

## RESEARCH ARTICLE

10.1029/2018JC014307

## Key Points:

- Turbulence closures are assessed in 3-D simulations in tidally energetic environment
- In the shelf seas, all closures simulate a seasonal pycnocline that is too shallow and bottom temperatures that are too warm; there is no clear winner out of the choice of closures
- Choice of turbulence closure affects the ecosystem behavior, shifting the timing of peak chlorophyll by 1 month and regional chlorophyll differences of order 100%

## Supporting Information:

- Data Set S1
- Data Set S2

## Correspondence to:

M. V. Luneva,  
manel@noc.ac.uk

## Citation:

Luneva, M. V., Wakelin, S., Holt, J. T., Inall, M. E., Kozlov, I. E., Palmer, M. R., et al. (2019). Challenging vertical turbulence mixing schemes in a tidally energetic environment: 1. 3-D shelf-sea model assessment. *Journal of Geophysical Research: Oceans*, 124, 6360–6387. <https://doi.org/10.1029/2018JC014307>

Received 2 AUG 2018

Accepted 29 MAY 2019

Accepted article online 1 JUL 2019

Published online 31 AUG 2019

## Challenging Vertical Turbulence Mixing Schemes in a Tidally Energetic Environment: 1. 3-D Shelf-Sea Model Assessment

Maria V. Luneva<sup>1</sup> , Sarah Wakelin<sup>1</sup> , Jason T. Holt<sup>1</sup> , Mark E. Inall<sup>2</sup> , Igor E. Kozlov<sup>3,4</sup> , Matthew R. Palmer<sup>1</sup> , Matthew Toberman<sup>2</sup> , Evgenia V. Zubkova<sup>3</sup>, and Jeff A. Polton<sup>1</sup> 

<sup>1</sup>National Oceanography Centre, Liverpool, UK, <sup>2</sup>Scottish Alliance for Geoscience Environment and Society (SAGES), Scottish Association for Marine Science, Oban, UK, <sup>3</sup>Satellite Oceanography Laboratory, Russian State Hydrometeorological University, Saint Petersburg, Russia, <sup>4</sup>Marine Hydrophysical Institute of RAS, Sevastopol, Russia

**Abstract** Mixing in the ocean and shelf seas is critical for the vertical distribution of dynamically active properties, such as density and biogeochemical tracers. Eight different decadal simulations are used to assess the skill of vertical turbulent mixing schemes (TMS) in a 3-D regional model of tidally active shelf seas. The TMS differ in the type of stability functions used and in the Ozmidov/Deardorff/Galperin limiter of the turbulence length scales. We review the dependence of the critical Richardson and Prandtl numbers to define the “diffusiveness” of the TMS. The skill in representing bias and variability of stratification profiles is assessed with five different metrics: surface and bottom temperatures and pycnocline depth, thickness, and strength. The assessment is made against hydrography from three data sets (28,000 profiles in total). Bottom and surface temperatures are found to be as sensitive to TMS choice as to horizontal resolution or heat flux formulation, as reported in other studies. All TMS underrepresent the pycnocline depth and benthic temperatures. This suggests physical processes are missing from the model, and these are discussed. Different TMSs show the best results for different metrics, and there is no outright winner. Simulations coupled with an ecosystem model show the choice of TMS strongly affects the ecosystem behavior: shifting the timing of peak chlorophyll by 1 month, showing regional chlorophyll differences of order 100%, and redistributing the production of chlorophyll between the pycnocline and mixed layer.

### 1. Introduction

Seasonal stratification is a ubiquitous feature of the upper ocean. Stable stratification typically results from increased summer heating and reduced wind stress and convection, which suppresses mixing as well as the exchange of heat, momentum, and nutrients between the surface layer and deeper ocean. Tidally energetic seas alternate between well-mixed and stratified states, both seasonally and across tidal mixing fronts. Strong summer stratification isolates the upper layer from the tidally driven benthic boundary layer, affecting ocean-atmosphere exchange (e.g., of carbon and oxygen) and biological activity (Rippeth et al. 2005, 2014, 2009). Indeed, cross-pycnocline mixing becomes the main mechanism to resupply nutrients to the euphotic layer and supports continued summertime primary production. Hence, accurately modeling mixing in seasonally stratified seas is critical to modeling ecosystem processes and understanding the environmental impacts of future change (Holt et al., 2016).

Shelf seas modeling requires simulation of the dynamics of the benthic tidally driven layer, the upper mostly wind- and buoyancy-driven layer, and the pycnocline. Small-scale turbulence and diapycnal mixing in shelf seas vary in time and space, with time scales of tidal cycle and spring neap modulations. Short-term (~1- to 3-hr) mixing spikes in the thermocline associated with the alignment of the wind with inertially rotating surface currents or tidal shear have been observed, with temporally averaged nutrient fluxes increasing by between 4 and 20 times during the shear bursts ((Burchard & Rippeth, 2009; Lenn et al., 2011; Williams, Sharples, Green, et al., 2013; Williams, Sharples, Mahaffey, et al., 2013). Strong tides produce areas mixed from the surface to bottom, forming tidal fronts and associated jet currents. To simulate time and spatially dependent profiles of vertical diffusion/viscosity, many of the widely used regional models (e.g., Carniel et al., 2009; O’Dea et al., 2017) include a generic length-scale “GLS” (Umlauf & Burchard, 2005) module. The GLS module (conveniently deployed as a stand-alone model in the General Ocean Turbulence Model,

GOTM, at gotm.net) proposes the choice of turbulence model with two prognostic equations along with various closures. The prognostic equations describe the evolution of turbulent kinetic energy (TKE) and turbulent length scale (or their combination) with time.

The closures describe the dependence of diffusivity/viscosity on velocity shear and stratification and are central to the modeling assumptions made in the mixing scheme, particularly under stratified conditions. There has been significant recent progress in closure development (e.g., Canuto et al., 2010), which now includes parameterizations of double diffusion, Langmuir circulation (LC; Harcourt, 2015, Kantha & Clayson, 2004), and wave breaking (Gerbi et al., 2009). All these closures are derived from the assumption of equilibrium and locality, which assumes a balance between production, dissipation, and transfer of energy between components of Reynolds stresses and heat/salt fluxes, while time derivative terms and turbulent transports are neglected. Existing nonlocal models (e.g., Cheng et al., 2005) are still too complex and to our knowledge are not included in any current ocean circulation model.

Most turbulent closures have been verified through the simulation of mixed layer depth (MLD) and sea surface temperature (SST), with little systematic assessment of their impact on the whole vertical structure of stratification and pycnocline properties in 3-D simulations.

This paper aims to fill this gap and assess widely used closures in the context of shelf sea applications. Using the GLS routines in the Nucleus for European Modelling of the Oceans (NEMO) system (Madec & NEMO team, 2016), comparisons are made between simulated and observed characteristics of the stratification. Seasonally stratified shelf seas often have an approximately two water layer structure. We characterize stratification profiles by five characteristics: SST, benthic temperature, depth and thickness of the pycnocline, and the strength of the pycnocline, expressed by the maximum Brunt-Väisälä frequency. To assess model properties in multidimensional space of parameters, we employ Taylor (2001) diagrams, and, also, we introduce a “model skill” measure combining errors in variability and biases of all characteristics in one dimensionless parameter (Holloway et al., 2011).

The pycnocline is a critical biogeochemical interface, and so we also evaluate the effects of closure type on simulated ecosystem properties, using the Earth and Regional Seas Ecosystem Model (ERSEM; Butenschön et al., 2016) directly coupled to these simulations. To our knowledge, assessment of vertical (diapycnal) mixing schemes with a 3-D dynamically active ocean model coupled with ecosystem model has not previously been documented.

The paper is organized as follows: in section 2 we describe the turbulence closures under consideration; section 3 describes the model setup, validation data, and methods of comparisons; section 4 contains an inter-comparison between model runs and comparison with hydrography on decadal and shorter time scales; in section 5 we discuss the implication of the results and the effects of turbulence closures on an ecosystem model simulation. Conclusions are presented in section 6, and section 7 describes code availability.

