},$$

$$\begin{aligned} d_3 = & \frac{c_5 c_{1T}(1 - c_3)(2 - 2c_2 + c_5)}{3c_1^3 c_{1T}^2} \\ & + \frac{c_1 c_{1T} c_T(1 - c_{3T})\Big(3c_5^2 + 6c_5(1 - c_2) + 2(1 - c_2)^2\Big)}{3c_1^3 c_{1T}^2} \\ & + \frac{c_1(1 - c_3)\Big(3c_{4T}(1 - c_2 + c_5) - (1 - c_{2T})(2 - 2c_2 + 3c_5)\Big)}{3c_1^3 c_{1T}^2}, \end{aligned}$$

$$d_4 = \frac{(1 - c_3)(4 - 4c_3 + 3c_1 c_T(1 - c_{3T}))}{3c_1^2 c_{1T}^2},$$

$$d_5 = \frac{-c_{4T}(2 - 2c_{2T} - c_{4T})\Big(3c_5^2 + 6c_5(1 - c_2) + 2(1 - c_2)^2\Big)}{12c_1^2 c_{1T}^2}, \tag{E1}$$

Appendix F The Algebraic System of 9 Equations of the $k\varepsilon t$ Parameterization

For clarity, we give here the explicit writing of the 9 equations presented in Eqs (6) - (7) and that are the basis of the $k\varepsilon t$ parameterization:

$$
\begin{aligned}
0 &= -c_1\frac{\varepsilon}{k}(\overline{u'^2}-\frac{2}{3}k)+(1-c_2)(-\frac{4}{3}\,\overline{u'w'}\,\partial_z U+\frac{2}{3}\,\overline{v'w'}\,\partial_z V)-\frac{2}{3}(1-c_3)\beta_3\,\overline{w'\theta'}-c_5\overline{u'w'}\,\partial_z U\\
0 &= -c_1\frac{\varepsilon}{k}(\overline{v'^2}-\frac{2}{3}k)+(1-c_2)(-\frac{4}{3}\,\overline{v'w'}\,\partial_z V+\frac{2}{3}\,\overline{u'w'}\,\partial_z U)-\frac{2}{3}(1-c_3)\beta_3\,\overline{w'\theta'}-c_5\overline{v'w'}\,\partial_z V\\
0 &= -c_1\frac{\varepsilon}{k}(\overline{w'^2}-\frac{2}{3}k)+(\frac{2}{3}-\frac{2}{3}c_2+c_5)(\overline{u'w'}\,\partial_z U+\overline{v'w'}\,\partial_z V)+\frac{4}{3}(1-c_3)\beta_3\,\overline{w'\theta'}\\
0 &= -c_1\frac{\varepsilon}{k}\,\overline{u'v'}-(1-c_2)(\overline{v'w'}\,\partial_z U+\overline{u'w'}\,\partial_z V)-\frac{1}{2}c_5(\overline{v'w'}\,\partial_z U+\overline{u'w'}\,\partial_z V)\\
0 &= -c_1\frac{\varepsilon}{k}\,\overline{u'w'}-(1-c_2)\overline{w'^2}\,\partial_z U+(1-c_3)\beta_3\,\overline{u'\theta'}-\frac{1}{2}c_4k\,\partial_z U-\frac{1}{2}c_5(\overline{w'^2}\,\partial_z U-\overline{u'^2}\,\partial_z U-\overline{u'v'}\,\partial_z V)\\
0 &= -c_1\frac{\varepsilon}{k}\,\overline{v'w'}-(1-c_2)\overline{w'^2}\,\partial_z V+(1-c_3)\beta_3\,\overline{v'\theta'}-\frac{1}{2}c_4k\,\partial_z V-\frac{1}{2}c_5(\overline{w'^2}\,\partial_z V-\overline{v'^2}\,\partial_z V-\overline{u'v'}\,\partial_z U)\\
0 &= -c_{1T}\frac{\varepsilon}{k}\,\overline{u'\theta'}-(1-c_{2T}-\frac{1}{2}c_{4T})\,\overline{w'\theta'}\,\partial_z U-\overline{u'w'}\,\partial_z \Theta\\
0 &= -c_{1T}\frac{\varepsilon}{k}\,\overline{v'\theta'}-(1-c_{2T}-\frac{1}{2}c_{4T})\,\overline{w'\theta'}\,\partial_z V-\overline{v'w'}\,\partial_z \Theta\\
0 &= -c_{1T}\frac{\varepsilon}{k}\,\overline{w'\theta'}-\overline{w'^2}\,\partial_z\Theta+(1-c_{3T})\beta_3\,\overline{\theta'^2}-\frac{1}{2}c_{4T}(\overline{u'\theta'}\,\partial_z U+\overline{v'\theta'}\,\partial_z V)
\end{aligned}
\tag{F1}
$$

Appendix G Coefficients of the Stability Functions of the $k\varepsilon t$ Parameterization

We give hereafter the expressions of the coefficients $n_0, n_1, n_2, n_{0T}, n_{1T}, n_{2T}, n^*_{0T}, n^*_{1T}, n^*_{2T}, d_0, d_1, d_2, d_3, d_4, d_5$ of the $k\varepsilon t$ stability functions (Eqs. (21) - (23)). We point out that the expressions of the coefficient n_{1T} and all the coefficients not multiplying α_n (i.e. n_0, n_2, n_{0T}, n_{2T}, d_0, d_1 and d_5) stay unchanged compared to the GLS ones (given in Appendix E).

$$
n_0=\frac{4-4c_2+3c_4}{6c_1},\qquad n_1=\frac{(1-c_3)\Big(2c_{1T}(c_4-c_5)-c_1(2-2c_{2T}-c_{4T})\Big)}{3c_1^2c_{1T}^2},
$$

$$
n_2=\frac{-c_{4T}(4-4c_2+3c_4)(2-2c_{2T}-c_{4T})}{24c_1c_{1T}^2},
$$

$$
n_3=\frac{(1-c_3)(1-c_{3T})\Big(2c_{1T}(4-4c_2+3c_5)+3c_1(2-2c_{2T}-c_{4T})\Big)}{6c_1^2c_{1T}^2},
$$

$$
n_{0T}=\frac{2}{3c_{1T}},\qquad n_{1T}=\frac{2(1-c_3)}{3c_1c_{1T}^2},
$$

$$n_{2T} = \frac{c_1 c_{4T}(4-4c_2+3c_4) + 8c_5 c_{1T}(1-c_2+c_5) - 2c_4 c_{1T}(2-2c_2+3c_5)}{12c_1^2 c_{1T}^2},$$

$$n_{0T}^* = \frac{1-c_{3T}}{c_{1T}}, \qquad n_{1T}^* = \frac{(1-c_3)(1-c_{3T})}{c_1 c_{1T}^2},$$

$$n_{2T}^* = \frac{(1-c_{3T})\Big(3c_5^2 + 6c_5(1-c_2) + 2(1-c_2)^2\Big)}{3c_1^2 c_{1T}},$$

$$d_0 = 1, \qquad d_1 = \frac{7(1-c_3)}{3c_1 c_{1T}}, \qquad d_2 = \frac{3c_5^2 + 6c_5(1-c_2) + 2(1-c_2)^2}{3c_1^2} - \frac{c_{4T}(2-2c_{2T}-c_{4T})}{4c_{1T}^2}$$

$$d_3 = \frac{(1-c_3)\Big(3c_1 c_{4T}(1-c_2+c_5) + c_5 c_{1T}(2-2c_2+c_5) - c_1(1-c_{2T})(2-2c_2+3c_5)\Big)}{3c_1^3 c_{1T}^2}$$

$$d_4 = \frac{4(1-c_3)^2}{3c_1^2 c_{1T}^2}, \qquad d_5 = \frac{-c_{4T}(2-2c_{2T}-c_{4T})\Big(3c_5^2 + 6c_5(1-c_2) + 2(1-c_2)^2\Big)}{12c_1^2 c_{1T}^2}. \tag{G1}$$

Appendix H Results of the $k-\varepsilon-\overline{\theta'^2}-\overline{w'^2}$ model

We detail here the results of the k–ε–$\overline{\theta'^2}$–$\overline{w'^2}$ model which is a possible extension of the $k\varepsilon t$ model where the non-equilibrium is also considered for the $\overline{w'^2}$ equation. By doing that, equations (6) and (7) now form a system of 8 equations with 8 unknowns: $(\overline{u'^2}, \overline{v'^2}, \overline{u'v'}, \overline{u'w'}, \overline{v'w'}, \overline{u'\theta'}, \overline{v'\theta'}, \overline{w'\theta'})$. We solved this system with Mathematica and we obtained the following expressions:

$$\overline{u'w'} = -c_\mu \frac{k^2}{\varepsilon} \partial_z U, \tag{H1}$$

$$\overline{v'w'} = -c_\mu \frac{k^2}{\varepsilon} \partial_z V, \tag{H2}$$

$$\overline{w'\theta'} = -c'_\mu \frac{k^2}{\varepsilon} \partial_z \Theta + c'^*_\mu \frac{k}{\varepsilon} \beta_3 \overline{\theta'^2}, \tag{H3}$$

which have the same shape as the ones found for the $k\varepsilon t$ model. Particularly, even if $\overline{w'^2}$ is not in equilibrium anymore, the velocity fluxes $\overline{u'w'}$ and $\overline{v'w'}$ are still fully downgradient. The expressions of the stability functions c_μ, c'_μ and c'^*_μ are:

$$c_\mu = \frac{n_0 + n_2\alpha_M + n_3\alpha_T + n_4\alpha_W + n_5\alpha_W\alpha_N + n_6\alpha_W\alpha_M}{d_0 + d_1\alpha_N + d_2\alpha_M + d_3\alpha_N\alpha_M + d_5\alpha_M^2}, \tag{H4}$$

$$c'_\mu = \frac{n_{2T}\alpha_M + n_{4T}\alpha_W + n_{5T}\alpha_W\alpha_N + n_{6T}\alpha_W\alpha_M}{d_0 + d_1\alpha_N + d_2\alpha_M + d_3\alpha_N\alpha_M + d_5\alpha_M^2}, \tag{H5}$$

$$c'^*_\mu = \frac{n^*_{0T} + n^*_{1T}\alpha_N + n^*_{2T}\alpha_M}{d_0 + d_1\alpha_N + d_2\alpha_M + d_3\alpha_N\alpha_M + d_5\alpha_M^2}, \tag{H6}$$

with $\alpha_N = \frac{k^2}{\varepsilon^2}N^2$, $\alpha_M = \frac{k^2}{\varepsilon^2}M^2$, $\alpha_T = \frac{k}{\varepsilon^2}\beta_3^2\,\overline{\theta'^2}$, and $\alpha_W = \frac{1}{k}\,\overline{w'^2}$. Coefficients n_i, n_{iT} and d_i depends on the coefficients c_i and c_{iT}; the expressions are given hereafter. Taking the values of the c_i and c_{iT} given in Table 1, the stability functions are approximately as follows

$$c_\mu = \frac{0.04693 - 0.00005303\alpha_M + 0.001996\alpha_T + 0.0896\alpha_W - 0.002994\alpha_W\alpha_N - 0.0001012\alpha_W\alpha_M}{1 + 0.03361\alpha_N + 0.01342\alpha_M + 0.00006267\alpha_N\alpha_M - 0.00001644\alpha_M^2}, \tag{H7}$$

$$c'_\mu = \frac{0.0002651\alpha_M + 0.1681\alpha_W + 0.005649\alpha_W\alpha_N + 0.002952\alpha_W\alpha_M}{1 + 0.03361\alpha_N + 0.01342\alpha_M + 0.00006267\alpha_N\alpha_M - 0.00001644\alpha_M^2}, \tag{H8}$$

$$c'^*_\mu = \frac{0.1120 + 0.003766\alpha_N + 0.001631\alpha_M}{1 + 0.03361\alpha_N + 0.01342\alpha_M + 0.00006267\alpha_N\alpha_M - 0.00001644\alpha_M^2}. \tag{H9}$$

Here are the expressions of the coefficients n_i, n_{iT}, and d_i:

$$n_0 = \frac{3c_4 - 2c_5}{6c_1}, \qquad n_2 = \frac{-c_{4T}(3c_4 - 2c_5)(2 - 2c_{2T} - c_{4T})}{24c_1c_{1T}^2},$$

$$n_3 = \frac{(1-c_3)(1-c_{3T})\Big(2c_{1T}c_5 + 3c_1(2 - 2c_{2T} - c_{4T})\Big)}{6c_1^2c_{1T}^2}, \qquad n_4 = \frac{2 - 2c_2 + c_5}{2c_1},$$

$$n_5 = \frac{-(1-c_3)\Big(2c_{1T}c_5 + 3c_1(2 - 2c_{2T} - c_{4T})\Big)}{6c_1^2c_{1T}^2}, \qquad n_6 = \frac{-c_{4T}(2 - 2c_2 + c_5)(2 - 2c_{2T} - c_{4T})}{8c_1c_{1T}^2},$$

$$n_{2T} = \frac{c_{4T}(3c_4 - 2c_5)}{12c_1c_{1T}^2}, \qquad n_{4T} = \frac{1}{c_{1T}}, \qquad n_{5T} = \frac{1-c_3}{c_1c_{1T}^2},$$

$$n_{6T} = \frac{3c_1c_{4T}(2-2c_2+c_5) + 2c_5c_{1T}(4-4c_2+3c_5)}{12c_1^2c_{1T}^2},$$

$$n_{0T}^* = \frac{1-c_{3T}}{c_{1T}}, \qquad n_{1T}^* = \frac{(1-c_3)(1-c_{3T})}{c_1c_{1T}^2}, \qquad n_{2T}^* = \frac{c_5(1-c_{3T})(4-4c_2+3c_5)}{6c_1^2c_{1T}},$$

$$d_0 = 1, \qquad d_1 = \frac{1-c_3}{c_1c_{1T}}, \qquad d_2 = \frac{c_5(4-4c_2+3c_5)}{6c_1^2} - \frac{c_{4T}(2-2c_{2T}-c_{4T})}{4c_{1T}^2}$$

$$d_3 = \frac{c_5c_{4T}(1-c_3)}{6c_1^2c_{1T}^2}, \qquad d_5 = \frac{-c_5c_{4T}(4-4c_2+3c_5)(2-2c_{2T}-c_{4T})}{24c_1^2c_{1T}^2}. \tag{H10}$$

Acknowledgments

We thank Florian Lemarié for providing the 1D code. All the computations presented in this paper were performed using the GRICAD infrastructure (https://gricad.univ-grenoble-alpes.fr), which is supported by Grenoble research communities.

References

Arakawa, A., & Schubert, W. H. (1974). Interaction of a cumulus cloud ensemble with the large-scale environment, part I. *J. Atmos. Sci.*, *31*(3), 674–701. doi: 10.1175/1520-0469(1974)031⟨0674:IOACCE⟩2.0.CO;2

Burchard, H. (2002). *Applied Turbulence Modelling in Marine Waters* (Vol. 100; S. Bhattacharji, G. M. Friedman, H. J. Neugebauer, & A. Seilacher, Eds.). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/3-540-45419-5

Burchard, H., & Baumert, H. (1995). On the performance of a mixed-layer model based on the $\kappa - \epsilon$ turbulence closure. *J. Geophys. Res.*, *100*(C5), 8523–8540. doi: 10.1029/94JC03229

Burchard, H., & Bolding, K. (2001). Comparative Analysis of Four Second-Moment Turbulence Closure Models for the Oceanic Mixed Layer. *Journal of Physical Oceanography*, *31*(8), 1943–1968. doi: 10.1175/1520-0485(2001)031⟨1943:CAOFSM⟩2.0.CO;2

Burchard, H., & Petersen, O. (1999). Models of turbulence in the marine environment —a comparative study of two-equation turbulence models. *Journal of Marine Systems*, *21*(1-4), 29–53. doi: 10.1016/S0924-7963(99)00004-4

Canuto, V. M., Howard, A., Cheng, Y., & Dubovikov, M. S. (2001). Ocean Turbulence. Part I: One-Point Closure Model—Momentum and Heat Vertical Diffusivities. *Journal of Physical Oceanography*, *31*(6), 1413–1426. doi: 10.1175/1520-0485(2001)031⟨1413:OTPIOP⟩2.0.CO;2

Cheng, Y., Canuto, V. M., Howard, A. M., Ackerman, A. S., Kelley, M., Fridlind, A. M., ... Elsaesser, G. S. (2020). A Second-Order Closure Turbulence Model: New Heat Flux Equations and No Critical Richardson Number. *Journal of the Atmospheric Sciences*, *77*(8), 2743–2759. doi: 10.1175/JAS-D-19-0240.1

Deardorff, J. W. (1972). Theoretical expression for the countergradient vertical heat flux. *J. Geophys. Res.*, *77*(30), 5900–5904. doi: 10.1029/JC077i030p05900

Emanuel, K. A. (1991). A scheme for representing cumulus convection in large-scale models. *J. Atmos. Sci.*, *48*(21), 2313–2329. doi: 10.1175/1520-0469(1991)048⟨2313:ASFRCC⟩2.0.CO;2

Fearon, G., Herbette, S., Veitch, J., Cambon, G., Lucas, A. J., Lemarié, F., & Vichi, M. (2020). Enhanced vertical mixing in coastal upwelling systems driven by diurnal-inertial resonance: Numerical experiments. *Journal of Geophysical Research: Oceans*, *125*(9), e2020JC016208. doi: 10.1029/2020JC016208

Fox-Kemper, B., Ferrari, R., & Hallberg, R. (2008). Parameterization of mixed layer eddies. part I: Theory and diagnosis. *J. Phys. Oceanogr.*, *38*(6), 1145. doi: 10.1175/2007JPO3792.1

Garanaik, A., Pereira, F. S., Smith, K., Robey, R., Li, Q., Pearson, B., & Van Roekel, L. (2024). A New Hybrid Mass-Flux/High-Order Turbulence Closure for Ocean Vertical Mixing. *Journal of Advances in Modeling Earth Systems*, *16*(1), e2023MS003846. doi: 10.1029/2023MS003846

Garcia, J. R., & Mellado, J. P. (2014). The two-layer structure of the entrainment zone in the convective boundary layer. *J. Atmos. Sci.*, *71*(6), 1935–1955. doi:

10.1175/JAS-D-13-0148.1

Gaspar, P., Grégoris, Y., & Lefevre, J.-M. (1990). A simple eddy kinetic energy model for simulations of the oceanic vertical mixing: Tests at station papa and long-term upper ocean study site. *J. Geophys. Res.*, *95*, 16. doi: 10.1029/JC095iC09p16179

Ghannam, K., Duman, T., Salesky, S. T., Chamecki, M., & Katul, G. (2017). The non-local character of turbulence asymmetry in the convective atmospheric boundary layer. *Quarterly Journal of the Royal Meteorological Society*, *143*(702), 494–507. doi: 10.1002/qj.2937

Gibbs, J. A., Fedorovich, E., & Van Eijk, A. M. J. (2011). Evaluating Weather Research and Forecasting (WRF) Model Predictions of Turbulent Flow Parameters in a Dry Convective Boundary Layer. *Journal of Applied Meteorology and Climatology*, *50*, 2429–2444.