## 2. Review of Vertical Turbulence Closures

### 2.1. General Approach

Turbulent coefficients of vertical diffusion of heat and salt  $k_h, k_s$  and viscosity  $k_{mu}, k_{mv}$  are defined as

$$\begin{pmatrix} k_{mu} \\ k_{mv} \\ k_h \\ k_s \end{pmatrix} = - \begin{pmatrix} \overline{u'w'}/\partial_z U \\ \overline{v'w'}/\partial_z V \\ \overline{T'w'}/\partial_z T \\ \overline{S'w'}/\partial_z S \end{pmatrix}, \quad (1)$$

where the second-order turbulent moments  $\overline{u'w'}$ ,  $\overline{v'w'}$ ,  $\overline{T'w'}$ , and  $\overline{S'w'}$  are Reynolds stresses in the east and northward directions and vertical turbulent heat and salt fluxes, respectively. The partial derivative with respect to depth is denoted by  $\partial_z$ . Variables  $U$ ,  $V$ ,  $T$ , and  $S$  are ensemble means over all sizes of turbulent eddies of velocity components, temperature, and salinity. Correspondingly,  $u'$ ,  $v'$ ,  $w'$ ,  $T'$ , and  $S'$  are fluctuating, turbulent components; overlining in the right-hand side of (1) denotes averaging over an ensemble of turbulent eddies. In isotropic fully developed turbulence these coefficients ( $k_{mu,mv,T,S}$ ) are the function of two turbulent scalar characteristics with independent dimensions. Without losing generality, we have chosen

turbulent velocity scales,  $q = (\overline{u'^2} + \overline{v'^2} + \overline{w'^2})^{1/2}$ , and turbulent length scale,  $l$ , as the governing parameters. From scaling arguments, following Kolmogorov (1942), one gets

$$k_{mu,mv,T,S} = ql A_{u,v,T,S} \quad (2)$$

where  $A_{u,v,T,S}$  are constants in this particular case. More generally, when the turbulence is anisotropic (perhaps due to shear, rotation, and/or stratification),  $A_{u,v,T,S}$  are not constants and are the so-called stability functions. These stability functions then become dependent on dimensionless parameters that characterize the sources of anisotropy of the flow, such as velocity squared shear ( $\Sigma^2 = \partial_z U^2 + \partial_z V^2$ ), stratification (defined by  $N^2 = g\rho_0^{-1} \partial_z \rho$ ), or frequency of rotation  $\Omega$ . Here  $\rho$  is density,  $\rho_0$  is the reference density, and  $g$  is gravitational acceleration. For these particular sources of anisotropy, the dimensionless parameters and turbulence-scale factors are

$$G_m = \tau^2 \Sigma^2, G_h = \tau^2 N^2, G_\Omega = \tau \Omega \quad (3)$$

where  $\tau = l/q$  is a turbulent time scale. The stability functions are then

$$A_{u,v,T,S} = A_{u,v,T,S}(G_m, h, \Omega, \dots) \quad (4)$$

The most important dimensionless characteristics of turbulent flows are the Richardson number,  $Ri$ , the flux Richardson number,  $Rf$ , and the Prandtl number,  $Pr$ . These are defined as follows:

$$Ri = \frac{N^2}{\Sigma^2} = \frac{G_h}{G_m}; \quad Rf = \frac{g \overline{\rho' w'}}{u' w' \partial_z U + v' w' \partial_z V} = \frac{Ri}{Pr}; \quad Pr = A_u / A_T. \quad (5)$$

Linear stability theory (Miles, 1961; Howard, 1961) for a flow with constant steady shear predicts the necessary, but not sufficient, condition for stability to be

$$Ri > Ri_{cr} = 1/4 \quad (6)$$

Nonlinear analysis, considered by Abarbanel et al. (1984), provides necessary and sufficient conditions for stability to be

$$Ri > Ri_{cr} = 1 \quad (7)$$

However, when internal waves are present, the stability condition of plane waves becomes (Orlanski & Bryan, 1969)

$$Ri > Ri_{cr} = 1 + |k_x/k_z| > 1 \quad (8)$$

where  $k_z, k_x$  are vertical and horizontal components of the wave number. Hence, the stability condition with internal waves present (8) is at least 4 times weaker than the condition induced by constant shear (6).

To close the problem (2)–(4), one must define the two-dimensional independent turbulent scales and find the associated stability functions (4). The hierarchy of turbulent models, based on ideas by Kolmogorov (1942), Prandtl and Wieghart (1945), and Rotta (1951), relies on the principle of separation of variables into turbulent and ensemble mean parts and further derivation of the set of equations of higher-order turbulent moments. This approach was first developed in the late 1960s and early 1970s when two-equation models became a practical tool for calculation of turbulent flows in engineering applications. Launder and Jones (1969) proposed a  $(k - \epsilon)$  model which solves prognostic equations for TKE  $k = q^2/2$  and its dissipation rate  $\epsilon$  (units of square meters per cubic second). Instead of  $(k - \epsilon)$  equations, Mellor and Yamada (1974, 1982) employed prognostic equations for  $q^2, q^2 l$ . The final number of equations depends on the order of anisotropy needed to be achieved: As an example, Cheng et al. (2005), simulating nonlocal convection properties, stop the chain of the equations at the fourth-order moments, finally solving the system of 36 equations. Most geophysical applications, however, employ no more than the 2.5 level closure models (according to the Mellor & Yamada, 1974, classification). These use two prognostic equations to find the turbulent length scale and

the TKE and then use simplified algebraic equations for second-order moments to describe the anisotropy of the flow and determine the stability functions. Different closures at level 2.5 use from 7 (e.g., Galperin et al., 1988) to 10 simplified algebraic equations (e.g., Canuto et al., 2001). Many of these models have been combined in the General Ocean Turbulence Model (Burchard & Bolding, 2001; Burchard et al., 2008; Umlauf & Burchard, 2005) and are implemented in NEMO as a GLS, which allows employment of prognostic equations for any combination ( $q^2, q^n, l^m$ ) models.

The GLS module in NEMO has a variety of “closures”—that is, different ways to define the stability functions in terms of the key turbulence-scale factors (3): dimensionless shear,  $G_m$ , and stratification,  $G_h$ .  $G_\Omega$  is not invoked as Coriolis effects are neglected in these closures. All these closures are “local,” that is, based on a set of algebraic equations for turbulent buoyancy fluxes and Reynolds stresses in which time derivative and turbulent transports are neglected. The resulting number of stability functions is reduced to two: one for momentum (effecting turbulent viscosity) and one for heat and salt (turbulent vertical diffusivity) such that

$$A_m = A_u = A_v, \quad A_h = A_S = A_T \quad (9)$$

## 2.2. Turbulent Closures of 2.5 Level in GLS Module

In this study we explore the following closures, present in NEMO:

**GKHR.** The (Galperin et al., 1988) closure is very similar to Mellor and Yamada (1982). That work considered equilibrium regimes of turbulence, where production of TKE balances dissipation rate. They found upper and lower limits for  $G_h$  for various flow regimes: In the convective regime  $G_h < 0$  and  $G_h > -0.029$ , with the upper limit being a corollary to the necessity for length scales to remain finite. In stably stratified flow  $G_h > 0$ , and a constraint comes from the Ozmidov length-scale limit (the size beyond which eddies are unable to overturn; see section 2.3 below).

**KC94.** The Kantha and Clayson (1994) closure is an extension of the GKHR and Mellor and Yamada (1982) models. Following Andren and Moeng (1993), they added a shear term and a term proportional to the scalar variance of the pressure-temperature correlation term. The critical flux Richardson number in the equilibrium case, if a limiter is not applied, becomes  $Rf_c = 0.213$  and critical gradient Richardson number  $Ri_c = 0.25$ , in accordance with the linear stability criteria (6).

**K3.** The (Kantha, 2003) closure is the same as KC94, but with adjusted constants that allow for the existence of turbulence at  $Ri > 0.25$ . The critical Richardson number in this closure in a steady state regime, if limiters are not applied, correspond well to the CHCD closure, discussed below.