Giordani, H., Bourdallé-Badie, R., & Madec, G. (2020). An eddy-diffusivity mass-flux parameterization for modeling oceanic convection. *J. Adv. Model. Earth Syst.*, *12*(9), e02078. doi: 10.1029/2020MS002078

Hanjalić, K., & Launder, B. E. (1972). A Reynolds stress model of turbulence and its application to thin shear flows. *Journal of Fluid Mechanics*, *52*(4), 609–638. doi: 10.1017/S002211207200268X

Hourdin, F., Couvreux, F., & Menut, L. (2002). Parameterization of the dry convective boundary layer based on a mass flux representation of thermals. *J. Atmos. Sci.*, *59*(6), 1105–1123. doi: 10.1175/1520-0469(2002)059⟨1105:POTDCB⟩2.0.CO;2

Large, W. G., McWilliams, J. C., & Doney, S. C. (1994). Oceanic vertical mixing: a review and a model with a nonlocal boundary layer parameterization. *Rev. Geophys.*, *32*, 363–404. doi: 10.1029/94RG01872

Lazier, J. (2001). Deep Convection. In *Encyclopedia of Ocean Sciences* (pp. 634–643). Elsevier. doi: 10.1006/rwos.2001.0113

Legay, A., Deremble, B., Penduff, T., Brasseur, P., & Molines, J.-M. (2024). A framework for evaluating ocean mixed layer depth evolution. *J. Adv. Model. Earth Syst.*, submitted. doi: 10.22541/essoar.168563421.17506622/v2

Lemarié, F., Samson, G., Redelsperger, J.-L., Giordani, H., Brivoal, T., & Madec, G. (2021). A simplified atmospheric boundary layer model for an improved representation of air–sea interactions in eddying oceanic models: implementation and first evaluation in NEMO (4.0). *Geoscientific Model Development*, *14*(1), 543–572. doi: 10.5194/gmd-14-543-2021

Luyten, J., Pedlosky, J., & Stommel, H. (1983). The ventilated thermocline. *J. Phys. Oceanogr.*, *13*(2), 292–309.

Marshall, J., & Schott, F. (1999). Open-ocean convection: Observations, theory, and models. *Reviews of geophysics*, *37*(1), 1–64. doi: 10.1029/98RG02739

Mellor, G. L., & Yamada, T. (1982). Development of a turbulence closure model for geophysical fluid problems. *Reviews of Geophysics*, *20*(4), 851. doi: 10.1029/RG020i004p00851

Mironov, D. V., Gryanik, V. M., Moeng, C., Olbers, D. J., & Warncke, T. H. (2000). Vertical turbulence structure and second-moment budgets in convection with rotation: A large-eddy simulation study. *Quarterly Journal of the Royal Meteorological Society*, *126*(563), 477–515. doi: 10.1002/qj.49712656306

Patankar, S. (1980). *Numerical heat transfer and fluid flow.* New York: McGraw-Hill.

Popinet, S. (2020). Basilisk flow solver and pde library. *available at available at http://basilisk. fr*.

Reffray, G., Bourdalle-Badie, R., & Calone, C. (2015). Modelling turbulent vertical mixing sensitivity using a 1-D version of NEMO. *Geoscientific Model Development*, *8*(1), 69–86. doi: 10.5194/gmd-8-69-2015

Reichl, B. G., & Hallberg, R. (2018). A simplified energetics based planetary boundary layer (ePBL) approach for ocean climate simulations. *Ocean Model.*, *132*, 112-129. doi: 10.1016/j.ocemod.2018.10.004

Rodi, W. (1987). Examples of calculation methods for flow and mixing in stratified fluids. *Journal of Geophysical Research*, *92*(C5), 5305. doi: 10.1029/JC092iC05p05305

Souza, A. N., Wagner, G. L., Ramadhan, A., Allen, B., Churavy, V., Schloss, J., ... Ferrari, R. (2020). Uncertainty quantification of ocean parameterizations: Application to the K-Profile-Parameterization for penetrative convection. *J. Adv. Model. Earth Syst.*, *12*(12), e2020MS002108. doi: 10.1029/2020MS002108

Stull, R. B. (1988). *An introduction to boundary layer meteorology.* doi: 10.1007/978-94-009-3027-8

Thangam, S., Abid, R., & Speziale, C. G. (1992). Application of a new K-tau model to near wall turbulent flows. *AIAA Journal*, *30*(2), 552–554. doi: 10.2514/3.10952

Treguier, A. M., de Boyer Montégut, C., Bozec, A., Chassignet, E. P., Fox-Kemper, B., Hogg, A. M., ... Yeager, S. (2023). The mixed layer depth in the ocean model intercomparison project (omip): Impact of resolving mesoscale eddies. *EGUsphere*, *2023*, 1–43. doi: 10.5194/egusphere-2023-310

Troen, I. B., & Mahrt, L. (1986). A simple model of the atmospheric boundary layer; sensitivity to surface evaporation. *Bound.-Lay. Meteorol.*, *37*(1-2), 129-148. doi: 10.1007/BF00122760

Umlauf, L., & Burchard, H. (2003). A generic length-scale equation for geophysical turbulence models. *Journal of Marine Research*, *61*(2), 235–265. doi: 10.1357/002224003322005087

Umlauf, L., & Burchard, H. (2005). Second-order turbulence closure models for geophysical boundary layers. A review of recent work. *Continental Shelf Research*, *25*(7-8), 795–827. doi: 10.1016/j.csr.2004.08.004

Umlauf, L., Burchard, H., & Hutter, K. (2003). Extending the k–omega turbulence model towards oceanic applications. *Ocean Modelling*, *5*(3), 195–218. doi: 10.1016/S1463-5003(02)00039-2

Van Roekel, L., Adcroft, A. J., Danabasoglu, G., Griffies, S. M., Kauffman, B., Large, W., ... Schmidt, M. (2018). The KPP Boundary Layer Scheme for the Ocean: Revisiting Its Formulation and Benchmarking One-Dimensional Simulations Relative to LES. *Journal of Advances in Modeling Earth Systems*, *10*(11), 2647–2685. doi: 10.1029/2018MS001336

Wagner, G., Hillier, A., Constantinou, N. C., Silvestri, S., Souza, A., Burns, K., ... Ferrari, R. (2023). CATKE: a turbulent-kinetic-energy-based parameterization for ocean microturbulence with dynamic convective adjustment. *arXiv e-prints*, arXiv:2306.13204. doi: 10.48550/arXiv.2306.13204

Warner, J. C., Sherwood, C. R., Arango, H. G., & Signell, R. P. (2005). Performance of four turbulence closure models implemented using a generic length scale method. *Ocean Modelling*, *8*(1-2), 81–113. doi: 10.1016/j.ocemod.2003.12.003

Wilcox, D. C. (1988). Reassessment of the scale-determining equation for advanced turbulence models. *AIAA Journal*, *26*(11), 1299–1310. doi: 10.2514/3.10041

Williams, R. G., Marshall, J. C., & Spall, M. A. (1995). Does stommel's mixed layer "demon" work? *J. Phys. Oceanogr.*, *25*(12), 3089–3102. doi: 10.1175/1520-0485(1995)025⟨3089:DSMLW⟩2.0.CO;2

Willis, G., & Deardorff, J. (1974). A laboratory model of the unstable planetary boundary layer. *Journal of Atmospheric Sciences*, *31*(5), 1297–1307. doi: 10.1175/1520-0469(1974)031⟨1297:ALMOTU⟩2.0.CO;2

Zeierman, S., & Wolfshtein, M. (1986). Turbulent time scale for turbulent-flow calculations. *AIAA Journal*, *24*(10), 1606–1610. doi: 10.2514/3.9490

Zhou, B., Sun, S., Yao, K., & Zhu, K. (2018). Reexamining the Gradient and Countergradient Representation of the Local and Nonlocal Heat Fluxes in the Convective Boundary Layer. *Journal of the Atmospheric Sciences*, *75*(7), 2317–2336. doi: 10.1175/JAS-D-17-0198.1

Derivation and implementation of a non-gradient term to improve the oceanic convection representation within the $k - \varepsilon$ parameterization

Alexandre Legay[1], Bruno Deremble[1], Hans Burchard[2]

[1]Univ. Grenoble Alpes, CNRS, INRAE, IRD, Grenoble INP, IGE, Grenoble, France
[2]Leibniz Institute for Baltic Sea Research Warnemünde, Rostock, Germany

Key Points:

- Analytical derivation of a non-gradient term within the $k-\varepsilon$ model for improving the oceanic convection representation
- Comparison with large eddy simulations in both wind-forced and buoyancy-driven regimes confirms the improvement due to the non-gradient term
- The vertical profile of the non-gradient term is compared to the one of the KPP the non-local term

Corresponding author: Alexandre Legay, `alexandre.legay@univ-grenoble-alpes.fr`

Abstract

The representation of turbulent fluxes during oceanic convective events is important to capture the evolution of the oceanic mixed layer. To improve the accuracy of turbulent fluxes, we examine the possibility of adding a non-gradient component in their expression in addition to the usual downgradient part. To do so, we extend the $k - \varepsilon$ algebraic second-moment closure by relaxing the assumption on the equilibrium of the temperature variance $\overline{\theta'^2}$. With this additional transport equation for the temperature variance, we obtain a $k{-}\varepsilon{-}\overline{\theta'^2}$ model (the "$k\varepsilon t$" model) which includes a non-gradient term for the temperature flux. We validate this new model against Large Eddy Simulations (LES) in both wind-forced and buoyancy-driven regimes. In both cases, we find that the vertical profile of temperature is well captured by the $k\varepsilon t$ model. Particularly, for the buoyancy-driven regime, the non-gradient term increases the portion of the mixed layer which is stably stratified. This is an improvement since this portion is too small with the $k{-}\varepsilon$ parameterization. Finally, comparison of the non-gradient term with the KPP non-local term gives insights for refining the KPP's ad hoc shape polynomial.

Plain Language Summary

In the ocean, vertical mixing of water occurs when cold air temperatures create dense cold water at the surface that tends to sink in the ocean or when a strong wind induces turbulence at the ocean surface. In numerical models, the classic approach to represent this vertical mixing is to consider that it is done entirely by diffusion. This means that the heat always goes from the warm water to the cold water, i.e. in the opposite direction of the gradient of the temperature. However, during intense events called "convection", some cold water parcels created at the top of the ocean can have enough thermal inertia and velocity to flow against the direction of the mean temperature gradient. This kind of phenomenon is often referred to coherent eddies or non-local turbulence. In this article, we perform an analytical derivation to give a mathematical expression of the impact of non-local mixing. We then compare our new model with more realistic three-dimensional models of convection and conclude that the new term derived here is important to reproduce the vertical profile of temperature in the ocean.

1 Introduction

In the realm of climate modeling, the oceanic mixed layer plays a critical role because it is responsible for regulating the oceanic heat uptake and carbon storage. Throughout much of the year, the mixed layer operates as a dynamic buffer, intimately interacting with the atmosphere. However, it is in late winter that the true importance of this layer becomes evident. In late winter, the mixed layer is deepest and direct contact is established with the deep ocean: it is during this period that the ocean effectively stores heat and CO2 (a mechanism sometimes pictured as Stommel's demon, see Luyten et al., 1983; Williams et al., 1995). Accurately representing the mixed layer is thus crucial because it directly affects our ability to make accurate predictions about future climate patterns (Treguier et al., 2023).

The depth of the mixed layer changes in response to various factors: it deepens when turbulent mixing is triggered by the mechanical effect of the wind and/or waves; or triggered by buoyancy effects: heat flux (cooling) and freshwater flux (evaporation, sea-ice formation). Conversely, the mixed layer becomes shallower typically during calm weather where there is less turbulence and restratifying mixed layer instabilities can develop, or when there is a stabilizing buoyancy flux due to warming (e.g. sunny condition) and/or freshwater input (e.g. precipitation, sea-ice melt, or river discharge). This restratification allows the surface layer to separate from the denser, deeper water (Stull, 1988). The explicit representation of the small-scale turbulence causing the mixing occurring in the mixed layer is of course impossible in climate models where the horizontal grid is often on the order of tens of kilometers. Instead, the ocean modeling community has developed parameterizations whose goal is to represent the mean effect of the turbulent fluctuations (Gaspar et al., 1990; Large et al., 1994; Burchard & Bolding, 2001; Umlauf & Burchard, 2003; Fox-Kemper et al., 2008; Reichl & Hallberg, 2018). The main purpose of a mixed layer parameterization is to propose a closure for the turbulent vertical fluxes $\overline{w'x'}$, where w' is the turbulent vertical velocity, x' the turbulent fluctuation of a property x (momentum, temperature, salinity, phytoplankton, etc...) and the overline denotes the ensemble averaging over small-scale fluctuations (see Stull, 1988). These turbulent fluxes, and all the other covariances $\overline{x'y'}$, are called the second-order moments. The traditional approach to close this problem consists of expressing these turbulent fluxes as a function of the vertical gradient of the mean property $X = \overline{x}$ (i.e. a downgradient parameterization), as shown here for the temperature

$$\overline{w'\theta'} = -K_t \partial_z \Theta\,, \tag{1}$$

with K_t being an eddy diffusivity coefficient. Among all the possibilities to compute K_t we would like to emphasize the Generic Length Scale (GLS) approach (Umlauf & Burchard, 2003) and more precisely the $k-\varepsilon$ closure (Burchard & Bolding, 2001). This closure consists in deriving two equations: one for the evolution of turbulent kinetic energy k, and one for dissipation ε. The downgradient formulation (1) also results from more complex algebraic second-moment closures even if it is not assumed a priori (Burchard & Baumert, 1995). The eddy diffusivity is obtained analytically and is a function of turbulent kinetic energy, dissipation, buoyancy frequency, and shear frequency. While this eddy diffusivity approach has been successfully applied in the oceanic and atmospheric modeling communities, it has also been quickly recognized that the shape of the temperature profile during a convective event is not well captured by this closure. In fact, Deardorff (1972) was among the first to realize that after a convective event, the stratification profile in the mixed layer is not neutral as one would expect for a perfectly well-mixed layer but is instead slightly stable. To illustrate this observation, we plot in Figure 1.a the typical shape of a normalized temperature profile in the mixed layer from a numerical model that explicitly resolves convection (see Mironov et al. (2000); details about the normalization are provided henceforth; we only wish to focus here on the shape of the temperature profile). This profile can be decomposed into two well-defined zones. Just below the air-sea interface, there is an unstable zone with cold water above warmer water ($\partial_z\Theta < 0$). Such layer is sometimes called the *thermal layer* (Lazier, 2001) and we define it here as the layer between the surface and the depth h_t at which $\partial_z\Theta = 0$. Below that depth h_t, we find the *convective layer*; a slightly stable layer that extends until the base of the mixed layer h_m. Both layers form the *mixed layer*. The position of h_t has been documented to be near $z = -0.4h_m$ (see Zhou et al., 2018) such that more than half of the mixed layer is stably stratified. The presence of such stable stratification in the convective layer has been attributed to downward propagating plumes which remain coherent during their descent and deposit their negative buoyancy anomaly at their neutral level, thus creating a stable stratification (see Arakawa and Schubert (1974) or Emanuel (1991) for the atmospheric scenario).