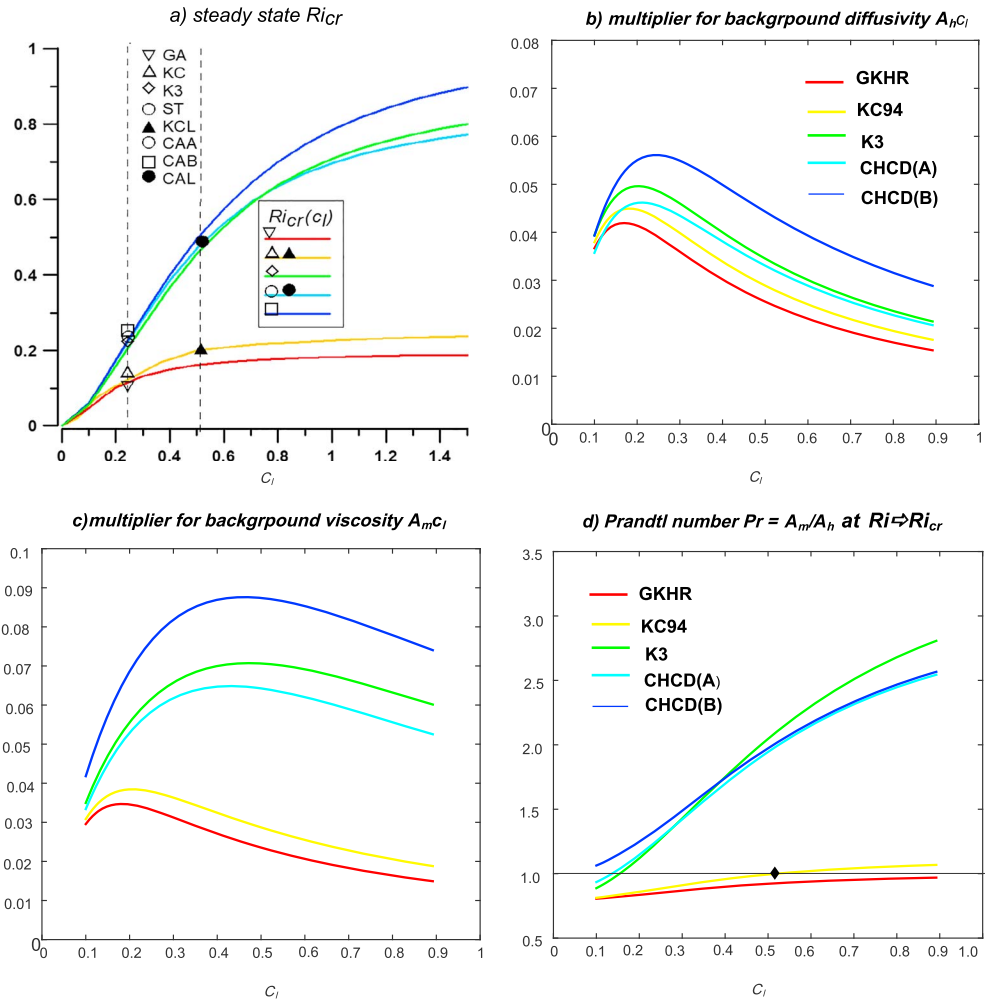
**CHCD.** The Canuto et al. (2001) and Cheng et al. (2002) closures add anisotropic shear production ( $\Sigma^4$  terms), buoyancy, salinity production, and the vorticity contribution to pressure-velocity, pressure-density, and pressure-salinity correlations. This model produces turbulent regimes for Richardson numbers  $Ri \sim 1$  (as demonstrated with nonlinear analysis by Abarbanel et al., 1984). The critical Richardson numbers are  $Ri_{cr} = 0.85, Ri_{cr} = 1.03$  for different sets of constants A and B (see Appendix A and Figure 1a), respectively.

## 2.3. Properties of the Turbulence Mixing Schemes as Applied to Regional Simulations: Effects of Deardorff/Galperin Limiter

All closures considered in this study allow the existence of turbulence at gradient Richardson number,  $Ri > 1$ , in unsteady, decaying stages of turbulence or where the vertical diffusion of TKE is important in the balance. Lateral transport of turbulent properties has been neglected throughout this work on the basis that the tidal excursion is at best marginally resolved. In the equilibrium state, where production due to shear balances the dissipation rate and work done opposing stable stratification, the gradient Richardson number depends on the type of stability function and the imposed Deardorff/Galperin limiter (Kantha, 2003, Cheng et al., 2002, 2003), which comes from the Ozmidov length-scale limit:  $L_o = c_l(\epsilon N^{-3})^{1/2}$ ; the size beyond which eddies are able to overturn, so

$$Nlq^{-1} \leq c_l, \quad (10a)$$

The range of  $c_l$  according different observational estimates is (0.2–0.6) (Deardorff, 1973; Galperin et al., 1988; Lewellen & Teske, 1973). It defines the maximal threshold for stability factor  $G_h$  in the wide range of



**Figure 1.** (a) Dependence of steady state critical gradient Richardson number  $Ri_{cr}$  on closure and Galperin/Deardorff limiter. (b, c) Multipliers for background diffusivity  $A_h c_l$  and viscosity  $A_m c_l$  in (10a) for different closures, (d) dependence of background Prandtl number on limiter  $c_l$  in conditions of  $Ri \geq Ri_{cr}$  and steady state. Blue = CHCD (B), cyan = CHCD (A), green K3, yellow-KC94 and red GKHR cases.

$$G_{h \leq} (c_l)^2 \approx [0.039 - 0.36], \quad (10b)$$

and the critical Richardson number (Appendix A) in steady state becomes dependant on  $c_l$  and type of closure (Figure 1a).

When the Richardson number exceeds the steady state limit, turbulence decays and one should be left with background diffusivity produced by unresolved subgrid processes such as internal waves and near-inertial oscillations (NIOs). In the NEMO GLS code the background TKE  $q_{bgr}^2$  is prescribed. The turbulent length scale is then set to the upper limit of (10a), that is,  $l = c_l q N^{-1}$ , and the background diffusivity becomes inversely proportional to the buoyancy frequency (Holt & Umlauf, 2008), which is consistent with the Gettett and Holloway (1984) parameterization of internal waves:

$$k_T = \frac{c_l A_h q_{bgr}}{N}, \quad k_m = \frac{c_l A_m q_{bgr}}{N}, \quad (11a)$$

and the dissipation rate of TKE is proportional to  $N$  and given by

$$\varepsilon = 0.5\check{c}_0 \check{q}_{bgr}^2 N / c_l \quad (11b)$$

where  $\check{c}_0 = 0.554^3$  for GKHR/KC94/K3 and  $\check{c}_0 = 0.527^3$  (CHCD).

The multipliers in (11a),  $c_l A_h, c_l A_m$  depend on the type of closure and limiter (Figures 1b and 1c), and thus the “background” diffusivity is closure dependent. Multipliers grow with closure in the order GKHR → KC94 → KC3 → CHCD(A) → CHCD(B). The Prandtl number  $Pr$  at  $Ri \rightarrow Ri_{cr}$  (Figure 1d) grows with limiter  $c_l$  for all closures and  $Pr > 1$  for Kantha (2003) and CHCD closures (see CHCD Figure 3) at the observed range of  $c_l \in [0.2, 0.6]$ . GKHR predicts  $Pr < 1$  in the full range of  $c_l$  and KC94 at ( $c_l < 0.527$ ) (see KC94 Figure 1) assuming more effective exchange of heat compared with exchange of momentum at the conditions of strong stratification. The latter is in agreement with a large-eddy simulation (LES) of a stably stratified atmospheric boundary layer (Zhou & Chow, 2012) and data from the Surface Heat Budget of the Arctic Ocean field program in the Arctic (Sorbjan & Grachev, 2010).

Cheng et al. (2003) pointed out that length-scale limiters are required in regimes of weak stratification ( $Ri < 0.25$ ), rather than at high  $Ri$ . This is supported by LESs and planetary boundary layer observations. Indeed, Holt and Umlauf (2008) demonstrated that in tidally active seas, when turbulence is not in equilibrium, a Deardorff/Galperin limiter is required in conditions of weak stratification in order to prevent a strong increase of turbulent length scale near regions of shear-driven convection. They show that in the absence of this limiter, the CHCD (A) closure resulted in a threefold reduction of stratification in the northwest European shelf seas, in conflict with observation.

#### 2.4. Prognostic Equations

Following Carniel et al. (2009), we use the  $(k - \varepsilon)$  model with constants consistent with each closure applied (see Umlauf & Burchard, 2005). TKE  $k = 0.5q^2$  and dissipation rate  $\varepsilon$  are highly variable both in the upper boundary layer due to diurnal variability and in the bottom boundary layer due to tidal friction effects (Simpson et al., 1996). This results in strong temporal variability nonequilibrium regimes of turbulence in both the boundary layers and the pycnocline. The existence of turbulence at Richardson numbers far above steady state critical values is found in observations (Palmer, Polton, et al., 2013) and models, even in the absence of (or unresolved) internal waves in the pycnocline. Time-dependent turbulence-scale factors affect the turbulent diffusivity and viscosity and feedback on the temporal evolution of  $k$  and  $\varepsilon$ .

Further, for simplicity we will call the combination of prognostic equations, stability function (closure), length-scale limiter application, and background mixing parameterization as the turbulence mixing scheme (TMS).