Several options have emerged in the literature to reproduce this vertical temperature profile with a stable stratification. The atmospheric community has favored the

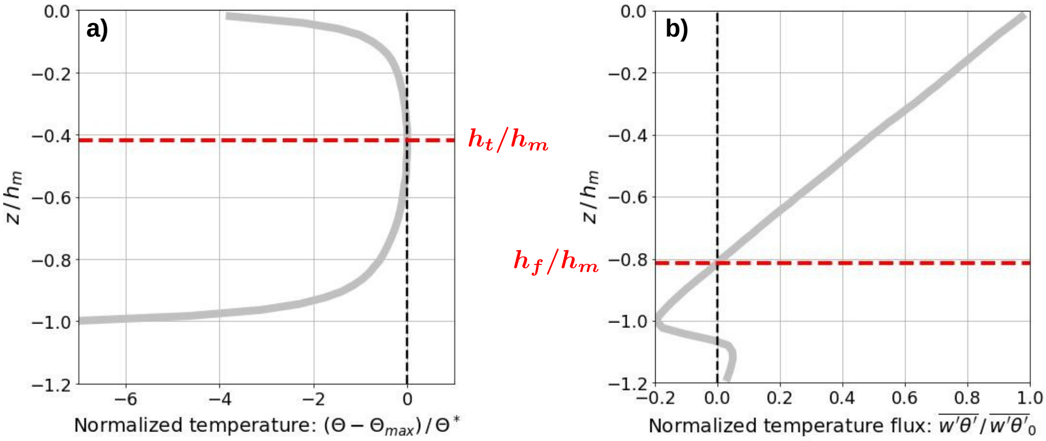

Figure 1. Normalized profile from LES data of Mironov et al. (2000) of a) the temperature and b) the vertical turbulent temperature flux. The depth is normalized by the mixed layer depth h_m, defined here as the minimum of the temperature flux. The temperature flux is normalized by its surface value $\overline{w'\theta'}|_0$. The temperature is normalized in $(\Theta - \Theta_{max})\,/\,\Theta^*$ with Θ_{max} the maximum of the temperature over the vertical and $\Theta^* = \overline{w'\theta'}|_0/w^*$ a scaling of the temperature, with $w^* = (\overline{w'\theta'}|_0\, h_m)^{1/3}$ a scaling of the velocity of the convective thermals (Willis & Deardorff, 1974; Marshall & Schott, 1999). Red dashed lines highlight the location h_t of the zero of the gradient $\partial_z\Theta$ and the location h_f of the zero of the temperature flux.

use of a mass flux parameterization which simulates the vertical movement of air parcels within convective clouds. It represents the ascent and descent of parcels, which transport heat, moisture, and other properties. These mass flux parameterizations have recently been introduced in ocean models (Giordani et al., 2020; Garanaik et al., 2024). Another, perhaps more ancient, approach taken by Large et al. (1994) was to add a positive non-gradient term $\mathit{\Gamma}$ in the parameterization of the flux in Equation (1): (see also Troen and Mahrt (1986); or Burchard and Petersen (1999) where the problem of missing non-gradient fluxes in downgradient parameterization is stated),

$$\overline{w'\theta'} = -K_t \partial_z \Theta + \mathit{\Gamma}. \tag{2}$$

$\mathit{\Gamma}$ being positive, it represents a positive turbulent temperature flux, i.e. a flux that follows the buoyancy effect (cold going down and hot going up). $\mathit{\Gamma}$ can thus be viewed as representing coherent structures ("non-local eddies", "coherent thermals") that are subjected to the buoyancy force. Particularly, we see in equation 2 that $\mathit{\Gamma}$ allows to keep a positive turbulent temperature flux in situations of neutral ($\partial_z \Theta = 0$) or slightly stable ($\partial_z \Theta > 0$) temperature profiles. In other words, this means that, in stably-stratified conditions, coherent structures can be strong enough to counter the downgradient flux that acts in a counter-buoyancy direction. Note that this term was often written $\overline{w'\theta'} = -K_t(\partial_z \Theta - \gamma)$ with $\gamma = \mathit{\Gamma}/K_t$ (e.g. Deardorff, 1972; Large et al., 1994). In this formulation, γ corresponds to the maximal stable stratification where a positive turbulent temperature flux can be maintained even if the downgradient flux generates a counter-buoyancy effect. In Large et al. (1994), $\mathit{\Gamma}$ was defined with some constraints: to be zero at the surface and at the base of the mixed layer such that it is merely a redistribution of heat. The magnitude and the exact shape of this term were however chosen in a relatively ad hoc way to respect some empirical rules of convection.

The term $\mathit{\Gamma}$ was often referred in the literature as a "non-local" term (Large et al., 1994; Ghannam et al., 2017) or a countergradient term (Deardorff, 1972; Troen & Mahrt, 1986; Gibbs et al., 2011). As we mentioned before, denomination "non-local" refers to the fact that it is supposed to represent non-local eddies (coherent thermals). However Zhou et al. (2018) argued that this often-implied association of the non-gradient term to the non-local eddies is partially wrong. "Non-local" can also indicate that the value of this term at a specific depth does not depend exclusively on properties evaluated at this depth. For example, in KPP, this term depends on the surface heat flux and on the total mixed layer thickness. The other used denomination, "countergradient", refers to the fact that, in the lower part of the mixed layer which is stable, this term acts with an opposite sign compared to the mean gradient. However, in the upper part of the mixed layer which is unstable, the denomination "countergradient" is very unsettling since this term acts as if it were a downgradient term. For these reasons, we will call this term "non-gradient", a more neutral denomination.

A key aspect of the addition of the non-gradient term is to relax the downgradient dependence and particularly the constraint that the depth at which $\overline{w'\theta'}$ vanishes is equal to the depth at which the gradient of the temperature profile vanishes (see Eq. 1). To better understand why this matters, we plot in Figure 1.b the vertical turbulent heat flux $\overline{w'\theta'}$ obtained in the same numerical model as presented before (Mironov et al., 2000). In this figure, we recover the traditional form of a linear decrease from the surface value (which corresponds to the magnitude of the surface flux) to a cancellation near the bottom of the mixed layer, which has been observed and described in several places (e.g. Large et al., 1994; Burchard & Bolding, 2001; Van Roekel et al., 2018). The exact depth at which the heat flux vanishes depends on the surface boundary conditions (wind and heat fluxes) but it has been documented to be close to $h_f = -0.8h_m$ at least in the free convection scenario (Garcia & Mellado, 2014). There is thus an obvious discrepancy between $h_t = -0.4h_m$ and $h_f = -0.8h_m$ such that Equation (1) cannot hold in most of the mixed layer and the addition of an extra term in the definition of the flux is physically relevant. Even if there is a consensus on the need to add a non-gradient component in the definition of the flux, the exact formulation of this flux remains a matter of debate. To develop a framework that is accurate, robust, and consistent with existing parameterizations, we have opted to focus on extending the $k - \varepsilon$ parameterization.

We first perform an analytical derivation of the non-gradient term. Since Deardorff (1972) and Cheng et al. (2020), we know that the non-gradient term is somehow related to the small-scale temperature variance $\overline{\theta'^2}$. We will therefore derive a second-moment closure that uses a full transport equation for the temperature variance $\overline{\theta'^2}$, in addition to the second-moment transport equations for k and ε, thus extending the $k-\varepsilon$ model to a $k - \varepsilon - \overline{\theta'^2}$ model (henceforth called the "$k\varepsilon t$" model). In this model, we get an analytical expression of a non-gradient term that shares several properties with the KPP non-local term: it is positive, and vanishes at the surface and at the bottom of the mixed layer. Last, we test the numerical implementation of $k\varepsilon t$ against Large Eddy Simulations (LES) and further compare its results to the predictions of a standard $k{-}\varepsilon$ model and KPP simulations.

2 Derivation and Implementation of the $k\varepsilon t$ Parameterization

This section introduces the second-order moments equations. We recall the hypotheses made in the GLS model to solve this system of equations. Then, we explain how we derive the $k\varepsilon t$ parameterization in the same formalism.

2.1 Formalism and Second-Order Moments Equations

The Reynolds Averaged Navier Stokes (RANS) equations used in ocean models are written for the mean velocities $\boldsymbol{U} = (U, V, W)$ and the mean temperature Θ. As in the original derivation of the $k{-}\varepsilon$ model, we consider here only one active tracer (temperature) that enters the equation of state. The RANS equations include the effect of turbulent fluctuations through the second-order moments $\overline{u_i'u_j'}$ and $\overline{u_i'\theta'}$. To close the system, we need to provide equations for these moments. We focus here on the procedure derived in Burchard and Bolding (2001). After adopting their closure assumptions for non-closed terms, and neglecting the rotational and viscous effects, the equations of second-order moments are

$$\begin{aligned}
\partial_t\overline{u_i'u_j'} + \partial_l(U_l\overline{u_i'u_j'} + \overline{u_i'u_j'u_l'}) = &- c_1\frac{\varepsilon}{k}(\overline{u_i'u_j'} - \frac{2}{3}\delta_{ij}k)\\
&+ P_{ij} - c_2(P_{ij} - \frac{2}{3}\delta_{ij}P)\\
&+ B_{ij} - c_3(B_{ij} - \frac{2}{3}\delta_{ij}B)\\
&- c_4kS_{ij}\\
&- c_5Z_{ij}\\
&- \frac{2}{3}\delta_{ij}\varepsilon,
\end{aligned} \tag{3}$$

$$\begin{aligned}
\partial_t\overline{u_i'\theta'} + \partial_j(U_j\overline{u_i'\theta'} + \overline{u_i'u_j'\theta'}) = &- c_{1T}\frac{\varepsilon}{k}\overline{u_i'\theta'}\\
&- (1 - c_{2T})\,\overline{u_j'\theta'}\,\partial_jU_i - \overline{u_i'u_j'}\,\partial_j\Theta\\
&+ (1 - c_{3T})\,\beta_i\,\overline{\theta'^2}\\
&+ c_{4T}\,\overline{u_j'\theta'}\,V_{ij},
\end{aligned} \tag{4}$$

$$\partial_t \overline{\theta'^2} + \partial_j (U_j \overline{\theta'^2} + \overline{u'_j \theta'^2}) = -2\,\overline{u'_j \theta'}\,\partial_j \Theta - 2\frac{1}{c_T}\frac{\varepsilon}{k}\,\overline{\theta'^2}, \tag{5}$$

with

- $P_{ij} = -\partial_l U_i\,\overline{u'_l u'_j} - \partial_l U_j\,\overline{u'_l u'_i}$: Production/destruction of $\overline{u'_i u'_j}$ by the shear
- $B_{ij} = \beta_i\,\overline{u'_j \theta'} + \beta_j\,\overline{u'_i \theta'}$: Production of $\overline{u'_i u'_j}$ by the buoyancy
- $S_{ij} = \frac{1}{2}(\partial_i U_j + \partial_j U_i)$: Shear tensor
- $V_{ij} = \frac{1}{2}(\partial_i U_j - \partial_j U_i)$: Vorticity tensor
- $Z_{ij} = V_{il}(\overline{u'_l u'_j} - \frac{2}{3}\delta_{lj} k) + V_{jl}(\overline{u'_l u'_i} - \frac{2}{3}\delta_{li} k)$: Symmetric tensor associated to the vorticity
- $k = \frac{1}{2}(\overline{u'^2} + \overline{v'^2} + \overline{w'^2})$: Turbulent Kinetic Energy (TKE)
- $P = \frac{1}{2}P_{ii}$: Production of TKE by the shear
- $B = \frac{1}{2}B_{ii}$: Production/destruction of TKE by the buoyancy
- ε: Dissipation of TKE

Further definitions are δ_{ij} the Kronecker delta, $\boldsymbol{\beta} = (0, 0, \alpha g)$, α the thermal expansion coefficient and g the gravitational acceleration. In the equations, the Einstein summation convention is adopted.

Coefficients c_1, c_2, c_3, c_4, c_5 are empirical coefficients for the parameterization of the pressure-velocity correlation tensor $\Pi_{ij} = \overline{u'_i \partial_j p} + \overline{u'_j \partial_i p}$, coefficients $c_{1T}, c_{2T}, c_{3T}, c_{4T}$ are for the parameterization of the pressure-temperature correlations $\Pi_i^\theta = \overline{\theta' \partial_i p}$, and c_T for the parameterization of the temperature variance dissipation. Further details about these parameterizations can be found in Canuto et al. (2001). We report the values of these coefficients in Table 1. These values are the ones of Canuto et al. (2001) model A, converted into the notations used here (it is the same as the values reported in Table 1 of Burchard and Bolding (2001) except for minor typos on c_3 and c_4 that have been identified. Exact formulations of these coefficients are given in Appendix B).

We are now going to explain the classic procedure used in the GLS models for solving the system, where the new model differs and what are the consequences.

Table 1. Values of the coefficients appearing in the second-order moment equations

c_1	c_2	c_3	c_4	c_5	c_{1T}	c_{2T}	c_{3T}	c_{4T}	c_T
2.5	0.984	0.5	0.512	0.416	5.95	0.6	0.33	0.4	1.44

2.2 GLS Procedure

The GLS procedure is as follows. Firstly, we consider the boundary layer approximation where the vertical scale is much less than the horizontal scale. Horizontal gradients are then neglected in comparison to the vertical gradients. A direct consequence is the simplification of the continuity equation in $\partial_z W = 0$. The resulting expressions of the tensors P_{ij}, B_{ij}, S_{ij}, V_{ij} and Z_{ij} are given in Appendix C.

Secondly, we consider that the moments $\overline{u_i'\theta'}$ and $\overline{\theta'^2}$ are in local equilibrium, meaning that the sum of the time variations, the advective transports and the turbulent transports of these moments is zero (i.e. the left-hand sides of equations (4) and (5) are zero). Concerning the moments $\overline{u_i'u_j'}$, the trick is to not make this assumption directly for $\overline{u_i'u_j'}$ but rather to the anisotropic part of these moments $\overline{u_i'u_j'}-2/3\,\delta_{ij}k$ to keep the time variation and the transports of the TKE to be non-zero. These assumptions correspond to the level $2\frac{1}{2}$ in the hierarchy of models proposed by Mellor and Yamada (1982). This hierarchy has been derived with scaling arguments based on the level of anisotropy of every term. The scaling at level 3 results naturally in neglecting transports and time variations for $\overline{u_i'u_j'} - 2/3\,\delta_{ij}k$ and $\overline{u_i'\theta'}$. However, neglecting these terms for the $\overline{\theta'^2}$ equation is not justified by the scaling process and is much more an ad hoc practical hypothesis that results in obtaining this so-called level $2\frac{1}{2}$ in which the system of equations is now algebraic. Indeed, we obtain the following set of equations

$$0 = -c_1\frac{\varepsilon}{k}(\overline{u_i'u_j'}-\frac{2}{3}\delta_{ij}k)+(1-c_2)(P_{ij}-\frac{2}{3}\delta_{ij}P)+(1-c_3)(B_{ij}-\frac{2}{3}\delta_{ij}B)-c_4kS_{ij}-c_5Z_{ij}, \quad (6)$$

$$0 = -c_{1T}\frac{\varepsilon}{k}\overline{u_i'\theta'} - (1-c_{2T})\,\overline{u_j'\theta'}\,\partial_j U_i - \overline{u_i'u_j'}\,\partial_j \Theta + (1-c_{3T})\,\beta_i\,\overline{\theta'^2} + c_{4T}\,\overline{u_j'\theta'}\,V_{ij}, \quad (7)$$

$$0 = -2\,\overline{u_j'\theta'}\,\partial_j \Theta - \frac{2}{c_T}\frac{\varepsilon}{k}\,\overline{\theta'^2}, \quad (8)$$

and, if we assume that k and ε are known, we have a linear system of 10 equations with 10 unknowns : $(\overline{u'^2}, \overline{v'^2}, \overline{w'^2}, \overline{u'v'}, \overline{u'w'}, \overline{v'w'}, \overline{u'\theta'}, \overline{v'\theta'}, \overline{w'\theta'}, \overline{\theta'^2})$. For clarity, these 10 equations are written explicitly in Appendix D. We solved this system thanks to the symbolic calculus software Mathematica and we confirmed the expressions obtained by Burchard and Bolding (2001):

$$\overline{u'w'} = -c_\mu \frac{k^2}{\varepsilon} \partial_z U, \tag{9}$$

$$\overline{v'w'} = -c_\mu \frac{k^2}{\varepsilon} \partial_z V, \tag{10}$$

$$\overline{w'\theta'} = -c'_\mu \frac{k^2}{\varepsilon} \partial_z \Theta, \tag{11}$$

which reflect downgradient fluxes with an eddy viscosity $K_m = c_\mu \frac{k^2}{\varepsilon}$ and an eddy diffusivity $K_t = c'_\mu \frac{k^2}{\varepsilon}$. The dimensionless functions c_μ and c'_μ are the so-called "stability functions" and can be expressed in the following forms

$$c_\mu = \frac{n_0 + n_1\alpha_N + n_2\alpha_M}{d_0 + d_1\alpha_N + d_2\alpha_M + d_3\alpha_N\alpha_M + d_4\alpha_N^2 + d_5\alpha_M^2}, \tag{12}$$

$$c'_\mu = \frac{n_{0T} + n_{1T}\alpha_N + n_{2T}\alpha_M}{d_0 + d_1\alpha_N + d_2\alpha_M + d_3\alpha_N\alpha_M + d_4\alpha_N^2 + d_5\alpha_M^2}, \tag{13}$$

with $\alpha_N = \frac{k^2}{\varepsilon^2}N^2$, $\alpha_M = \frac{k^2}{\varepsilon^2}M^2$, $N^2 = -g/\rho_0\,\partial_z\rho$ the (squared) buoyancy frequency, and $M^2 = (\partial_z U)^2 + (\partial_z V)^2$ the (squared) shear frequency. Coefficients n_i, n_{iT} and d_i depend on the coefficients c_i and c_{iT}. Their full expressions are given in Appendix E. Taking the values of the c_i and c_{iT} given in Table 1, the stability functions are approximately as follows

$$c_\mu = \frac{0.1067 + 0.01732\alpha_N - 0.0001205\alpha_M}{1 + 0.2398\alpha_N + 0.02872\alpha_M + 0.005154\alpha_N\alpha_M + 0.006930\alpha_N^2 - 0.00003372\alpha_M^2}, \tag{14}$$

$$c'_\mu = \frac{0.1120 + 0.003766\alpha_N + 0.0008871\alpha_M}{1 + 0.2398\alpha_N + 0.02872\alpha_M + 0.005154\alpha_N\alpha_M + 0.006930\alpha_N^2 - 0.00003372\alpha_M^2}. \tag{15}$$

To compute the fluxes in Eqs. (9) - (11), we still need to know k and ε. In a GLS model, we solve two prognostic equations, one for k and one for another variable that can be linked to ε. The choice of this second equation is the main difference between the different GLS models ($k{-}\varepsilon$: Hanjalić and Launder (1972); Rodi (1987), $k{-}kl$: Mellor and Yamada (1982), $k-\omega$: Wilcox (1988), $k-\tau$: Zeierman and Wolfshtein (1986); Thangam et al. (1992)). In this paper, we focus on the $k{-}\varepsilon$ model which solves directly the equation for ε. The TKE equation and the ε equation are as follows

$$D_t k = P + G - \varepsilon + \mathscr{D}_k, \tag{16}$$

$$D_t \varepsilon = \frac{\varepsilon}{k}(c_{\varepsilon 1} P + c_{\varepsilon 3} G - c_{\varepsilon 2}\varepsilon) + \mathscr{D}_\varepsilon, \tag{17}$$

with

- $D_t(\cdot) = [\partial_t + U\partial_x + V\partial_y](\cdot)$: Total derivative
- $\mathscr{D}_k = \partial_z(\frac{K_m}{\sigma_k}\partial_z k)$ and $\mathscr{D}_\varepsilon = \partial_z(\frac{K_m}{\sigma_\varepsilon}\partial_z \varepsilon)$: Diffusion terms
- σ_k and σ_ε: Schmidt numbers for TKE and dissipation
- $P \equiv (-\overline{u'w'}\,\partial_z U - \overline{v'w'}\,\partial_z V) = c_\mu \alpha_M \varepsilon$: Production of TKE by the shear
- $G \equiv \beta_3\,\overline{w'\theta'} = -c'_\mu \alpha_N \varepsilon + c'^{*}_\mu \alpha_T \varepsilon$: Production/destruction of TKE by the buoyancy
- $c_{\varepsilon 1}$, $c_{\varepsilon 2}$ and $c_{\varepsilon 3}$: Empirical coefficients

The TKE equation (16) was obtained by taking the trace of the Reynolds stress equations (3). With the boundary layer approximation which neglects the horizontal gradient in comparison to the vertical ones, taking this trace gives $D_t k + \frac{1}{2}(\partial_z \overline{w' u'_i u'_i}) = P + G - \varepsilon$. We then consider downgradient formulations for the third-order moments $\overline{w' u'_i u'_i}$ and finally results in Equation (16). We want to highlight that the diffusion term thus comes from the divergence of the third-order moments.