### 3. Methods

#### 3.1. Numerical Model and Experiments

We use the northwest European shelf NEMO configuration with ~7-km resolution (AMM7; Figure 1) and 51 vertical hybrid terrain-following/geopotential levels, compressed toward the bottom and surface. This is a shelf sea NEMO V3.6 (Madec & NEMO Team, 2016), currently used for operational forecasts/hindcasts by the U.K. Met Office (and known as CO6). Most of the model details are similar to O’Dea et al. (2017), which is the V3.4 version, and the validation of that version versus observations and a finer-resolution model can be found in Graham et al. (2018). The model explicitly resolves tides, driven by geopotential forcing, including 13 tidal constituents, present in NEMO code and detected in observations (Davies & Kwong, 2002). Tidal lateral boundary conditions are defined by the Flather (1976) algorithm and use sea surface elevations and velocity for nine harmonics from a tidal model of the North Atlantic (Flather, 1981). Temperature and salinity lateral boundary conditions are prescribed by the Global Seasonal Forecast system version 5 (MacLachlan et al., 2015) using the nominally 1/4 degree global ORCA NEMO model. For surface boundary conditions, we use the CORE formulation (Large & Yeager, 2009) with ERA-interim atmospheric reanalysis variables (Dee et al., 2011). Other details, and the sensitivity to forcing and physics options, can be found in O’Dea et al. (2017). The present model does not use data assimilation.

The standard run of V3.6 corresponding to ST conditions in this study, was initialized in January 1980 by interpolating temperature and salinity fields from the 1/4° ORCA025 hindcast of the standard global ocean

**Table 1**  
Numerical Experiments

Run name	Duration years	Closure and/ $c_l$ / (formula applied)	$Ri_{cL}(c_l)$	ERSEM	Surface boundary
GA	1996–2010	GKHR (11a,b):0.267	0.124	No	$\alpha = 1.3$
KC	1996–2010	KC94 (11a,b):0.267	0.143	No	$\alpha = 1.3$
K3	1996–2010	Kantha,2003 (11a,b):0.267	0.231	No	$\alpha = 1.3$
CAA	1996–2010	CHCD(A) (11a,b):0.267	0.25	No	$\alpha = 1.3$
CAB	1996–2010	CHCD(B) (11a,b):0.267	0.254	No	$\alpha = 1.3$
ST	1980–2015	CHCD(A) (11a,b):0.267	0.25	No	$\alpha_{ch} = 100,000 (\alpha \cong 1)$
KCL	1998–2010	KC (12c):0.53	0.193	No	$\alpha = 1.3$
CAL	1998–2010	CHCD(A) (12c):0.53	0.429	No	$\alpha = 1.3$
CAAE	2014–2015	CHCD(B) (11a,b): 0.267	0.25	Yes	$\alpha_{ch} = 100,000 (\alpha \cong 0.5)$
CABE	2014–2015	CHCD(B) (11a,b):0.267	0.254	Yes	$\alpha = 1.3$
GAE	2014–2015	GKHR (11a,b):0.267	0.124	Yes	$\alpha = 1.3$

Note. The third column describes the closure and limiting condition (see Appendices A and B);  $Ri_{cL}(c_l)$  is the actual steady state Richardson number for the value of limiter applied (see Appendix A and Figure 1). The sixth column shows the constant used for the roughness parameter.

configuration GO5.0 (Megann et al., 2014). Most other simulations considered here start from the ST run in January 1996.

This study considers a series of perturbation experiments, which are defined in Table 1. The difference between the runs arises only from the choice of stability functions, the Galperin limiter ( $c_l$ ), and the way of prescribing surface boundary conditions for TKE and dissipation rate. Five runs have been performed for 1996–2010 (GA, KC, K3, CAA, and CAB).

Simulation results have been shown to be very sensitive to the background TKE (Costa et al., 2017). We use

$$q_{bgr}^2 = 2 \times 10^{-6} \text{m}^2 \text{s}^{-2}, \quad (12a)$$

which is consistent with the minimal observed level of turbulence at the conditions of strong stratification (Monin & Ozmidov, 1985).

In conditions of very strong stratification, the limiting condition for diffusivity/viscosity is imposed

$$k_{T,m} = \max(k_{T,m}, 10^{-6} \text{m}^2 \text{s}^{-1}) \quad (12b)$$

The parameters of the “standard” run can be found in the supporting information (namelist\_cfg) and Appendix B. Condition “ln\_clim = .true.” activates background diffusivity option (10a). We performed two additional runs (CAL and KCL), where condition (10a) is not activated. The Galperin-Deandroff limiter is applied only to  $G_h = \min(G_h, c_l^2)$ , affecting the stability functions, but not turbulence length scale itself. In this case, an additional condition for background dissipation rate is required to define background diffusivity:

$$\varepsilon = \max(\varepsilon, \varepsilon_{bgr}) \quad k_T = \frac{\check{c}_{0l} \check{A}_T (q_{bgr})^4}{4\varepsilon}, \quad k_m = \frac{\check{c}_{0l} \check{A}_m (q_{bgr})^4}{4\varepsilon}, \quad (12c)$$

the NEMO default value of  $c_l = 0.52$  (GKHR) have been employed here. Stability functions  $A_T, A_m$  are in the range of (0.04–0.08) and (0.04–0.16), respectively. From (12c) it is evident that setting too small a background dissipation rate (which is a default of  $10^{-12} \text{m}^2/\text{s}^3$  in NEMO, in units normalized by water density  $10^3 \text{kg}/\text{m}^3$ ) will result in extremely high background viscosity/diffusivity exceeding  $10^{-3} \text{m}^2/\text{s}$  relevant to the active mixed layer (ML). We use  $\varepsilon_{bgr} = 10^{-9} \text{m}^2/\text{s}^3$  instead, which is appropriate to the observed background values in this region (Rippeth et al., 2014; Simpson et al., 1996; Simpson & Tinker, 2009). In that case resulting diffusivity/viscosity vary in the range of ( $6 \times 10^{-6}$  to  $1.3 \times 10^{-5}$ )  $\text{m}^2/\text{s}$  and ( $6 \times 10^{-6}$  to  $2.6 \times 10^{-5}$ )  $\text{m}^2/\text{s}$ , respectively.

There are two options to set the sea surface roughness parameter,  $z_0$ , in NEMO. In the first case,  $z_0$  is set as some portion of wave height,  $h_s$ :  $z_0 = \alpha h_s$ , and  $h_s$  is defined by empirical parameterization following

Raschle et al. (2008), where  $\alpha$  is a constant, estimated to be in the range 0.5–1.3. In the second case, a Charnock type parameterization (Charnok, 1955) is used:  $z_0 = \alpha_{ch} \tau_w / (\rho_o g)$ , where  $\tau_w, \rho_o$  are wind shear stress and water reference density,  $\alpha_{ch}$  is a constant in the range of  $(0.1–1.6) \times 10^5$  (Carniel et al., 2009). In equilibrium wind-wave conditions, both formulations are physically equivalent, and  $\alpha = 4.16 \alpha_{ch} \left(\frac{\rho_a}{\rho_w}\right) C_d^{-1} \approx 10^{-5} \alpha_{ch}$ .

Runs CAA and ST differ only by parameterizations of the surface roughness: ST has a smaller surface roughness (see Table 1) and hence less mechanical mixing due to wave breaking. Run KCL has a smaller critical steady state Richardson number, compared with K3 and CAA, but a weaker limit on the turbulent length scale, which potentially affects the upper limit of vertical diffusivity in the nonequilibrium regime. Note that runs CAL and KCL are restarted from conditions on 1 January 1998 from the CAA and KC runs, respectively, and simulated from 1998 to 2010.

ERSEM, used to examine the sensitivity of ecosystem variables to different turbulence schemes, is a lower trophic level biogeochemical model of phytoplankton, zooplankton, detritus, and bacteria, which explicitly resolves carbon, nitrogen, oxygen, phosphorous, and silicon cycles (Blackford et al., 2004; Butenschön et al., 2016). Online coupling to NEMO provides ERSEM with 3-D temperature and salinity, velocity, and mixing distributions (and hence water column structure) allowing advection and diffusion of the ecosystem variables. The ecosystem model is as much as 10 times more computationally expensive compared to the only-physics runs. In shelf seas, the annual biogeochemical cycle is characterized by the summer exhaustion of nutrients and regeneration after winter mixing. For these reasons, AMM7-ERSEM simulations are shorter: three runs (CAAE, CABE, and GAE) for 2014–2015 are conducted, restarted on 1 January 2014 and initialized with data from a multidecadal standard AMM7-ERSEM run (ST, using the CHCD stability function), beginning in January 1980 (Butenschön et al., 2016).