An exact equation for ε can be derived but, in practice, this equation needs drastic assumptions to be closed. We used in Equation (17) the classic assumptions of scaling the sources and sinks of ε with the ones of the TKE through empirical coefficients $c_{\varepsilon 1}$, $c_{\varepsilon 2}$ and $c_{\varepsilon 3}$ (see Burchard & Bolding, 2001).

Values $\sigma_k = 1$, $c_{\varepsilon 1} = 1.44$ and $c_{\varepsilon 2} = 1.92$ are frequently used in the literature (Rodi, 1987). Value $\sigma_\varepsilon = 1.20$ is found according to (14) of Umlauf and Burchard (2003). Finally, for $c_{\varepsilon 3}$, it is often considered two different values in order to keep $c_{\varepsilon 3}G$ always as a source term of ε (Rodi, 1987; Burchard & Bolding, 2001; Umlauf & Burchard, 2003; Warner et al., 2005; Reffray et al., 2015). A positive value $c^+_{\varepsilon 3}$ is used when G is positive (stable stratification) and a negative value $c^-_{\varepsilon 3}$ is used when G is negative (unstable stratification). However, Umlauf and Burchard (2005) argued that this is not necessary and that better results (particularly for the heat flux profile) are obtained with considering always a negative value. We do this choice and the value $c_{\varepsilon 3} = -0.65$ is obtained according to (26) of Umlauf et al. (2003) (by considering a steady state Richardson number equal to 0.25).

2.3 Procedure for the $k\varepsilon t$ Parameterization

The new procedure differs from the GLS one by considering that the temperature variance $\overline{\theta'^2}$ is not at equilibrium anymore. Relaxing this assumption takes us from the level $2\frac{1}{2}$ to the level 3 in the hierarchy of Mellor and Yamada (1982). Beyond this mathematical justification, the idea of keeping the non-equilibrium $\overline{\theta'^2}$ equation originated from the fact that the $\overline{\theta'^2}$ dependence appears only in the $\overline{w'\theta'}$ equation (see Eqs. (3) and (4)). Thus, a physical change in the shape of the $\overline{\theta'^2}$ profile will directly impact $\overline{w'\theta'}$. Because we now have an equation for the temperature variance, we are left with (6) and (7) that form a system of 9 equations with 9 unknowns: ($\overline{u'^2}$, $\overline{v'^2}$, $\overline{w'^2}$, $\overline{u'v'}$, $\overline{u'w'}$, $\overline{v'w'}$, $\overline{u'\theta'}$, $\overline{v'\theta'}$, $\overline{w'\theta'}$). For clarity, these 9 equations are written explicitly in Appendix F. We solve this system thanks to Mathematica and we obtain the following expressions:

$$\overline{u'w'} = -c_\mu \frac{k^2}{\varepsilon} \partial_z U, \tag{18}$$

$$\overline{v'w'} = -c_\mu \frac{k^2}{\varepsilon} \partial_z V, \tag{19}$$

$$\overline{w'\theta'} = -c'_\mu \frac{k^2}{\varepsilon} \partial_z \Theta + c'^*_\mu \frac{k}{\varepsilon} \beta_3 \, \overline{\theta'^2}. \tag{20}$$

The momentum fluxes are still downgradient with an eddy viscosity $K_m = c_\mu \frac{k^2}{\varepsilon}$ whereas the temperature flux now has a "non-gradient" contribution $\Gamma_{k\varepsilon t} = c'^*_\mu \frac{k}{\varepsilon} \beta_3 \, \overline{\theta'^2}$

related to the temperature variance in addition to the downgradient part with eddy diffusivity $K_t = c'_\mu \frac{k^2}{\varepsilon}$. The stability functions c_μ, c'_μ and c'^*_μ can be expressed in the following forms

$$c_\mu = \frac{n_0 + n_1\alpha_N + n_2\alpha_M + n_3\alpha_T}{d_0 + d_1\alpha_N + d_2\alpha_M + d_3\alpha_N\alpha_M + d_4\alpha_N^2 + d_5\alpha_M^2}, \tag{21}$$

$$c'_\mu = \frac{n_{0T} + n_{1T}\alpha_N + n_{2T}\alpha_M}{d_0 + d_1\alpha_N + d_2\alpha_M + d_3\alpha_N\alpha_M + d_4\alpha_N^2 + d_5\alpha_M^2}, \tag{22}$$

$$c'^*_\mu = \frac{n^*_{0T} + n^*_{1T}\alpha_N + n^*_{2T}\alpha_M}{d_0 + d_1\alpha_N + d_2\alpha_M + d_3\alpha_N\alpha_M + d_4\alpha_N^2 + d_5\alpha_M^2}, \tag{23}$$

with $\alpha_N = \frac{k^2}{\varepsilon^2}N^2$, $\alpha_M = \frac{k^2}{\varepsilon^2}M^2$, and $\alpha_T = \frac{k}{\varepsilon^2}\beta_3^2\overline{\theta'^2}$. Coefficients n_i, n_{iT} and d_i depends on the coefficients c_i and c_{iT}. Their full expressions are given in Appendix G. Taking the values of the c_i and c_{iT} given in Table 1, the stability functions are approximately as follows

$$c_\mu = \frac{0.1067 + 0.0001072\alpha_N - 0.0001205\alpha_M + 0.004673\alpha_T}{1 + 0.07843\alpha_N + 0.02872\alpha_M + 0.0003389\alpha_N\alpha_M + 0.001506\alpha_N^2 - 0.00003372\alpha_M^2}, \tag{24}$$

$$c'_\mu = \frac{0.1120 + 0.003766\alpha_N + 0.0008871\alpha_M}{1 + 0.07843\alpha_N + 0.02872\alpha_M + 0.0003389\alpha_N\alpha_M + 0.001506\alpha_N^2 - 0.00003372\alpha_M^2}, \tag{25}$$

$$c'^*_\mu = \frac{0.1120 + 0.003766\alpha_N + 0.003344\alpha_M}{1 + 0.07843\alpha_N + 0.02872\alpha_M + 0.0003389\alpha_N\alpha_M + 0.001506\alpha_N^2 - 0.00003372\alpha_M^2}. \tag{26}$$

As in the GLS procedure, the TKE and ε equations (equations (16) and (17)) are solved prognostically. The only difference in these equations is about the $c_{\varepsilon 3}$ coefficient which is now calculated to be $c_{\varepsilon 3} = -1.83$ according to (26) of Umlauf et al. (2003) (by considering a steady state Richardson number equal to 0.25).

Beyond this minor change, one key difference is that the temperature variance is now also solved prognostically through:

$$D_t\overline{\theta'^2} = -2\,\overline{w'\theta'}\,\partial_z\Theta - \frac{2}{c_T}\frac{\varepsilon}{k}\,\overline{\theta'^2} + \mathscr{D}_{\overline{\theta'^2}}, \tag{27}$$

with $\mathscr{D}_{\overline{\theta'^2}} = \partial_z(\frac{K_m}{\sigma_{\overline{\theta'^2}}}\partial_z\overline{\theta'^2})$ the diffusion and $\sigma_{\overline{\theta'^2}}$ the Schmidt number for the temperature variance. As for the TKE equation, the diffusion term $\mathscr{D}_{\overline{\theta'^2}}$ results from the closure of the third-order moment $\overline{w'\theta'\theta'}$ by a downgradient formulation. We did not find any estimations of the Schmidt number $\sigma_{\overline{\theta'^2}}$ in the literature and, as a first guess, we took $\sigma_{\overline{\theta'^2}} = \sigma_k = 1$, meaning that the temperature variance is diffused with the same intensity as TKE.

We add several comments about the non-gradient term $\Gamma_{k\varepsilon t} = c'^{*}_{\mu}\frac{k}{\varepsilon}\beta_3\,\overline{\theta'^2}$ we obtained for the temperature flux. Firstly, we recall that, by writing $\overline{w'\theta'} = -K_m(\partial_z\Theta - \gamma_{k\varepsilon t})$, we highlight that $\gamma_{k\varepsilon t} = \frac{c'^{*}_{\mu}}{c'_{\mu}}\frac{1}{k}\beta_3\overline{\theta'^2}$ gives the stable stratification towards which $\partial_z\Theta$ tends to relax. Secondly, the form of $\Gamma_{k\varepsilon t}$ can be compared to the one found by Deardorff (1972). By reasoning with the $\overline{w'\theta'}$ equation, Deardorff (1972) found a non-gradient term $\Gamma_{\mathrm{Deardoff}} \propto l/k^{1/2}\,\overline{\theta'^2}$ with l a mixing length introduced for the parameterization of the pressure-temperature correlation. If we consider the classic scaling $l \propto k^{3/2}/\varepsilon$ (see for example Rodi, 1987; Umlauf & Burchard, 2003, 2005), we obtain $\Gamma_{\mathrm{Deardoff}} \propto k/\varepsilon\,\overline{\theta'^2}$. The non-gradient expressions of $\Gamma_{k\varepsilon t}$ and $\Gamma_{\mathrm{Deardoff}}$ thus both exhibit the same dependence on the turbulence time scale k/ε and on the temperature variance $\overline{\theta'^2}$. This is fundamentally different from $\Gamma_{\mathrm{KPP}} \propto G\,\overline{w'\theta'}|_{z=0}$ which is written explicitly as a redistribution of the surface temperature flux $\overline{w'\theta'}|_{z=0}$ according to an empirical shape function G that is a third-order polynomial of the dimensionless vertical coordinate z/h with h the mixed layer depth.

Finally, we point out that, just as we retained the non-equilibrium equation of $\overline{\theta'^2}$ to obtain a non-gradient term for $\overline{w'\theta'}$, it would be tempting to retain the non-equilibrium equation of $\overline{w'^2}$ to obtain non-gradient terms for the velocity fluxes $\overline{u'w'}$ and $\overline{v'w'}$. We solved this problem and, astonishingly, the velocity fluxes $\overline{u'w'}$ and $\overline{v'w'}$ in this context are still downgradient. Results of this $k{-}\varepsilon{-}\overline{\theta'^2}{-}\overline{w'^2}$ model are detailed in Appendix H.

2.4 1D Models Simulations

We implemented the $k\varepsilon t$ parameterization, with the formalism described in section 2.3, in the 1D code presented in Fearon et al. (2020). This code is a standalone 1D vertical version of the Coastal and Regional Ocean COmmunity model (CROCO, `https://www.croco-ocean.org/`) and allows to run simulations with KPP, TKE, and several GLS schemes (note that we also re-implemented the $k-\varepsilon$ model with the formalism presented in section 2.2, that is equivalent to using the Canuto et al. (2001) stability functions).

The temperature variance equation (27) is discretized using a backward Euler scheme in time. To preserve the positivity of $\overline{\theta'^2}$, the Patankar trick is used (Patankar, 1980; Burchard, 2002; Lemarié et al., 2021). Boundary conditions for the temperature variance are zero at the bottom of the domain (Dirichlet condition), while at the surface a homogeneous Neumann condition is used (no flux of temperature variance).

For every test case, we performed the simulations using the $k-\varepsilon$ model, the $k\varepsilon t$ model, and the KPP model. The changes induced by the $k\varepsilon t$ model, particularly the influence of the non-gradient term, will be analyzed by comparing with the $k-\varepsilon$ model. Concerning the KPP scheme, the simulations were done with and without its non-gradient term. The goal is to compare this term and its effect to the non-gradient term obtained in the $k\varepsilon t$ parameterization. The version of KPP used here is the original one described in Large et al. (1994).

2.5 LES Simulation

In order to validate the $k\varepsilon t$ model, we performed some LES simulations. Practically, we use the Basilisk code (`http://basilisk.fr`, Popinet, 2020) to solve the three-dimensional Boussinesq equations in a small oceanic patch near the air-sea interface. We intend to explicitly compute the turbulent fluxes and the mean vertical profiles of temperature for buoyancy-driven convection and wind-driven convection. We can then compare these fluxes with the parameterization. The size of the domain is $L_x = L_y = 1200$ m (periodic in the horizontal direction), and $L_z = 600$ m. The grid resolution is isotropic (2.3 m) with $512 \times 512 \times 256$ cells. All variables are discretized at the cell center and are advected using the Bell-Collela-Glaz method. There is no explicit viscosity and no explicit diffusivity: both these terms are handled implicitly by the advection scheme. The surface forcing (wind and heat flux) is applied at the upper grid cell with a relaxation term. The bottom boundary condition is free slip for the velocity and inhomogeneous Neumann for the temperature (set to the initial stratification). The model is initialized with zero velocity and prescribed stratification for temperature (see next paragraph) to which we add a small random perturbation of magnitude 10^{-3} °C. We use an adaptive time step adjusted with a CFL condition of 0.6. Averages are computed in a post-processing step: the overbar is interpreted here as a horizontal average and primes are deviations from this horizontal average.

2.6 The Two Test Cases: Cooling-Dominant and Wind-Dominant

Two simulation setups were defined in order to capture the different convective regimes highlighted in Legay et al. (2024). The first configuration is a cooling-dominant simulation forced by a surface net heat flux of $Q_0 = -320\,\mathrm{W\,m^{-2}}$ and a wind stress of $\tau_x = 0.64\,\mathrm{N\,m^{-2}}$; it is initialized with a surface temperature of 293 K and a constant stratification of 3.9 K/1000 m. The second one is a wind-dominant simulation forced by a surface net heat flux of $Q_0 = -8\,\mathrm{W\,m^{-2}}$ and a wind stress of $\tau_x = 0.41\,\mathrm{N\,m^{-2}}$; it is initialized with a surface temperature of 293 K and a constant stratification of 1.2 K/1000 m. Rotation is included with a Coriolis frequency of $f = 10^{-4}\ \mathrm{s^{-1}}$, this corresponds to a latitude of 44 °N. The two cases are simulated with 10 days of constant forcing conditions. For the 1D simulations, the domain is discretized on the same vertical grid as the 3D model (uniformly spaced vertical grid of 256 points), and the time step is 360 s.

2.7 Nondimensionalization

In order to compare the shape of the different profiles, variables are made dimensionless. For the depth, we found that using the depth of the maximum temperature variance $z[\max(\overline{\theta'^2})]$ as a proxy of the mixed layer depth h_m is the best choice for two main reasons. Firstly, the temperature variance is well converged with a maximum that is prominent, easy to identify, and located at the same depth as the classic definition of the minimum of $\overline{w'\theta'}$ (see Figure 2). Second, this definition holds for wind-dominant simulations whereas in this case, the temperature flux profile can be far from the idealized version presented in Figure 1. We mention that while this method works in most cases, there are some conditions where $\overline{\theta'^2}$ is maximum at the surface. In this case, we simply considered the second maximum strictly below the surface. We tested other definitions of h_m such as the minimum of the temperature flux $\overline{w'\theta'}$ or other definitions of the mixed layer depth h_m, but they appeared to be less robust definitions (subject to noisy variations).

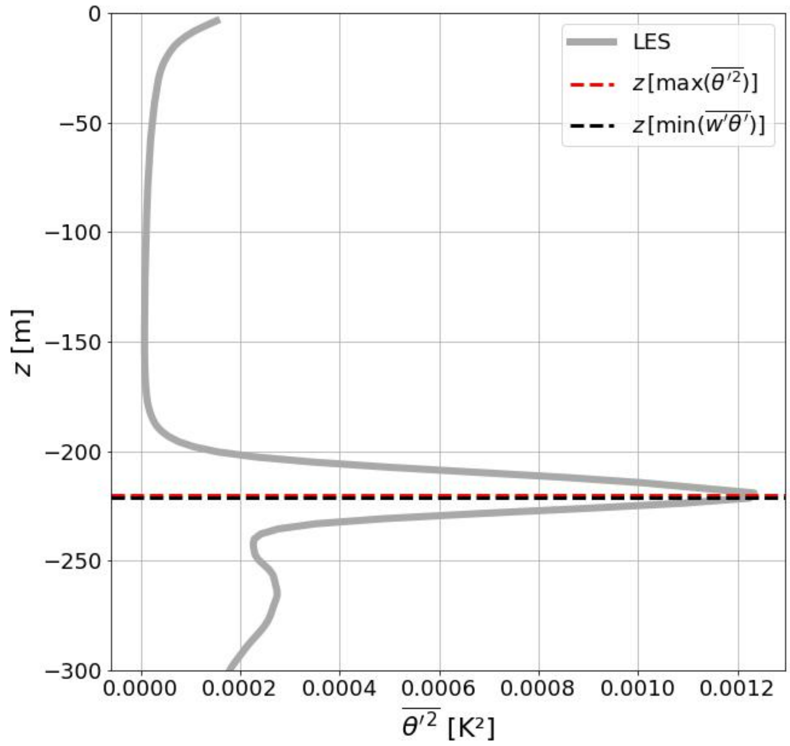

Figure 2. Temperature variance profile of the LES simulation at the end of the simulation for the cooling-dominant case. Dashed lines indicate two different proxies of the mixed layer depth: the maximum of the temperature variance and the minimum of the temperature flux. In this case, these two proxies are localized at the same depth.

The other nondimensionalizations consist in normalizing the temperature flux by its surface value $\overline{w'\theta'}|_0 = Q_0/(\rho_0 c_p)$, with $\rho_0 = 1027\,\mathrm{kg\,m^{-3}}$ the reference density and $c_p = 4000\,\mathrm{J\,kg^{-1}\,K^{-1}}$ the specific heat capacity; and normalizing the temperature in $(\Theta - \Theta_{max}) / \Theta^*$ with Θ_{max} the maximum of the temperature over the vertical and $\Theta^* = \overline{w'\theta'}|_0/w^*$ a scaling of the temperature, with $w^* = (-B_0 h_m)^{1/3}$ a scaling of the velocity of the convective thermals (Willis & Deardorff, 1974; Marshall & Schott, 1999), $B_0 = g\alpha Q_0/(\rho_0 c_p)$ the surface buoyancy flux, g the gravitational acceleration and α the thermal expansion coefficient taken equal to $2.6 \times 10^{-4}\,\mathrm{K^{-1}}$.

3 Results and Discussion

3.1 Cooling-Dominant Case

Figure 3 presents the dimensionless temperature flux profile of the $k{-}\varepsilon$ and the $k\varepsilon t$ simulations at the end of the 10 days of simulations for the cooling-dominant case. The $k\varepsilon t$ flux is further decomposed into its downgradient $(-K_t \partial_z \Theta)$ and its non-gradient $(\Gamma_{k\varepsilon t})$ components (see Eq. 20). It is remarkable that even if the expression of the total flux changed drastically between the two parameterizations, the $k\varepsilon t$ profile is very similar to the $k{-}\varepsilon$ one that exhibits the classic pattern expected for a cooling-dominant simulation: a linear decrease from the surface to the bottom of the mixed layer where it reaches a minimum which is approximately -0.2 times the surface flux. The non-gradient flux is positive (by definition), and it is zero at the surface and at the bottom of the mixed layer; hence, it does not add or remove any heat but rather redistributes heat among the mixed layer. This term is responsible for warming the upper part of the mixed layer and cooling the lower part of the mixed layer (the temperature equation is of the form $D_t\Theta = ... {-}\partial_z \overline{w'\theta'}$ and it is then the sign of $-\partial_z \Gamma_{k\varepsilon t}$ that is important to distinguish between cooling and warming). This is qualitatively the effect we expect from a coherent thermal: thermals grow by entraining cold water near the surface, resulting in a warming of the upper part of the mixed layer, and then detrain in the environment which results in a cooling of the bottom part of the mixed layer.

Figure 4 presents the dimensionless temperature profile of the $k{-}\varepsilon$, the $k\varepsilon t$, the KPP, and the LES simulations at the end of the 10 days of simulations. Dashed lines highlight the location h_t, the depth at which $\partial_z \Theta = 0$ for each case. The overall comparison with the LES is better with $k\varepsilon t$ scheme than with $k{-}\varepsilon$: while the $k{-}\varepsilon$ model

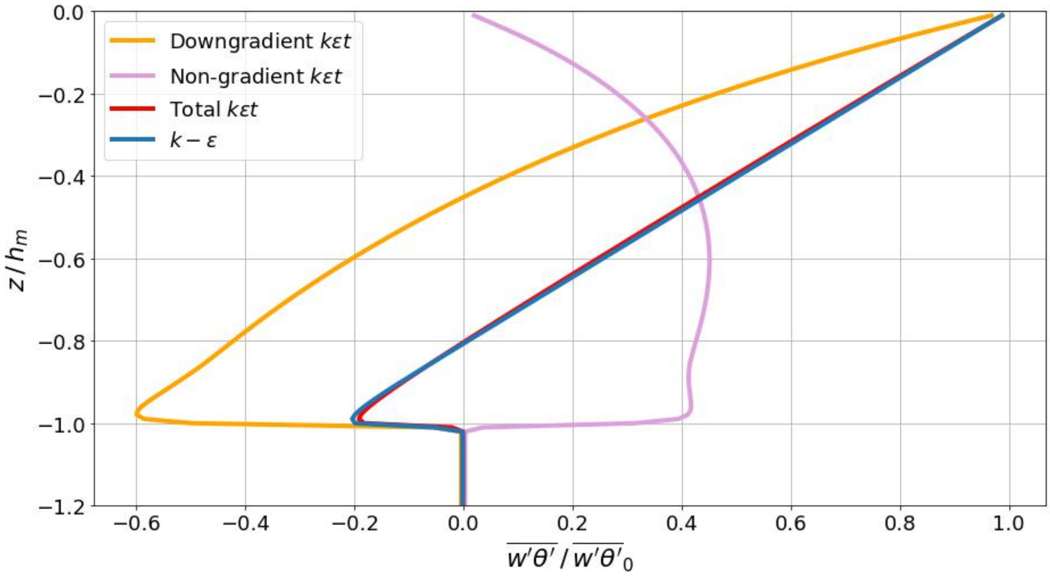

Figure 3. Dimensionless temperature flux profiles of the $k - \varepsilon$ and the $k\varepsilon t$ simulations at the end of the 10 days of simulations for the cooling-dominant case. The $k\varepsilon t$ flux is further decomposed into its downgradient ($-K_t \partial_z \Theta$) and its non-gradient ($\mathit{\Gamma}_{k\varepsilon t}$) components (see Eq. 20).

predicts $h_t = h_f = -0.8h_m$ (by definition of a pure downgradient flux, see Fig. 3), this co-location constraint is relaxed in the $k\varepsilon t$ simulation, for which $h_t = -0.44h_m$, which is closer to the LES ($h_t = -0.41h_m$). The KPP scheme predicts $h_t = -0.2h_m$, whereas the KPP simulation without the non-local term $\mathit{\Gamma}_{\mathrm{KPP}}$ gives $h_t = -0.93h_m$. Therefore, $\mathit{\Gamma}_{\mathrm{KPP}}$ has the same expected effect to raise h_t up as the non-gradient term of $k\varepsilon t$ but, none of the two KPP simulations (with or without $\mathit{\Gamma}_{\mathrm{KPP}}$) give a satisfactory h_t in comparison to the LES.

Figure 5 shows the temporal evolution of h_t/h_m for the 10 days of the simulation. The evolution of this quantity in the LES, although a bit noisy, shows h_t/h_m between -0.4 and -0.6 at the end of the simulation. The $k - \varepsilon$ values decrease then stabilize around -0.8. The KPP simulation quickly stabilizes near -0.2 whereas KPP without the non-local term gives a continuous decrease of h_t/h_m with values reaching -0.93 at the end of the 10 days. The $k\varepsilon t$ curve, among all schemes, exhibits the closest values to those of the LES. However, it results in a continuous increase during 10 days. This behavior can be modified by considering a different value of $\sigma_{\overline{\theta'^2}}$. Thus, another simulation of the $k\varepsilon t$ model with $\sigma_{\overline{\theta'^2}} = 10$ (a temperature variance that diffuses 10 times less

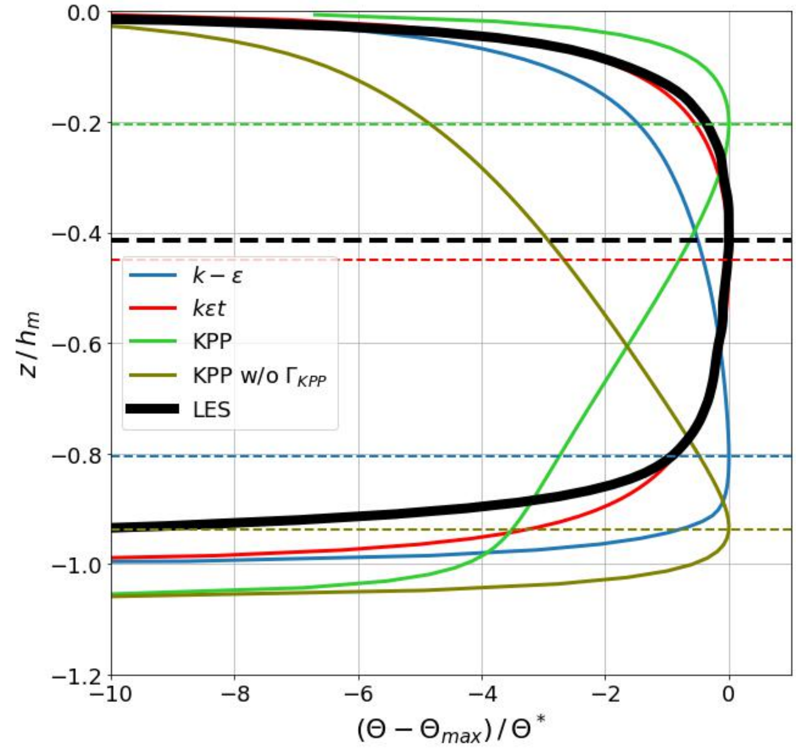

Figure 4. Dimensionless temperature profiles of the $k - \varepsilon$, the $k\varepsilon t$, the KPP, and the LES simulations at the end of the 10 days of simulations for the cooling-dominant case. The KPP model was run with and without its non-gradient term $\mathit{\Gamma}_{\mathrm{KPP}}$. Dashed lines highlight the location h_t of the zero of the gradient $\partial_z \Theta$.

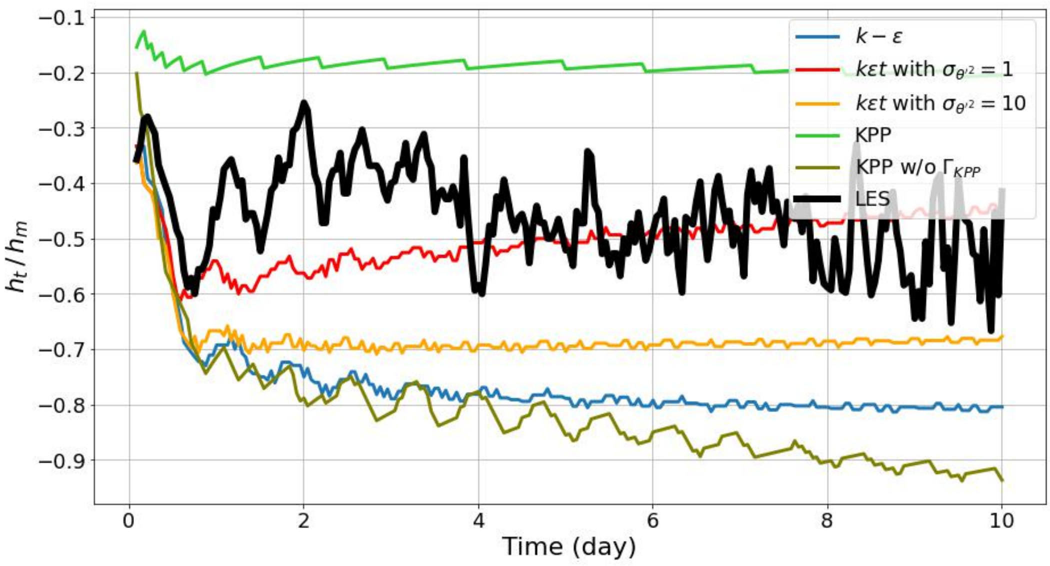

Figure 5. Temporal evolution of h_t/h_m in the cooling-dominant case for the $k - \varepsilon$, the $k\varepsilon t$, the KPP and the LES simulations. The $k\varepsilon t$ simulation was run with two different values of the Schmidt number for the temperature variance: $\sigma_{\overline{\theta'^2}} = 1$ and $\sigma_{\overline{\theta'^2}} = 10$. The KPP model was run with and without its non-gradient term $\mathit{\Gamma}_{\text{KPP}}$.

than the velocities) gives h_t/h_m that stabilizes around -0.7. This preliminary test highlights the need to adjust all parameters of this closure with advanced Bayesian methods such as the ones used in Souza et al. (2020) and Wagner et al. (2023). This calibration procedure would require an ensemble of LES simulations in order to not overfit the parameters to the two LES used here and this task is beyond the scope of this study.

Figure 6 shows a comparison between the non-gradient term of $k\varepsilon t$ (run with two different values $\sigma_{\overline{\theta'^2}} = 1$ and $\sigma_{\overline{\theta'^2}} = 10$) and the non-local term of KPP at the end of the 10 days of simulation. These profiles share the property of vanishing at the surface and at the bottom of the mixed layer, they therefore both act as a redistribution of heat in the mixed layer. The KPP term appears to have a single-mode shape. In fact, $\mathit{\Gamma}_{\text{KPP}}$ can be written as $\mathit{\Gamma}_{\text{KPP}}(z) = C_s G(z) \overline{w'\theta'}|_0$ with C_s a constant (see for example Equation (20) of Van Roekel et al., 2018). The vertical dependence is entirely contained in G which is a third-order polynomial. Hence, $\mathit{\Gamma}_{\text{KPP}}$ can only have a single positive mode. Instead, $\mathit{\Gamma}_{k\varepsilon t}$ presents a bi-modal shape for both $\sigma_{\overline{\theta'^2}} = 1$ and $\sigma_{\overline{\theta'^2}} = 10$. For $\sigma_{\overline{\theta'^2}} = 1$, the two modes are close one to the other but, for $\sigma_{\overline{\theta'^2}} = 10$, the non-gradient term presents two clear distinct modes. In the latter case, the simple qualitative way of see-

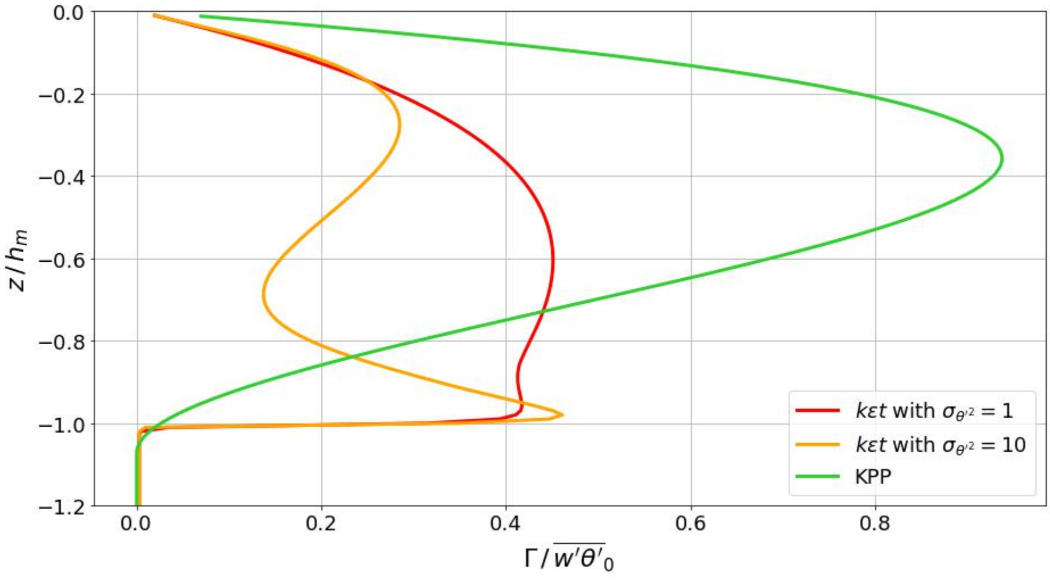

Figure 6. Dimensionless profiles of the non-gradient term of $k\varepsilon t$ (run with two different values $\sigma_{\overline{\theta'^2}} = 1$ and $\sigma_{\overline{\theta'^2}} = 10$) and KPP at the end of the 10 days of simulation for the cooling-dominant case.

ing the non-gradient term as the effect of a thermal is no longer relevant. This point is supported by Zhou et al. (2018) who proved that the often-implied association of the gradient and non-gradient term terms to the local and non-local eddies is partially wrong. Analyses of the contribution of the different factors of $\Gamma_{k\varepsilon t} = c'^*_\mu \frac{k}{\varepsilon} \beta_3 \overline{\theta'^2}$ (not shown) indicated that the mode close to the mixed layer bottom is mainly due to a maximum of $\overline{\theta'^2}$ whereas the mode closest to the surface is a result of a complex interaction of all the terms in the expression of the non-gradient term. Knowing that $\Gamma_{k\varepsilon t}$ presents a bi-modal shape could be of interest for adapting the KPP non-gradient term. For example, it would be possible to consider Γ_{KPP} as a sum of two polynomials rather than one for trying to catch this bi-modal feature.