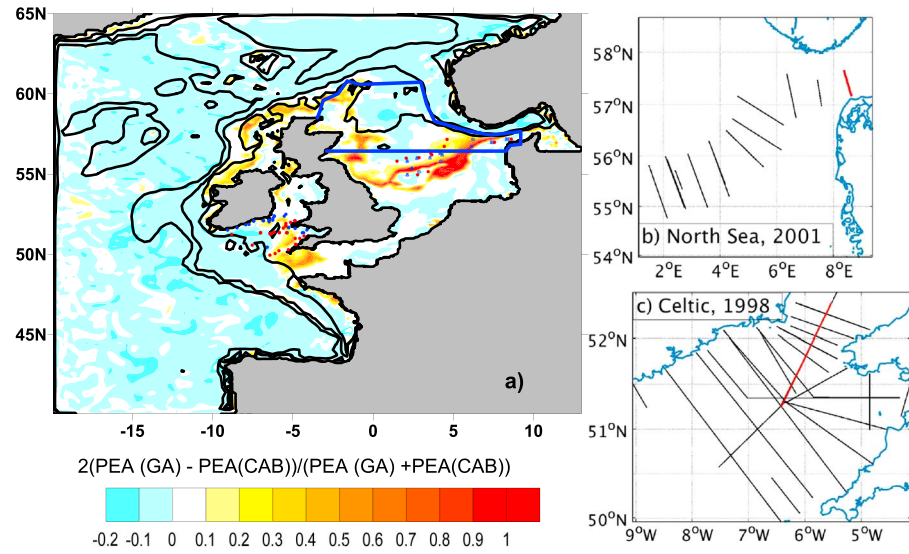
### 3.2. Observational Data

1. *Long-term variability.* For the assessment of the turbulence closures on decadal time scale we use the Met Office Hadley Centre “EN.4.2.0” data set of quality controlled ocean temperature and salinity profiles (Good et al., 2013). After the model spin-up we select all temperature and salinity profiles in the area shallower than 200 m during the stratified period (March to October 1998–2010), with  $N_{\max}^2 > 2 \times 10^{-5} \text{ s}^{-2}$ . We filter out profiles with  $N_{\max}^2 > 0.5 \text{ s}^{-2}$ , associated with river plumes, the exact location of which is not well resolved by this relatively coarse resolution model. Daily mean model variables are interpolated as a nearest neighbor in horizontal and linearly in vertical onto the coordinates of the observations with the resulting data set containing 22,000 profiles. Note that the model has a relatively fine vertical resolution on the shelf, especially near the surface and bottom, where the hybrid coordinates are compressed.
2. *Evaluation of synoptic scale variability.* High-resolution towed-undulating CTD (SCANFISH) hydrography data are used to compare model skill at fine spatial scales. We use data collected during July–September 1998 in 22 transects (~3,000 profiles) in the Celtic Sea (Brown et al., 2003) and two surveys in June–July 2001 in the North Sea (Cruise R/V *Corystes*, Centre for Environment, Fisheries and Aquaculture Science; Badin et al., 2009), with 17 transects (1,400 profiles in total). Both papers give detailed descriptions of the regions and the dominant physical processes there. The Celtic Sea survey region is a very energetic region with internal waves and mesoscale eddies, while the North Sea region is shallower, with weaker internal tides and baroclinic jets dominating the summer currents. Given the spacing of the sawtooth undulation (0.3–1 km) of the SCANFISH is much finer than the model resolution considered here (~7 km), we consider each down and up pass as a vertical profile at the location of its midpoint. Locations of the SCANFISH transects are shown in Figures 2a–2c. The times of each SCANFISH transect are restricted to dates only, without details of timing of individual profiles. For this reason we used model daily mean temperature and salinity profiles for comparison. This filters tidal components and diurnal variability, resulting in weaker variability in model variables compared with unfiltered observations.

### 3.3. Methods of Intercomparison

The evaluation of model skill is not a trivial task and depends on the metrics, databases chosen, and length of observations (Allen et al., 2007; Gleckler et al., 2008). We evaluate model skill in simulating pycnocline characteristics and bulk properties related to vertical stratification. Previous modeling studies (Graham





**Figure 2.** (a) Model domain and locations of SCANFISH surveys. Contours show 100-, 300-, and 1,000-m isobaths. Origins and final locations of the SCANFISH data are shown by blue and red dots. Colors show the relative difference in potential energy anomaly in July 2001 between the most and least diffusive runs, discussed in section 3. The northern part of the North Sea, discussed in Figure 3, is shown by blue contour; (b, c) zoomed SCANFISH transects in the North and Celtic Seas. Sections in red indicate the locations of transects, shown in Figure 5.

et al., 2018; O’Dea et al., 2017) demonstrate persistent strong warm biases of benthic temperatures in this region, comparing with observations. In our analysis, we examine how local properties of the pycnocline are related to simulation errors in bulk characteristics. All stratification characteristics are calculated on the same vertical grid and using the same procedures both in observations and model outputs.

The total stratification of the water column on the shelf can be evaluated using the potential energy anomaly (*PEA*):

$$PEA = \int_{-H}^0 g(\rho - \rho_m)z dz, \quad (13)$$

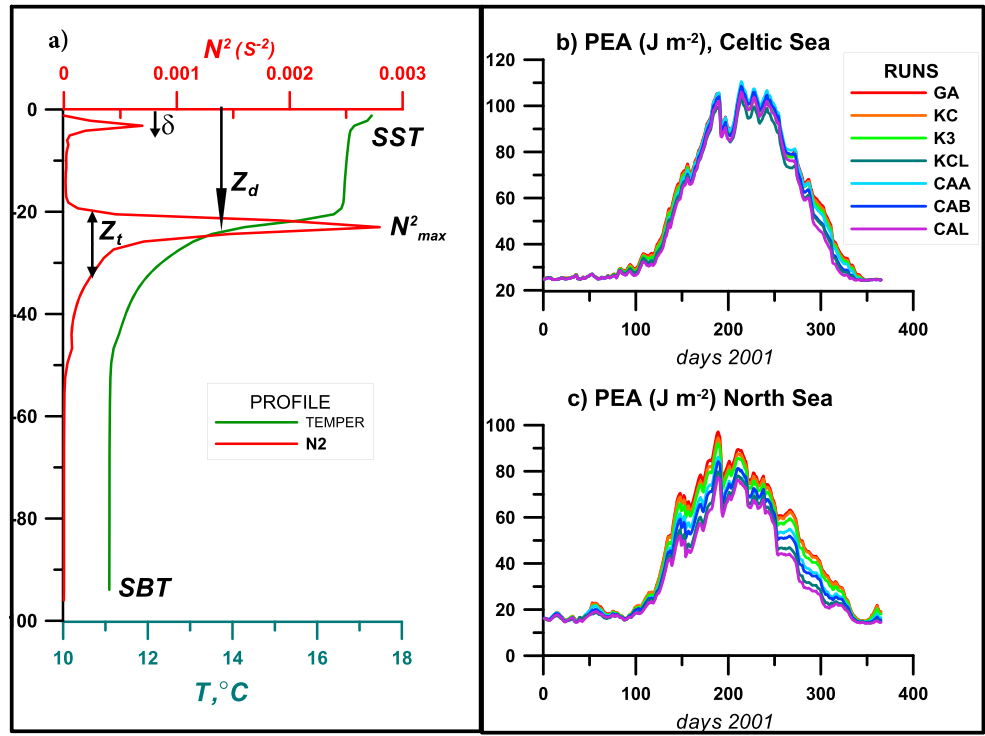
where  $\rho_m = H^{-1} \int_{-H}^0 \rho dz$  is depth-mean density of the fluid,  $g$  is gravity,  $H = \min(200 \text{ m}, h)$ ,  $h$  is the depth of the sea. We use this characteristic to classify models from less diffusive (high *PEA*) to more diffusive (lower *PEA*).

For model evaluation we also use the MLD, which is defined by a density criteria, at the depth where the difference with density at 10 m does not exceed 0.01 sigma units.