3.2 Wind-Dominant Case

Figure 7 presents the dimensionless temperature flux profile of the $k-\varepsilon$ and the $k\varepsilon t$ simulations at the end of the 10 days of simulations for the wind-dominant case. For the shape of the flux, we get similar conclusions as in the cooling-dominant case: we ob-

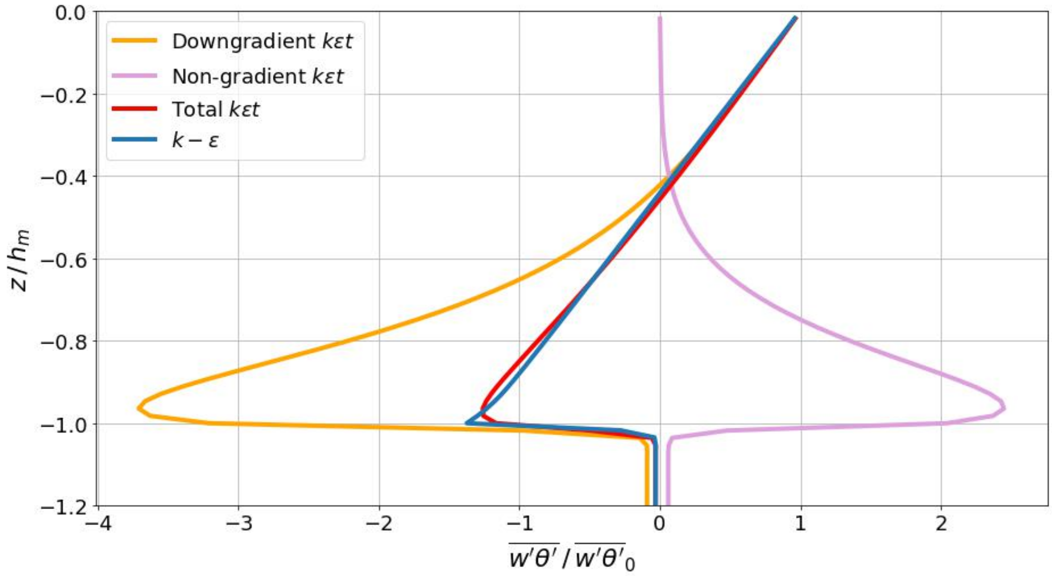

Figure 7. Dimensionless temperature flux profiles of the $k - \varepsilon$ and the $k\varepsilon t$ simulations at the end of the 10 days of simulations for the wind-dominant case. The $k\varepsilon t$ flux is further decomposed into its downgradient and its non-gradient components.

tain a remarkable agreement between the $k\varepsilon t$ total profile and the $k{-}\varepsilon$ profile even if the expression of the total flux changed between the two parameterizations.

Figure 8 presents the dimensionless temperature profile of the $k{-}\varepsilon$, the $k\varepsilon t$, the KPP, and the LES simulations at the end of the 10 days of simulations. Here again, dashed lines highlight the location of h_t in all cases. The effect of the non-gradient term of $k\varepsilon t$ of raising h_t is negligible here, and this is fine since $k{-}\varepsilon$ correctly predicts the LES profile. Instead, the difference between KPP and KPP without $\mathit{\Gamma}_{\mathrm{KPP}}$ is substantial. KPP without $\mathit{\Gamma}_{\mathrm{KPP}}$ gives a good profile while the full KPP results in a profile that presents a high value of h_t. The fact that $k{-}\varepsilon$ and KPP without $\mathit{\Gamma}_{\mathrm{KPP}}$ are already satisfactory suggests that non-gradient effects are less important in this wind-dominant case than in the cooling-dominant case. If we adopt the disputed view of associating non-gradient effects to non-local eddies, this suggests that the deepening is here dominated by local eddies driven by shear while the deepening in the cooling-dominant case is driven by non-local thermals.

Figure 9 shows the temporal evolution of h_t/h_m for all models. The LES evolution consists of a continuous decrease until near -0.45 at the end of the simulation (with no

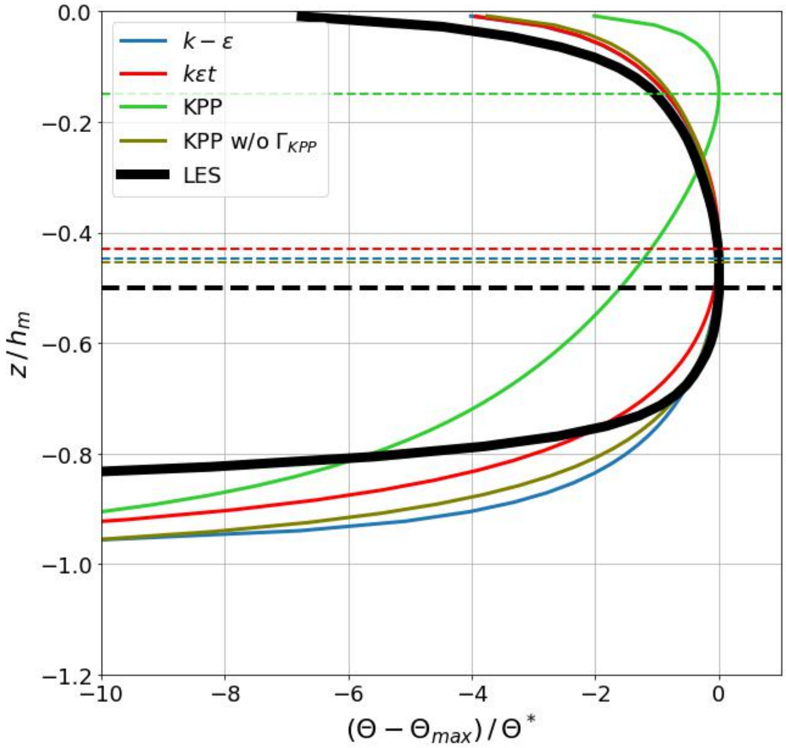

Figure 8. Dimensionless temperature profiles of the $k - \varepsilon$, the $k\varepsilon t$, the KPP, and the LES simulations at the end of the 10 days of simulations for the wind-dominant case. The KPP model was run with and without its non-gradient term Γ_{KPP}. Dashed lines highlight the location h_t of the zero of the gradient $\partial_z \Theta$.

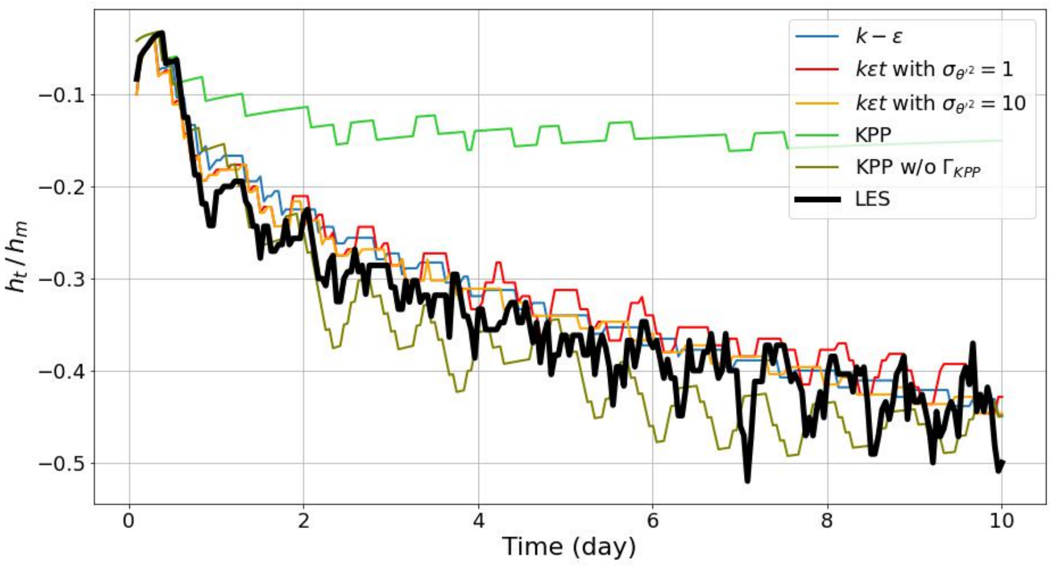

Figure 9. Temporal evolution of h_t/h_m in the wind-dominant case for the $k - \varepsilon$, the $k\varepsilon t$, the KPP and the LES simulations. The $k\varepsilon t$ simulation was run with two different values of the Schmidt number for the temperature variance: $\sigma_{\overline{\theta'^2}} = 1$ and $\sigma_{\overline{\theta'^2}} = 10$. The KPP model was run with and without its non-gradient term $\mathit{\Gamma}_{\mathrm{KPP}}$.

clear convergence). This evolution is reproduced by $k{-}\varepsilon$, $k\varepsilon t$ and KPP without $\mathit{\Gamma}_{\mathrm{KPP}}$. Instead, the comparison of the full KPP with the LES is not in favor of KPP, since h_t/h_m stabilizes around -0.15 in this case. The LES evolution presents inertial oscillations of h_t/h_m at the inertial period $T_f = 2\pi/f = 17\,\mathrm{h}\,30$ min. This is captured by $k\varepsilon t$ and KPP without $\mathit{\Gamma}_{\mathrm{KPP}}$ but not by the full KPP and $k{-}\varepsilon$. Changing the value of $\sigma_{\overline{\theta'^2}}$ gives here almost no effect, contrary to what was highlighted in the cooling-dominant case. This supports the fact that non-gradient effects are probably negligible in wind-dominant regimes.

Figure 10 shows a comparison between the non-gradient term of $k\varepsilon t$ (run with two different values $\sigma_{\overline{\theta'^2}} = 1$ and $\sigma_{\overline{\theta'^2}} = 10$) and KPP at the end of the 10 days of simulation. The KPP shape is very similar to the one of the cooling-dominant case (Figure 6). On the opposite, the $k\varepsilon t$ profiles changed drastically in comparison to the cooling-dominant case. Indeed, these profiles present now a single-mode shape. The mode closest to the surface disappeared because $\overline{\theta'^2}$ is here equal to zero at the surface. This behavior can again inspire the construction of the KPP non-gradient term. If, as suggested, it is constructed by the sum of two polynomials, the polynomial with its maximum close

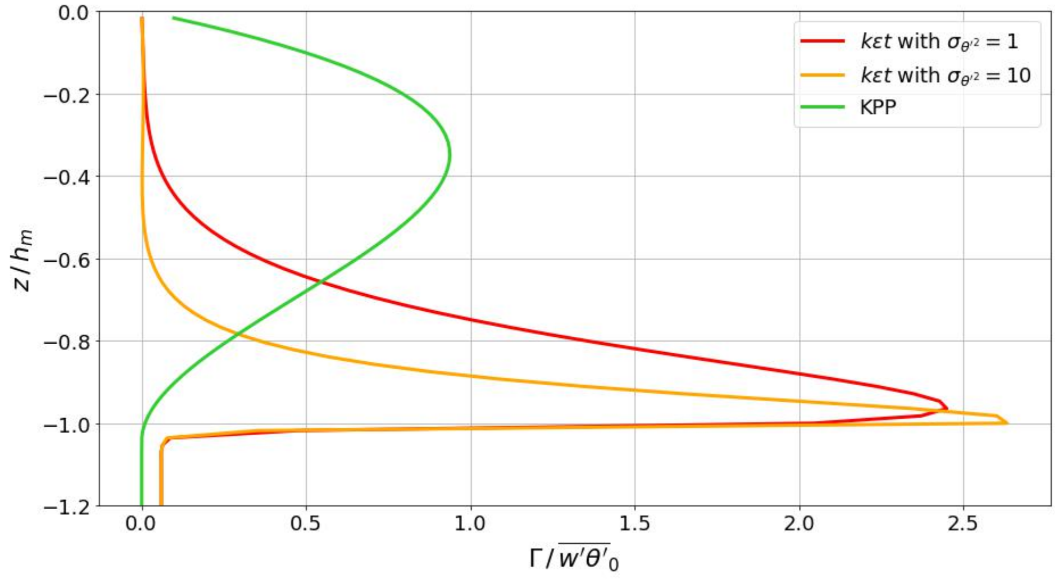

Figure 10. Dimensionless profiles of the non-gradient term of $k\varepsilon t$ (run with two different values $\sigma_{\overline{\theta'^2}} = 1$ and $\sigma_{\overline{\theta'^2}} = 10$) and KPP at the end of the 10 days of simulation for the wind-dominant case.

to the surface must vanish in the wind-dominant case whereas the second polynomial with its maximum near the mixed layer bottom must be present both in the cooling-dominant and wind-dominant conditions.

4 Conclusion

The primary motivation behind this research was the need to improve the representation of oceanic convection processes in ocean models. Indeed, most parameterizations adopt a downgradient approach for which the mixing of tracers and momentum primarily occurs in the direction of their gradients, such as from regions of high temperature to low temperature. However, in convective situations, this simplistic assumption falls short, and turbulent fluxes cannot be solely explained or formulated as a downgradient process (Zhou et al., 2018). While this property was originally recognized for atmospheric convection (Hourdin et al., 2002), oceanographers were also aware of this aspect of convection when they introduced a non-local term in KPP (Large et al., 1994). In this context, the non-local term represented the influence and transport of tracers across different spatial locations within a convective system, even when these locations are not immediately adjacent. In simpler terms, the non-local term accounted for the long-range mixing of properties that occur in convective events. While there is an ongoing effort to capture this property with a mass flux parameterization (Giordani et al., 2020), our approach has been to seek an analytical formulation of this non-local (or non-gradient) term within the GLS (Generic Length Scale) models and we can retrospectively comment on that choice based on the results obtained in this article. One key argument for this approach is the desire for consistency and integration within the modeling framework. By deriving the non-gradient term analytically within the GLS framework, we aim to ensure that all components of the parameterization align seamlessly. This approach avoids potential mismatches or inconsistencies that may arise when adding external components to existing parameterizations. Another crucial argument is the need to deepen our physical understanding of oceanic convection processes. By analytically deriving the non-gradient term within the GLS framework, we gain insights into the underlying physics and dynamics governing this term. This understanding can lead to more robust and physically grounded parameterization, improving our ability to capture convective processes accurately. Last, our approach offers flexibility for optimization and adaptation. As the GLS framework provides a versatile platform for parameterization, we can adapt and refine the derived non-gradient term to suit specific oceanic conditions or scenarios. This adaptability is valuable for tailoring the parameterization to different modeling and research needs.

In order to assess the validity of our approach we have compared our scheme with Large eddy simulations (LES). The main metric that we analyzed was the depth of the thermal layer h_t which corresponds to the depth at which $\partial_z \Theta = 0$. We have verified that the effect of the non-gradient term is to raise h_t such that a significant part of the mixed layer is stably stratified (at least in the thermally driven convection), an aspect that was not well reproduced by the $k - \varepsilon$ model. We also noted that since the main effect of this term is to redistribute heat, the addition of the non-gradient term does not have a profound impact on the evolution of the depth of the mixed layer. We have also conducted extensive comparisons with KPP. With these comparisons, we have unveiled common aspects between our newly derived non-gradient term and certain aspects of the KPP non-local term. This comparison suggests that our work has the potential to serve as a source of inspiration for enhancing and fine-tuning the KPP parameterization. Particularly, it could be used to modify the definition of the ad hoc polynomial that shape the diffusivity and the non-local term in KPP.

Extending the derivation to include salinity will allow us to more comprehensively capture the behavior of oceanic convection. We are currently working on this approach: the main challenge is that the non-gradient term for salinity involves coupled equations with temperature, making the analytical derivation significantly more complex. Solving these coupled equations analytically is mathematically challenging and may require additional hypotheses. Another issue is also that the computational demands of implementing a coupled temperature-salinity non-gradient term within ocean models may increase. This can affect model efficiency and require adjustments in computational resources. Despite these difficulties, the extension of the non-gradient term derivation to salinity promises a more comprehensive and accurate representation of oceanic convection. In the near future, our research plans also entail a systematic re-evaluation of all GLS parameters and $k\varepsilon t$ parameters. To achieve this, we will employ an ensemble of LES simulations, with a resolution high enough to capture the energetic eddies in entrainment layers, in conjunction with Bayesian methods (Wagner et al., 2023). Bayesian methods offer a data-driven approach to parameter estimation, allowing us to incorporate real-world observations and LES data into the parameterization process.

Appendix A Open Research

All the codes and the data used for the study are available through the GitHub repository `https://github.com/legaya/James2024-ket/` or the following DOI: `https://doi.org/10.5281/zenodo.10562734`. These archives contain the two Jupyter Notebooks used for performing the 1D simulations and all the analyses, the 1D model described in section 2.4 as Fortran Modules, the Fortran codes needed for generating these modules, the files needed to perform the LES simulations, and the LES results as netCDF files.