In seasonal pycnoclines in shelf seas, the structure of stratification (temperature and salinity profiles) is approximately two layered, consisting of upper and bottom weakly stratified layers, separated by a strong pycnocline (Figure 3a). To evaluate the model skill in reproducing stratification characteristics compared with hydrographic data, we schematically describe each profile by the following parameters: (i) SST, (ii) near-bed temperature (SBT), and local pycnocline characteristics: (iii) maximum of Brunt-Väisälä frequency in the pycnocline  $N_{\max}^2$ , (iv) depth of pycnocline ( $Z_d$ ), and (v) thickness of pycnocline ( $Z_T$ ), given by

$$Z_d = \frac{\int_{-H}^{-\delta} N^2 z dz}{\int_{-H}^{-\delta} N^2 dz}, \quad Z_T^2 = \frac{\int_{-H}^{-\delta} N^2 (z - Z_d)^2 dz}{\int_{-H}^{-\delta} N^2 dz}, \quad N^2 = -g\rho_0^{-1} \partial_z \rho. \quad (14)$$

Here  $\delta$  is a masking depth, corresponding to the depth of the shallow diurnal pycnocline (Figure 3a). Typically we take  $\delta = 5 \text{ m}$ , unless observations start from deeper levels. The depth of the pycnocline,  $Z_d$  is



**Figure 3.** Example of modelled temperature and  $N^2$  profiles in the Celtic Sea and five pycnocline characteristics of stratification (section 3.3), used in model assessment,  $\delta$  is a “masking” depth used in (14). (b, c) Evolution of shelf-wide mean PEA in 2001 for representative models from Table 1: (b) the Celtic Sea and (c) the North Sea. SST = sea surface temperature; SBT = near-bed temperature; PEA = potential energy anomaly.

a “center of mass” of Brunt-Väisälä frequency and can be expressed also in terms of mean and boundary density:

$$Z_d = \frac{\int_{-H}^{-\delta} N^2 z dz}{\int_{-H}^{-\delta} N^2 dz} = \frac{(\rho(H) - \rho_m) * H - \rho(\delta) \delta}{\rho(H) - \rho(\delta)}, \quad (14a)$$

As most of the domain is temperature stratified,  $Z_d$  is strongly related to the ability of the model to predict mean, surface, and bottom temperatures.

To examine the model’s skill in reproducing the variability of (i)–(v) (Figure 3a), we employ Taylor diagrams (Taylor, 2001), providing a graphical summary of how closely the model patterns match the observations. The similarity between two patterns is quantified in terms of their correlation:  $R$ , their centered root-mean-square difference:  $E'$ , and the amplitude of their variations, which is represented by their standard deviations (here  $\sigma_{T_m}$  and  $\sigma_{T_0}$  are modeled and observed standard deviations of variable  $T$ ). Being normalized by their standard deviations, all observed variables correspond to a single point (1,0) on the diagram, each model variable corresponds to the vector with a length-scale  $r = \sigma_{T_m}/\sigma_{T_0}$  and angle equal to  $\arccos(R_{(o,m)})$ , and  $m$  denotes model,  $R_{(o,m)}$  is the correlation between observations ( $o$ ) and modeled ( $m$ ) variables. The root-mean-square difference  $E'$  is equal to the distance between model and observed point (Taylor, 2001) and is given by

$$\frac{E'^2}{\sigma_{T_0}^2} = 1 + \sigma_{T_m}^2/\sigma_{T_0}^2 - 2R\sigma_{T_m}/\sigma_{T_0} \quad (15)$$

We introduce a “skill measure,” following ideas proposed by Holloway et al. (2011), to combine variability and bias errors into a single diagnostic variable:

$$i = \{1 : 5\} = \{SST, SBT, Z_d, Z_T, N_{max}^2\}$$

$$Ski^i = \left\{ (1-E') \left( 1 - \frac{|M^i - O^i|}{\max(M^i) + \max(O^i) - \min(M^i) - \min(O^i)} \right) \right\}^{1/2} \quad (16a)$$

$$Ski^i = \left\{ (1-E') \left( 1 - \frac{|M^i - O^i|}{|M^i| + |O^i|} \right) \right\}^{1/2} \quad (16b)$$

$$Ski = \left\{ \prod_{i=1}^5 Ski^i \right\}^{1/5} \quad (17)$$

where “ $O^i$ ” and “ $M^i$ ” denote variables (i–v),  $\|\cdot\|$  denotes spatial and temporary averaging of variables (i.e., over 22,000 data in the EN4 comparison, ~3,000 profiles in the Celtic Sea, and 1,400 in the North Sea SCANFISH data sets). For temperature (and salinity), where mean values are far above unity, we use (16a) (e.g.,  $M = T_{\text{model}}$  and  $O = T_{\text{observed}}$ ), so biases are normalized by maximum differences over all the observed period; for other pycnocline characteristics we use (16b). Each multiplier in (16) or (17) is less than unity. The square root in (16) and 5<sup>th</sup> root in (17) is taken to avoid the metric vanishing with increasing number of parameters under consideration. Perfect agreement between modeled and observed parameters occurs when every term in the product is unity. We subjectively consider model skills as high with  $Ski > 0.7$ , satisfactory if  $0.5 < Ski < 0.7$ , moderate low if  $0.4 < Ski < 0.5$ , low if  $Ski < 0.4$ , and absent of skill when  $Ski$  is imaginary, that is, when errors are larger than 100% for one of the variables (i)–(v).

## 4. Results

### 4.1. TMS Gradation

Area-mean differences in PEA, simulated with different TMSs, are not strong (Figure 3b) in the Celtic Sea, with the maximum difference reached in late summer–early autumn (about 4%). In the North Sea differences are much stronger (25%; see Figure 3c). While in Figures 3b and 3c 2001 has been chosen in the context of the SCANFISH survey in the North Sea, when high-resolution data are available, this behavior repeats from year to year. Locally, differences in total stratification on the shelf (taken here to be shallower than 200 m) reach 50–100% in the North Sea (Figure 2a), in the location close to the SCANFISH survey in 2001. Higher domain-mean PEA corresponds to stronger shelf-wide stratification. Using the domain-mean PEA criteria, the TMSs are sorted from the less diffusive (more stratified) to more diffusive (less stratified) as

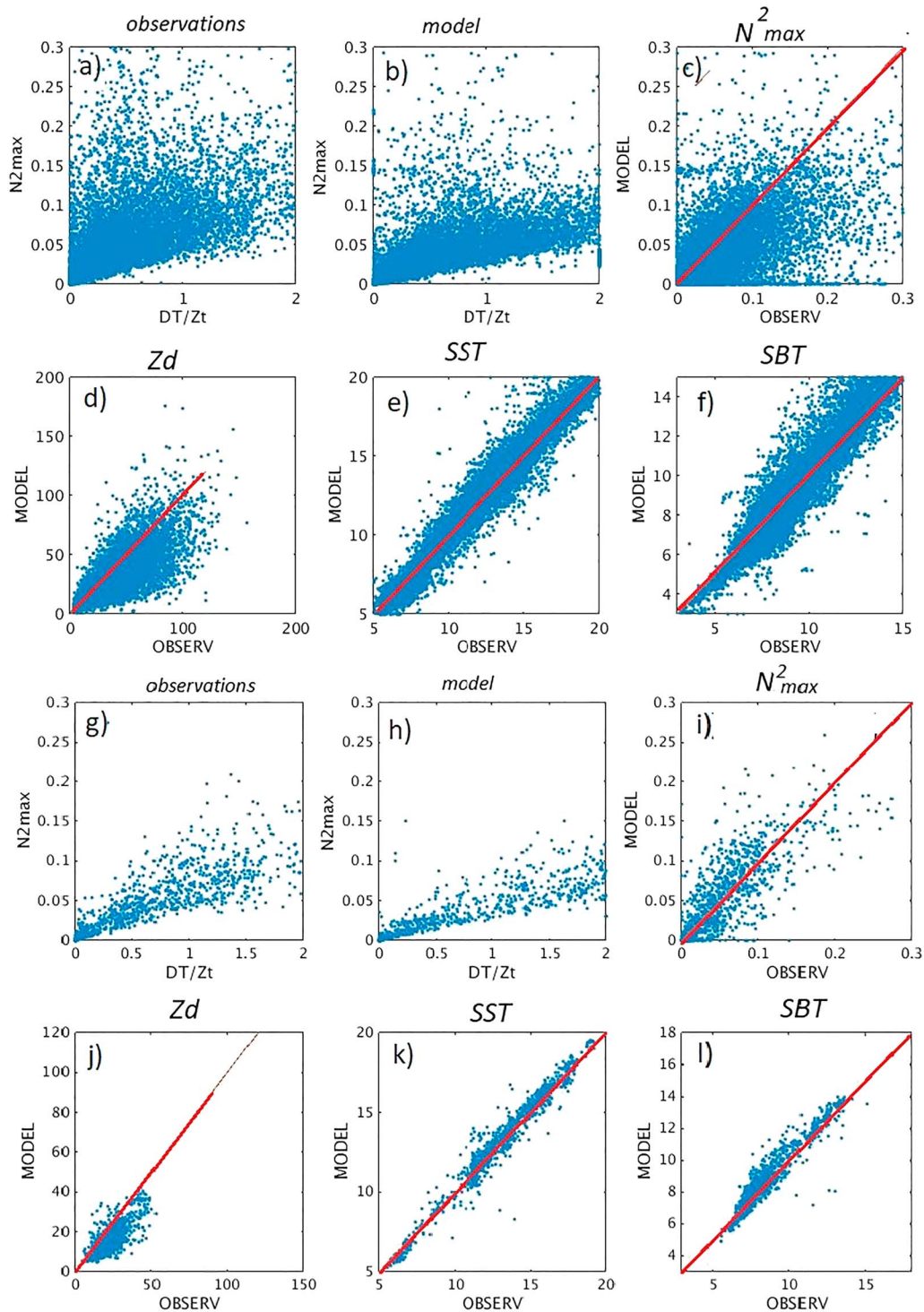
$$GA \rightarrow KC \rightarrow K3 \rightarrow \{KCL, ST, CAA\} \rightarrow CAB \rightarrow CAL \quad (18)$$