Appendix B Coefficients in the Second-Order Moment Equations

Coefficients $c_1, c_2, c_3, c_4, c_5, c_{1T}, c_{2T}, c_{3T}, c_{4T}, c_T$ used in Eqs (3) - (5) are linked to the coefficients introduced by Canuto et al. (2001) through the following formulas:

$$\begin{gathered} c_1 = 1/\lambda, \qquad c_2 = \alpha_1, \qquad c_3 = 1 - \beta_5, \qquad c_4 = 4/3\,\alpha_1 - 4/5, \qquad c_5 = \alpha_1 - \alpha_2, \\ c_{1T} = \lambda_5/2, \qquad c_{2T} = 3/4\,\alpha_3, \qquad c_{3T} = \gamma_1, \qquad c_{4T} = \alpha_3/2, \qquad c_T = 2\,\lambda_8/(1 - \gamma_1). \end{gathered} \tag{B1}$$

Appendix C Expressions of the Main Tensors under the Boundary Layer Approximation

After applying the boundary layer approximation, the tensors $P_{ij}, B_{ij}, S_{ij}, V_{ij}, Z_{ij}$ used in Eqs (3) - (5) simplify to

$$P_{ij} = \begin{pmatrix} -2\,\partial_z U\,\overline{u'w'} & -\partial_z U\,\overline{v'w'} - \partial_z V\,\overline{u'w'} & -\partial_z U\,\overline{w'^2} \\ -\partial_z U\,\overline{v'w'} - \partial_z V\,\overline{u'w'} & -2\,\partial_z V\,\overline{v'w'} & -\partial_z V\,\overline{w'^2} \\ -\partial_z U\,\overline{w'^2} & -\partial_z V\,\overline{w'^2} & 0 \end{pmatrix} \tag{C1}$$

$$B_{ij} = \begin{pmatrix} 0 & 0 & \beta_3\,\overline{u'\theta'} \\ 0 & 0 & \beta_3\,\overline{v'\theta'} \\ \beta_3\,\overline{u'\theta'} & \beta_3\,\overline{v'\theta'} & 2\,\beta_3\,\overline{w'\theta'} \end{pmatrix} \tag{C2}$$

$$S_{ij} = \frac{1}{2}\begin{pmatrix} 0 & 0 & \partial_z U \\ 0 & 0 & \partial_z V \\ \partial_z U & \partial_z V & 0 \end{pmatrix} \tag{C3}$$

$$V_{ij} = \frac{1}{2}\begin{pmatrix} 0 & 0 & \partial_z U \\ 0 & 0 & \partial_z V \\ -\partial_z U & -\partial_z V & 0 \end{pmatrix} \tag{C4}$$

$$Z_{ij} = \begin{pmatrix} \overline{u'w'}\,\partial_z U & \frac{1}{2}\,\overline{v'w'}\,\partial_z U + \frac{1}{2}\,\overline{u'w'}\,\partial_z V & \frac{1}{2}\,\partial_z U\,(\overline{w'^2 - u'^2}) - \frac{1}{2}\,\partial_z V\,\overline{u'v'} \\ \frac{1}{2}\,\overline{v'w'}\,\partial_z U + \frac{1}{2}\,\overline{u'w'}\,\partial_z V & \overline{v'w'}\,\partial_z V & \frac{1}{2}\,\partial_z V\,(\overline{w'^2 - v'^2}) - \frac{1}{2}\,\partial_z U\,\overline{u'v'} \\ \frac{1}{2}\,\partial_z U\,(\overline{w'^2 - u'^2}) - \frac{1}{2}\,\partial_z V\,\overline{u'v'} & \frac{1}{2}\,\partial_z V\,(\overline{w'^2 - v'^2}) - \frac{1}{2}\,\partial_z U\,\overline{u'v'} & -\overline{u'w'}\,\partial_z U - \overline{v'w'}\,\partial_z V \end{pmatrix} \tag{C5}$$

Appendix D The Algebraic System of 10 Equations of the GLS Formalism

For clarity, we give here the explicit writing of the 10 equations presented in Eqs (6) - (8) and that are the basis of the GLS formalism:

$$
\begin{aligned}
0 &= -c_1\frac{\varepsilon}{k}(\overline{u'^2}-\frac{2}{3}k)+(1-c_2)(-\frac{4}{3}\,\overline{u'w'}\,\partial_z U+\frac{2}{3}\,\overline{v'w'}\,\partial_z V)-\frac{2}{3}(1-c_3)\beta_3\,\overline{w'\theta'}-c_5\overline{u'w'}\,\partial_z U\\
0 &= -c_1\frac{\varepsilon}{k}(\overline{v'^2}-\frac{2}{3}k)+(1-c_2)(-\frac{4}{3}\,\overline{v'w'}\,\partial_z V+\frac{2}{3}\,\overline{u'w'}\,\partial_z U)-\frac{2}{3}(1-c_3)\beta_3\,\overline{w'\theta'}-c_5\overline{v'w'}\,\partial_z V\\
0 &= -c_1\frac{\varepsilon}{k}(\overline{w'^2}-\frac{2}{3}k)+(\frac{2}{3}-\frac{2}{3}c_2+c_5)(\overline{u'w'}\,\partial_z U+\overline{v'w'}\,\partial_z V)+\frac{4}{3}(1-c_3)\beta_3\,\overline{w'\theta'}\\
0 &= -c_1\frac{\varepsilon}{k}\,\overline{u'v'}-(1-c_2)(\overline{v'w'}\,\partial_z U+\overline{u'w'}\,\partial_z V)-\frac{1}{2}c_5(\overline{v'w'}\,\partial_z U+\overline{u'w'}\,\partial_z V)\\
0 &= -c_1\frac{\varepsilon}{k}\,\overline{u'w'}-(1-c_2)\overline{w'^2}\,\partial_z U+(1-c_3)\beta_3\,\overline{u'\theta'}-\frac{1}{2}c_4k\,\partial_z U-\frac{1}{2}c_5(\overline{w'^2}\,\partial_z U-\overline{u'^2}\,\partial_z U-\overline{u'v'}\,\partial_z V)\\
0 &= -c_1\frac{\varepsilon}{k}\,\overline{v'w'}-(1-c_2)\overline{w'^2}\,\partial_z V+(1-c_3)\beta_3\,\overline{v'\theta'}-\frac{1}{2}c_4k\,\partial_z V-\frac{1}{2}c_5(\overline{w'^2}\,\partial_z V-\overline{v'^2}\,\partial_z V-\overline{u'v'}\,\partial_z U)\\
0 &= -c_{1T}\frac{\varepsilon}{k}\,\overline{u'\theta'}-(1-c_{2T}-\frac{1}{2}c_{4T})\,\overline{w'\theta'}\,\partial_z U-\overline{u'w'}\,\partial_z\Theta\\
0 &= -c_{1T}\frac{\varepsilon}{k}\,\overline{v'\theta'}-(1-c_{2T}-\frac{1}{2}c_{4T})\,\overline{w'\theta'}\,\partial_z V-\overline{v'w'}\,\partial_z\Theta\\
0 &= -c_{1T}\frac{\varepsilon}{k}\,\overline{w'\theta'}-\overline{w'^2}\,\partial_z\Theta+(1-c_{3T})\beta_3\,\overline{\theta'^2}-\frac{1}{2}c_{4T}(\overline{u'\theta'}\,\partial_z U+\overline{v'\theta'}\,\partial_z V)\\
0 &= -2\,\overline{w'\theta'}\,\partial_z\Theta-\frac{2}{c_T}\frac{\varepsilon}{k}\,\overline{\theta'^2}
\end{aligned}
\tag{D1}
$$

Appendix E Coefficients of the Stability Functions for the GLS Formalism

Coefficients $n_0, n_1, n_2, n_{0T}, n_{1T}, n_{2T}, d_0, d_1, d_2, d_3, d_4, d_5$ of the GLS stability functions (Eqs. (12) and (13)) have the following definitions:

$$
n_0 = \frac{4-4c_2+3c_4}{6c_1},
$$

$$
n_1 = \frac{c_1c_{1T}c_T(1-c_{3T})(4-4c_2+3c_4)-2c_1(1-c_3)(2-2c_{2T}-c_{4T})+4c_{1T}(1-c_3)(c_4-c_5)}{6c_1^2c_{1T}^2},
$$

$$
n_2 = \frac{-c_{4T}(4-4c_2+3c_4)(2-2c_{2T}-c_{4T})}{24c_1c_{1T}^2},
$$

$$
n_{0T} = \frac{2}{3c_{1T}}, \qquad n_{1T} = \frac{2(1-c_3)}{3c_1c_{1T}^2},
$$

$$
n_{2T} = \frac{c_1c_{4T}(4-4c_2+3c_4)+8c_5c_{1T}(1-c_2+c_5)-2c_4c_{1T}(2-2c_2+3c_5)}{12c_1^2c_{1T}^2},
$$

$$d_0 = 1, \qquad d_1 = \frac{7 - 7c_3 + 3c_1 c_T (1 - c_{3T})}{3c_1 c_{1T}},$$

$$d_2 = \frac{3c_5^2 + 6c_5(1 - c_2) + 2(1 - c_2)^2}{3c_1^2} - \frac{c_{4T}(2 - 2c_{2T} - c_{4T})}{4c_{1T}^2},$$

$$\begin{aligned} d_3 = &\frac{c_5 c_{1T}(1 - c_3)(2 - 2c_2 + c_5)}{3c_1^3 c_{1T}^2} \\ &+ \frac{c_1 c_{1T} c_T (1 - c_{3T})\Big(3c_5^2 + 6c_5(1 - c_2) + 2(1 - c_2)^2\Big)}{3c_1^3 c_{1T}^2} \\ &+ \frac{c_1(1 - c_3)\Big(3c_{4T}(1 - c_2 + c_5) - (1 - c_{2T})(2 - 2c_2 + 3c_5)\Big)}{3c_1^3 c_{1T}^2}, \end{aligned}$$

$$d_4 = \frac{(1 - c_3)(4 - 4c_3 + 3c_1 c_T (1 - c_{3T}))}{3c_1^2 c_{1T}^2},$$

$$d_5 = \frac{-c_{4T}(2 - 2c_{2T} - c_{4T})\Big(3c_5^2 + 6c_5(1 - c_2) + 2(1 - c_2)^2\Big)}{12c_1^2 c_{1T}^2}, \tag{E1}$$

Appendix F The Algebraic System of 9 Equations of the $k\varepsilon t$ Parameterization

For clarity, we give here the explicit writing of the 9 equations presented in Eqs (6) - (7) and that are the basis of the $k\varepsilon t$ parameterization:

$$
\begin{aligned}
0 &= -c_1 \frac{\varepsilon}{k}(\overline{u'^2} - \frac{2}{3}k) + (1-c_2)(-\frac{4}{3}\,\overline{u'w'}\,\partial_z U + \frac{2}{3}\,\overline{v'w'}\,\partial_z V) - \frac{2}{3}(1-c_3)\beta_3\,\overline{w'\theta'} - c_5\overline{u'w'}\,\partial_z U \\
0 &= -c_1 \frac{\varepsilon}{k}(\overline{v'^2} - \frac{2}{3}k) + (1-c_2)(-\frac{4}{3}\,\overline{v'w'}\,\partial_z V + \frac{2}{3}\,\overline{u'w'}\,\partial_z U) - \frac{2}{3}(1-c_3)\beta_3\,\overline{w'\theta'} - c_5\overline{v'w'}\,\partial_z V \\
0 &= -c_1 \frac{\varepsilon}{k}(\overline{w'^2} - \frac{2}{3}k) + (\frac{2}{3} - \frac{2}{3}c_2 + c_5)(\overline{u'w'}\,\partial_z U + \overline{v'w'}\,\partial_z V) + \frac{4}{3}(1-c_3)\beta_3\,\overline{w'\theta'} \\
0 &= -c_1 \frac{\varepsilon}{k}\,\overline{u'v'} - (1-c_2)(\overline{v'w'}\,\partial_z U + \overline{u'w'}\,\partial_z V) - \frac{1}{2}c_5(\overline{v'w'}\,\partial_z U + \overline{u'w'}\,\partial_z V) \\
0 &= -c_1 \frac{\varepsilon}{k}\,\overline{u'w'} - (1-c_2)\overline{w'^2}\,\partial_z U + (1-c_3)\beta_3\,\overline{u'\theta'} - \frac{1}{2}c_4 k\,\partial_z U - \frac{1}{2}c_5(\overline{w'^2}\,\partial_z U - \overline{u'^2}\,\partial_z U - \overline{u'v'}\,\partial_z V) \\
0 &= -c_1 \frac{\varepsilon}{k}\,\overline{v'w'} - (1-c_2)\overline{w'^2}\,\partial_z V + (1-c_3)\beta_3\,\overline{v'\theta'} - \frac{1}{2}c_4 k\,\partial_z V - \frac{1}{2}c_5(\overline{w'^2}\,\partial_z V - \overline{v'^2}\,\partial_z V - \overline{u'v'}\,\partial_z U) \\
0 &= -c_{1T} \frac{\varepsilon}{k}\,\overline{u'\theta'} - (1 - c_{2T} - \frac{1}{2}c_{4T})\,\overline{w'\theta'}\,\partial_z U - \overline{u'w'}\,\partial_z \Theta \\
0 &= -c_{1T} \frac{\varepsilon}{k}\,\overline{v'\theta'} - (1 - c_{2T} - \frac{1}{2}c_{4T})\,\overline{w'\theta'}\,\partial_z V - \overline{v'w'}\,\partial_z \Theta \\
0 &= -c_{1T} \frac{\varepsilon}{k}\,\overline{w'\theta'} - \overline{w'^2}\,\partial_z \Theta + (1-c_{3T})\beta_3\,\overline{\theta'^2} - \frac{1}{2}c_{4T}(\overline{u'\theta'}\,\partial_z U + \overline{v'\theta'}\,\partial_z V)
\end{aligned}
\tag{F1}
$$

Appendix G Coefficients of the Stability Functions of the $k\varepsilon t$ Parameterization

We give hereafter the expressions of the coefficients $n_0, n_1, n_2, n_{0T}, n_{1T}, n_{2T}, n^*_{0T}, n^*_{1T}, n^*_{2T}, d_0, d_1, d_2, d_3, d_4, d_5$ of the $k\varepsilon t$ stability functions (Eqs. (21) - (23)). We point out that the expressions of the coefficient n_{1T} and all the coefficients not multiplying α_n (i.e. n_0, n_2, n_{0T}, n_{2T}, d_0, d_1 and d_5) stay unchanged compared to the GLS ones (given in Appendix E).

$$
n_0 = \frac{4 - 4c_2 + 3c_4}{6c_1}, \qquad n_1 = \frac{(1-c_3)\Big(2c_{1T}(c_4 - c_5) - c_1(2 - 2c_{2T} - c_{4T})\Big)}{3c_1^2 c_{1T}^2},
$$

$$
n_2 = \frac{-c_{4T}(4 - 4c_2 + 3c_4)(2 - 2c_{2T} - c_{4T})}{24 c_1 c_{1T}^2},
$$

$$
n_3 = \frac{(1-c_3)(1-c_{3T})\Big(2c_{1T}(4 - 4c_2 + 3c_5) + 3c_1(2 - 2c_{2T} - c_{4T})\Big)}{6c_1^2 c_{1T}^2},
$$

$$
n_{0T} = \frac{2}{3c_{1T}}, \qquad n_{1T} = \frac{2(1-c_3)}{3c_1 c_{1T}^2},
$$

$$n_{2T} = \frac{c_1 c_{4T}(4 - 4c_2 + 3c_4) + 8c_5 c_{1T}(1 - c_2 + c_5) - 2c_4 c_{1T}(2 - 2c_2 + 3c_5)}{12 c_1^2 c_{1T}^2},$$

$$n_{0T}^* = \frac{1 - c_{3T}}{c_{1T}}, \qquad n_{1T}^* = \frac{(1 - c_3)(1 - c_{3T})}{c_1 c_{1T}^2},$$

$$n_{2T}^* = \frac{(1 - c_{3T})\Big(3c_5^2 + 6c_5(1 - c_2) + 2(1 - c_2)^2\Big)}{3 c_1^2 c_{1T}},$$

$$d_0 = 1, \qquad d_1 = \frac{7(1 - c_3)}{3 c_1 c_{1T}}, \qquad d_2 = \frac{3c_5^2 + 6c_5(1 - c_2) + 2(1 - c_2)^2}{3c_1^2} - \frac{c_{4T}(2 - 2c_{2T} - c_{4T})}{4c_{1T}^2}$$

$$d_3 = \frac{(1 - c_3)\Big(3c_1 c_{4T}(1 - c_2 + c_5) + c_5 c_{1T}(2 - 2c_2 + c_5) - c_1(1 - c_{2T})(2 - 2c_2 + 3c_5)\Big)}{3 c_1^3 c_{1T}^2}$$

$$d_4 = \frac{4(1 - c_3)^2}{3 c_1^2 c_{1T}^2}, \qquad d_5 = \frac{-c_{4T}(2 - 2c_{2T} - c_{4T})\Big(3c_5^2 + 6c_5(1 - c_2) + 2(1 - c_2)^2\Big)}{12 c_1^2 c_{1T}^2}. \tag{G1}$$

Appendix H Results of the $k - \varepsilon - \overline{\theta'^2} - \overline{w'^2}$ model