In general, this classification is in agreement with the critical Richardson number gradation, increasing from GA ( $Ri_{cL} = 0.12$ ) to CAL ( $Ri_{cL} = 0.49$ ); see Table 1. However, despite the higher critical steady state Richardson number in the K3 run ( $Ri_{cL} = 0.231$ ) compared with KCL run ( $Ri_{cL} = 0.193$ ), the latter TMS is more diffusive, pointing to the importance of the length-scale limiter. In the above, KCL, ST, and CAA are combined in (18) into a “moderately diffusive group” with relatively similar properties, despite the differences in critical Richardson numbers and surface boundary conditions.

### 4.2. Comparison of Pycnocline Characteristics on Decadal Time Scales

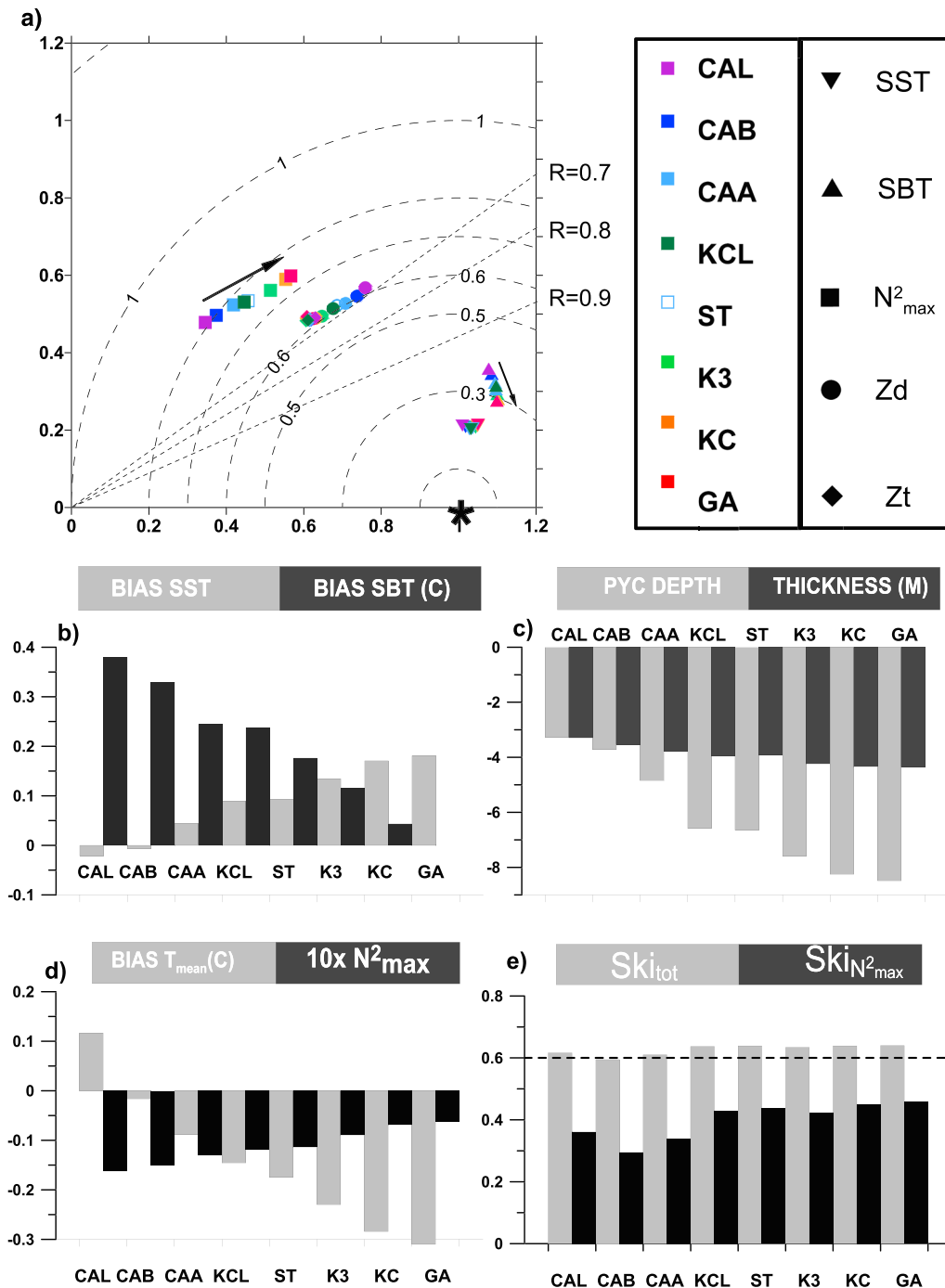
The SBT is affected by the history of horizontal and vertical heat fluxes on seasonal and interannual time scales (Holt et al., 2012), without being directly constrained by surface boundary conditions in the summer. Similarly, differences in temperature over the water column,  $DT = SST - SBT$ , are influenced by integral characteristics of shelf sea dynamics, that is, “leakage” of heat through the pycnocline and lateral transport on seasonal time scales. The depth and thickness of the seasonal pycnocline, as well as  $N_{max}^2$ , depend more on local pycnocline properties controlled by short-term variability at synoptic, semidiurnal, diurnal, and inertial time scales, for example, internal waves (e.g., Palmer et al., 2015) and small-scale eddies (Badin et al., 2009).

Both modeled and observed local pycnocline properties, such as  $N_{max}^2$ , demonstrate positive correlations with bulk  $DT/Z_T$  but exhibit strong scatter (Figures 4a and 4b). Different models show similar patterns; in Figure 4 we show an example of the “moderately” diffusive K3 run. Modeled and observed  $N_{max}^2$  are



**Figure 4.** Scatterplots of model variables (i)–(iv) versus EN4 observations for K3 closure during 1998–2012, (a–f) in whole shelf domain; (g–l) for the northern part of the North Sea, shown by blue contour in Figure 1. SST = sea surface temperature; SBT = near-bed temperature.

positively correlated, also with strong scatter (Figure 4c). This scatter suggests a contribution from processes other than vertical mixing, that is, advection or lateral mixing. These processes may be related to the presence of multiple small-scale tidal fronts within the domain (see Figure 5 from Graham et al., 2018). Indeed, when we consider the northern part of the North Sea in isolation, where small-scale tidal fronts



**Figure 5.** (a) Taylor diagram (after Taylor, 2001) summarizing the relative skill of model variables relative to observations in 1998–2012. Assessed variables (i)–(v) are denoted by varying symbol shapes. Assessed closure models are denoted by symbol colors. For each variable/model pairing the simulated standard deviation (normalized by the observed standard deviation) and the correlation between the model and observations are plotted as a radius and a phase (clockwise from the y axis) such that the star denotes a perfect simulation with a correlation  $R = 1$  and standard deviation of 1, and distance from the star (dashed lines) denotes the model error; Biases in (b) SST and SBT; (c) pycnocline depth  $Z_d$  and thickness  $Z_t$ ; (d) shelf mean temperature and  $N^2_{max}$  at pycnocline; (e) model skill evaluated by (12) for  $N^2_{max}$  and total skill. Models are ordered in increasing overall diffusivity, as characterized by the potential energy anomaly assessment (18). SST = sea surface temperature; SBT = near-bed temperature.

are less prevalent, stronger similarity between observed and modeled properties in the pycnocline becomes evident (Figures 4g–4i). Generally, SST and SBT are well predicted by all TMSs, with very high correlations (close to 0.99; Figure 5a) and positive biases in SBT (model too warm—indicative of a general oversupply of

heat from surface to depth). The model error decreases by approximately 10% for both bottom temperature and  $N^2_{\max}$ , from more diffusive to less diffusive TMSs.