We detail here the results of the k–ε–$\overline{\theta'^2}$–$\overline{w'^2}$ model which is a possible extension of the $k\varepsilon t$ model where the non-equilibrium is also considered for the $\overline{w'^2}$ equation. By doing that, equations (6) and (7) now form a system of 8 equations with 8 unknowns: ($\overline{u'^2}$, $\overline{v'^2}$, $\overline{u'v'}$, $\overline{u'w'}$, $\overline{v'w'}$, $\overline{u'\theta'}$, $\overline{v'\theta'}$, $\overline{w'\theta'}$). We solved this system with Mathematica and we obtained the following expressions:

$$\overline{u'w'} = -c_\mu \frac{k^2}{\varepsilon} \partial_z U, \tag{H1}$$

$$\overline{v'w'} = -c_\mu \frac{k^2}{\varepsilon} \partial_z V, \tag{H2}$$

$$\overline{w'\theta'} = -c'_\mu \frac{k^2}{\varepsilon} \partial_z \Theta + c'^*_\mu \frac{k}{\varepsilon} \beta_3 \, \overline{\theta'^2}, \tag{H3}$$

which have the same shape as the ones found for the $k\varepsilon t$ model. Particularly, even if $\overline{w'^2}$ is not in equilibrium anymore, the velocity fluxes $\overline{u'w'}$ and $\overline{v'w'}$ are still fully downgradient. The expressions of the stability functions c_μ, c'_μ and c'^*_μ are:

$$c_\mu = \frac{n_0 + n_2\alpha_M + n_3\alpha_T + n_4\alpha_W + n_5\alpha_W\alpha_N + n_6\alpha_W\alpha_M}{d_0 + d_1\alpha_N + d_2\alpha_M + d_3\alpha_N\alpha_M + d_5\alpha_M^2}, \tag{H4}$$

$$c'_\mu = \frac{n_{2T}\alpha_M + n_{4T}\alpha_W + n_{5T}\alpha_W\alpha_N + n_{6T}\alpha_W\alpha_M}{d_0 + d_1\alpha_N + d_2\alpha_M + d_3\alpha_N\alpha_M + d_5\alpha_M^2}, \tag{H5}$$

$$c'^*_\mu = \frac{n^*_{0T} + n^*_{1T}\alpha_N + n^*_{2T}\alpha_M}{d_0 + d_1\alpha_N + d_2\alpha_M + d_3\alpha_N\alpha_M + d_5\alpha_M^2}, \tag{H6}$$

with $\alpha_N = \frac{k^2}{\varepsilon^2}N^2$, $\alpha_M = \frac{k^2}{\varepsilon^2}M^2$, $\alpha_T = \frac{k}{\varepsilon^2}\beta_3^2\,\overline{\theta'^2}$, and $\alpha_W = \frac{1}{k}\,\overline{w'^2}$. Coefficients n_i, n_{iT} and d_i depends on the coefficients c_i and c_{iT}; the expressions are given hereafter. Taking the values of the c_i and c_{iT} given in Table 1, the stability functions are approximately as follows

$$c_\mu = \frac{0.04693 - 0.00005303\alpha_M + 0.001996\alpha_T + 0.0896\alpha_W - 0.002994\alpha_W\alpha_N - 0.0001012\alpha_W\alpha_M}{1 + 0.03361\alpha_N + 0.01342\alpha_M + 0.00006267\alpha_N\alpha_M - 0.00001644\alpha_M^2}, \tag{H7}$$

$$c'_\mu = \frac{0.0002651\alpha_M + 0.1681\alpha_W + 0.005649\alpha_W\alpha_N + 0.002952\alpha_W\alpha_M}{1 + 0.03361\alpha_N + 0.01342\alpha_M + 0.00006267\alpha_N\alpha_M - 0.00001644\alpha_M^2}, \tag{H8}$$

$$c'^*_\mu = \frac{0.1120 + 0.003766\alpha_N + 0.001631\alpha_M}{1 + 0.03361\alpha_N + 0.01342\alpha_M + 0.00006267\alpha_N\alpha_M - 0.00001644\alpha_M^2}. \tag{H9}$$

Here are the expressions of the coefficients n_i, n_{iT}, and d_i:

$$n_0 = \frac{3c_4 - 2c_5}{6c_1}, \qquad n_2 = \frac{-c_{4T}(3c_4 - 2c_5)(2 - 2c_{2T} - c_{4T})}{24c_1c_{1T}^2},$$

$$n_3 = \frac{(1-c_3)(1-c_{3T})\Big(2c_{1T}c_5 + 3c_1(2 - 2c_{2T} - c_{4T})\Big)}{6c_1^2c_{1T}^2}, \qquad n_4 = \frac{2 - 2c_2 + c_5}{2c_1},$$

$$n_5 = \frac{-(1-c_3)\Big(2c_{1T}c_5 + 3c_1(2 - 2c_{2T} - c_{4T})\Big)}{6c_1^2c_{1T}^2}, \qquad n_6 = \frac{-c_{4T}(2 - 2c_2 + c_5)(2 - 2c_{2T} - c_{4T})}{8c_1c_{1T}^2},$$

$$n_{2T} = \frac{c_{4T}(3c_4 - 2c_5)}{12c_1 c_{1T}^2}, \qquad n_{4T} = \frac{1}{c_{1T}}, \qquad n_{5T} = \frac{1 - c_3}{c_1 c_{1T}^2},$$

$$n_{6T} = \frac{3c_1 c_{4T}(2 - 2c_2 + c_5) + 2c_5 c_{1T}(4 - 4c_2 + 3c_5)}{12c_1^2 c_{1T}^2},$$

$$n_{0T}^* = \frac{1 - c_{3T}}{c_{1T}}, \qquad n_{1T}^* = \frac{(1 - c_3)(1 - c_{3T})}{c_1 c_{1T}^2}, \qquad n_{2T}^* = \frac{c_5(1 - c_{3T})(4 - 4c_2 + 3c_5)}{6c_1^2 c_{1T}},$$

$$d_0 = 1, \qquad d_1 = \frac{1 - c_3}{c_1 c_{1T}}, \qquad d_2 = \frac{c_5(4 - 4c_2 + 3c_5)}{6c_1^2} - \frac{c_{4T}(2 - 2c_{2T} - c_{4T})}{4c_{1T}^2}$$

$$d_3 = \frac{c_5 c_{4T}(1 - c_3)}{6c_1^2 c_{1T}^2}, \qquad d_5 = \frac{-c_5 c_{4T}(4 - 4c_2 + 3c_5)(2 - 2c_{2T} - c_{4T})}{24c_1^2 c_{1T}^2}. \tag{H10}$$

Acknowledgments

We thank Florian Lemarié for providing the 1D code. All the computations presented in this paper were performed using the GRICAD infrastructure (https://gricad.univ-grenoble-alpes.fr), which is supported by Grenoble research communities.

References

Arakawa, A., & Schubert, W. H. (1974). Interaction of a cumulus cloud ensemble with the large-scale environment, part I. *J. Atmos. Sci.*, *31*(3), 674–701. doi: 10.1175/1520-0469(1974)031⟨0674:IOACCE⟩2.0.CO;2

Burchard, H. (2002). *Applied Turbulence Modelling in Marine Waters* (Vol. 100; S. Bhattacharji, G. M. Friedman, H. J. Neugebauer, & A. Seilacher, Eds.). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/3-540-45419-5

Burchard, H., & Baumert, H. (1995). On the performance of a mixed-layer model based on the $\kappa - \epsilon$ turbulence closure. *J. Geophys. Res.*, *100*(C5), 8523–8540. doi: 10.1029/94JC03229

Burchard, H., & Bolding, K. (2001). Comparative Analysis of Four Second-Moment Turbulence Closure Models for the Oceanic Mixed Layer. *Journal of Physical Oceanography*, *31*(8), 1943–1968. doi: 10.1175/1520-0485(2001)031⟨1943:CAOFSM⟩2.0.CO;2

Burchard, H., & Petersen, O. (1999). Models of turbulence in the marine environment —a comparative study of two-equation turbulence models. *Journal of Marine Systems*, *21*(1-4), 29–53. doi: 10.1016/S0924-7963(99)00004-4

Canuto, V. M., Howard, A., Cheng, Y., & Dubovikov, M. S. (2001). Ocean Turbulence. Part I: One-Point Closure Model—Momentum and Heat Vertical Diffusivities. *Journal of Physical Oceanography*, *31*(6), 1413–1426. doi: 10.1175/1520-0485(2001)031⟨1413:OTPIOP⟩2.0.CO;2

Cheng, Y., Canuto, V. M., Howard, A. M., Ackerman, A. S., Kelley, M., Fridlind, A. M., … Elsaesser, G. S. (2020). A Second-Order Closure Turbulence Model: New Heat Flux Equations and No Critical Richardson Number. *Journal of the Atmospheric Sciences*, *77*(8), 2743–2759. doi: 10.1175/JAS-D-19-0240.1

Deardorff, J. W. (1972). Theoretical expression for the countergradient vertical heat flux. *J. Geophys. Res.*, *77*(30), 5900–5904. doi: 10.1029/JC077i030p05900

Emanuel, K. A. (1991). A scheme for representing cumulus convection in large-scale models. *J. Atmos. Sci.*, *48*(21), 2313–2329. doi: 10.1175/1520-0469(1991)048⟨2313:ASFRCC⟩2.0.CO;2

Fearon, G., Herbette, S., Veitch, J., Cambon, G., Lucas, A. J., Lemarié, F., & Vichi, M. (2020). Enhanced vertical mixing in coastal upwelling systems driven by diurnal-inertial resonance: Numerical experiments. *Journal of Geophysical Research: Oceans*, *125*(9), e2020JC016208. doi: 10.1029/2020JC016208

Fox-Kemper, B., Ferrari, R., & Hallberg, R. (2008). Parameterization of mixed layer eddies. part I: Theory and diagnosis. *J. Phys. Oceanogr.*, *38*(6), 1145. doi: 10.1175/2007JPO3792.1

Garanaik, A., Pereira, F. S., Smith, K., Robey, R., Li, Q., Pearson, B., & Van Roekel, L. (2024). A New Hybrid Mass-Flux/High-Order Turbulence Closure for Ocean Vertical Mixing. *Journal of Advances in Modeling Earth Systems*, *16*(1), e2023MS003846. doi: 10.1029/2023MS003846

Garcia, J. R., & Mellado, J. P. (2014). The two-layer structure of the entrainment zone in the convective boundary layer. *J. Atmos. Sci.*, *71*(6), 1935–1955. doi:

10.1175/JAS-D-13-0148.1

Gaspar, P., Grégoris, Y., & Lefevre, J.-M. (1990). A simple eddy kinetic energy model for simulations of the oceanic vertical mixing: Tests at station papa and long-term upper ocean study site. *J. Geophys. Res.*, *95*, 16. doi: 10.1029/JC095iC09p16179

Ghannam, K., Duman, T., Salesky, S. T., Chamecki, M., & Katul, G. (2017). The non-local character of turbulence asymmetry in the convective atmospheric boundary layer. *Quarterly Journal of the Royal Meteorological Society*, *143*(702), 494–507. doi: 10.1002/qj.2937

Gibbs, J. A., Fedorovich, E., & Van Eijk, A. M. J. (2011). Evaluating Weather Research and Forecasting (WRF) Model Predictions of Turbulent Flow Parameters in a Dry Convective Boundary Layer. *Journal of Applied Meteorology and Climatology*, *50*, 2429–2444.

Giordani, H., Bourdallé-Badie, R., & Madec, G. (2020). An eddy-diffusivity mass-flux parameterization for modeling oceanic convection. *J. Adv. Model. Earth Syst.*, *12*(9), e02078. doi: 10.1029/2020MS002078

Hanjalić, K., & Launder, B. E. (1972). A Reynolds stress model of turbulence and its application to thin shear flows. *Journal of Fluid Mechanics*, *52*(4), 609–638. doi: 10.1017/S002211207200268X

Hourdin, F., Couvreux, F., & Menut, L. (2002). Parameterization of the dry convective boundary layer based on a mass flux representation of thermals. *J. Atmos. Sci.*, *59*(6), 1105–1123. doi: 10.1175/1520-0469(2002)059⟨1105:POTDCB⟩2.0.CO;2

Large, W. G., McWilliams, J. C., & Doney, S. C. (1994). Oceanic vertical mixing: a review and a model with a nonlocal boundary layer parameterization. *Rev. Geophys.*, *32*, 363–404. doi: 10.1029/94RG01872

Lazier, J. (2001). Deep Convection. In *Encyclopedia of Ocean Sciences* (pp. 634–643). Elsevier. doi: 10.1006/rwos.2001.0113

Legay, A., Deremble, B., Penduff, T., Brasseur, P., & Molines, J.-M. (2024). A framework for evaluating ocean mixed layer depth evolution. *J. Adv. Model. Earth Syst.*, submitted. doi: 10.22541/essoar.168563421.17506622/v2

Lemarié, F., Samson, G., Redelsperger, J.-L., Giordani, H., Brivoal, T., & Madec, G. (2021). A simplified atmospheric boundary layer model for an improved representation of air–sea interactions in eddying oceanic models: implementation and first evaluation in NEMO (4.0). *Geoscientific Model Development*, *14* (1), 543–572. doi: 10.5194/gmd-14-543-2021

Luyten, J., Pedlosky, J., & Stommel, H. (1983). The ventilated thermocline. *J. Phys. Oceanogr.*, *13* (2), 292–309.

Marshall, J., & Schott, F. (1999). Open-ocean convection: Observations, theory, and models. *Reviews of geophysics*, *37* (1), 1–64. doi: 10.1029/98RG02739

Mellor, G. L., & Yamada, T. (1982). Development of a turbulence closure model for geophysical fluid problems. *Reviews of Geophysics*, *20* (4), 851. doi: 10.1029/RG020i004p00851

Mironov, D. V., Gryanik, V. M., Moeng, C., Olbers, D. J., & Warncke, T. H. (2000). Vertical turbulence structure and second-moment budgets in convection with rotation: A large-eddy simulation study. *Quarterly Journal of the Royal Meteorological Society*, *126* (563), 477–515. doi: 10.1002/qj.49712656306

Patankar, S. (1980). *Numerical heat transfer and fluid flow.* New York: McGraw-Hill.

Popinet, S. (2020). Basilisk flow solver and pde library. *available at available at http://basilisk. fr*.

Reffray, G., Bourdalle-Badie, R., & Calone, C. (2015). Modelling turbulent vertical mixing sensitivity using a 1-D version of NEMO. *Geoscientific Model Development*, *8* (1), 69–86. doi: 10.5194/gmd-8-69-2015

Reichl, B. G., & Hallberg, R. (2018). A simplified energetics based planetary boundary layer (ePBL) approach for ocean climate simulations. *Ocean Model.*, *132*, 112-129. doi: 10.1016/j.ocemod.2018.10.004

Rodi, W. (1987). Examples of calculation methods for flow and mixing in stratified fluids. *Journal of Geophysical Research*, *92* (C5), 5305. doi: 10.1029/JC092iC05p05305

Souza, A. N., Wagner, G. L., Ramadhan, A., Allen, B., Churavy, V., Schloss, J., ... Ferrari, R. (2020). Uncertainty quantification of ocean parameterizations: Application to the K-Profile-Parameterization for penetrative convection. *J. Adv. Model. Earth Syst.*, *12* (12), e2020MS002108. doi: 10.1029/2020MS002108

Stull, R. B. (1988). *An introduction to boundary layer meteorology.* doi: 10.1007/978-94-009-3027-8

Thangam, S., Abid, R., & Speziale, C. G. (1992). Application of a new K-tau model to near wall turbulent flows. *AIAA Journal*, *30*(2), 552–554. doi: 10.2514/3.10952

Treguier, A. M., de Boyer Montégut, C., Bozec, A., Chassignet, E. P., Fox-Kemper, B., Hogg, A. M., … Yeager, S. (2023). The mixed layer depth in the ocean model intercomparison project (omip): Impact of resolving mesoscale eddies. *EGUsphere*, *2023*, 1–43. doi: 10.5194/egusphere-2023-310

Troen, I. B., & Mahrt, L. (1986). A simple model of the atmospheric boundary layer; sensitivity to surface evaporation. *Bound.-Lay. Meteorol.*, *37*(1-2), 129-148. doi: 10.1007/BF00122760

Umlauf, L., & Burchard, H. (2003). A generic length-scale equation for geophysical turbulence models. *Journal of Marine Research*, *61*(2), 235–265. doi: 10.1357/002224003322005087

Umlauf, L., & Burchard, H. (2005). Second-order turbulence closure models for geophysical boundary layers. A review of recent work. *Continental Shelf Research*, *25*(7-8), 795–827. doi: 10.1016/j.csr.2004.08.004

Umlauf, L., Burchard, H., & Hutter, K. (2003). Extending the k–omega turbulence model towards oceanic applications. *Ocean Modelling*, *5*(3), 195–218. doi: 10.1016/S1463-5003(02)00039-2

Van Roekel, L., Adcroft, A. J., Danabasoglu, G., Griffies, S. M., Kauffman, B., Large, W., … Schmidt, M. (2018). The KPP Boundary Layer Scheme for the Ocean: Revisiting Its Formulation and Benchmarking One-Dimensional Simulations Relative to LES. *Journal of Advances in Modeling Earth Systems*, *10*(11), 2647–2685. doi: 10.1029/2018MS001336

Wagner, G., Hillier, A., Constantinou, N. C., Silvestri, S., Souza, A., Burns, K., … Ferrari, R. (2023). CATKE: a turbulent-kinetic-energy-based parameterization for ocean microturbulence with dynamic convective adjustment. *arXiv e-prints*, arXiv:2306.13204. doi: 10.48550/arXiv.2306.13204

Warner, J. C., Sherwood, C. R., Arango, H. G., & Signell, R. P. (2005). Performance of four turbulence closure models implemented using a generic length scale method. *Ocean Modelling*, *8*(1-2), 81–113. doi: 10.1016/j.ocemod.2003.12.003

Wilcox, D. C. (1988). Reassessment of the scale-determining equation for advanced turbulence models. *AIAA Journal*, *26*(11), 1299–1310. doi: 10.2514/3.10041

Williams, R. G., Marshall, J. C., & Spall, M. A. (1995). Does stommel's mixed layer "demon" work? *J. Phys. Oceanogr.*, *25*(12), 3089–3102. doi: 10.1175/1520-0485(1995)025⟨3089:DSMLW⟩2.0.CO;2

Willis, G., & Deardorff, J. (1974). A laboratory model of the unstable planetary boundary layer. *Journal of Atmospheric Sciences*, *31*(5), 1297–1307. doi: 10.1175/1520-0469(1974)031⟨1297:ALMOTU⟩2.0.CO;2

Zeierman, S., & Wolfshtein, M. (1986). Turbulent time scale for turbulent-flow calculations. *AIAA Journal*, *24*(10), 1606–1610. doi: 10.2514/3.9490

Zhou, B., Sun, S., Yao, K., & Zhu, K. (2018). Reexamining the Gradient and Countergradient Representation of the Local and Nonlocal Heat Fluxes in the Convective Boundary Layer. *Journal of the Atmospheric Sciences*, *75*(7), 2317–2336. doi: 10.1175/JAS-D-17-0198.1