In the majority of the simulations, modeled surface temperatures exhibit positive biases with respect to the observations, with stronger errors in SST (0.13–0.18 °C) in less diffusive GA-KC models and small negative biases (<0.02 °C) in CAL-CAB runs (Figure 4b). Positive biases in the benthic temperature increase with more diffusive models: from 0.03 °C in the GA case to 0.39 °C in the CAL case.

The choice of TMS is shown to have a larger impact on the model properties than the choice of surface boundary conditions: the ST and CAA runs differ only in the surface boundary conditions and result in changes in SBT biases of approximately 0.05 °C, while differences in choice of stability functions produce changes in biases of 0.1–0.4 °C (Figure 5b).

While in the majority of runs both surface and bottom temperatures exhibit positive biases, the mean temperature (over all observed profiles, showing changes in heat content) has a negative bias in all runs except CAL. Biases change (Figure 5d), from positive 0.11 °C (CAL) to increasingly negative toward less diffusive mixing schemes (−0.31 °C, GA), which is about a 12% difference in bias, normalized by the standard deviation of 3.22 °C. This is due to an underestimation of pycnocline depth in all runs: from 3.2 m in CAL to 8 m in GA. These differences arise due to the feedback between SST and surface heat fluxes, presumably the sensible heat flux, dominating during nighttime and autumn cooling. SSTs also affect the latent heat flux and thus evaporation, although this is a smaller effect than the sensible heat flux, as demonstrated by small differences in mean salinity biases: ~0.017 PSU (practical salinity units) between runs (from 0.07 PSU in CAB run to 0.087 PSU in GA), which is only 2.4% of the salinity standard deviation of 0.69 PSU. Mean temperatures and so water column heat content on the shelf are better simulated by the CAB and CAA models, while  $N^2_{\max}$  is better predicted by the low-diffusive TMS runs GA, KC, and K3 (Figure 5d).

Model skill, defined by (16), is very high for SST and SBT in all cases (above 0.8), satisfactory (in the range 0.5 to 0.6) for pycnocline depth and thickness and varies from low (from 0.3 to 0.33) to moderately low (0.43–0.46) for  $N^2_{\max}$  in the pycnocline (Figure 5e). Combined skills, evaluated using statistics of all five variables (17) are satisfactory (0.58–0.63) with a slight growth from CAB to GA models.

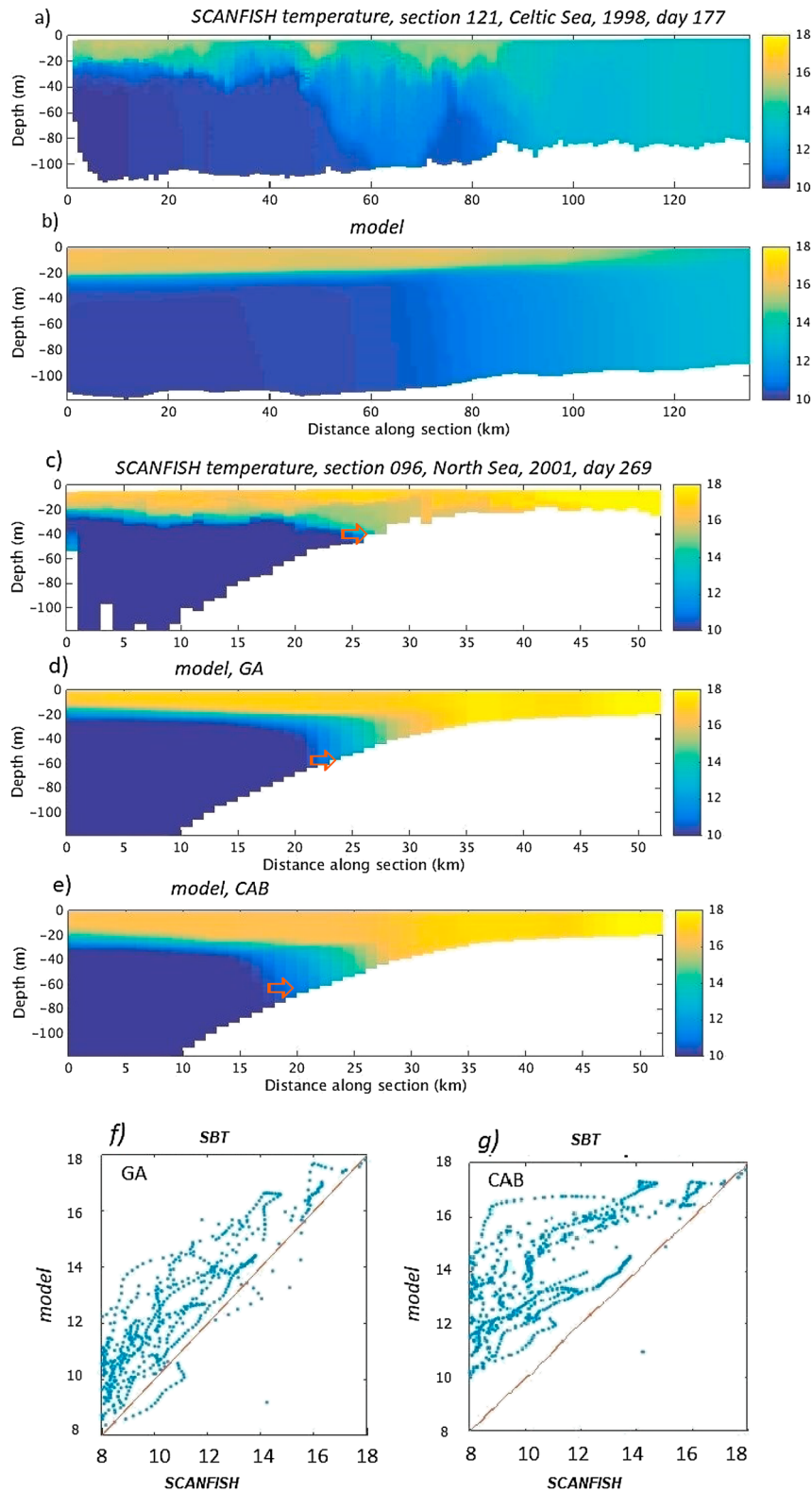
Surprisingly, KCL and ST runs have very similar characteristics, despite the difference in closures, critical Richardson numbers, surface boundary conditions, and limiters applied. This similarity persists in further analysis on the shorter time scale.

#### 4.3. Simulating Synoptic Variability: Comparison With SCANFISH Data

The SCANFISH transects, performed in the Celtic Sea and southern part of the North Sea, have much higher spatial resolution (~0.3–1 km) than the model, ~7 km. The Celtic Sea survey covers a region of the shelf that is approximately 100 m deep and reveals intense mesoscale and submesoscale activity. Typical eddy length scales are of the order 10 km (see Figure 6a; and Holt & James, 2006), and sharper, small-scale variations in the thermocline are observed that are indicative of internal waves. Neither of these baroclinic processes are resolved by the hydrodynamic model. As the exact time of individual profiles is not known, for comparison we used daily averaged model data. For these reasons, as expected, the simulated thermocline is a comparatively flat and smooth (Figure 6b).

The North Sea SCANFISH survey is conducted in significantly shallower water of around 40–70 m at the southern end of the transects and deeper (~100 m) at the northern end (Figures 6c–6e and 2b). Most of the transects cross a thermal front, which is most pronounced at the east part of domain. The models accurately reproduce the locations of lateral thermal fronts (Figures 6c–6e); however, these are much thicker and smoother in more diffusive TMSs (Figure 6e), inducing stronger warm biases of benthic temperatures (Figures 6f and 6g). Observed pycnocline features appear less dynamic, with generally less vertical variability along isopycnals compared with those in the Celtic Sea (Figure 6a).

The terrain following model coordinates (O'Dea et al., 2017) produce a vertical resolution of about 1–2 m over these shallow depths, which is close to the resolution of the SCANFISH survey (1 m, in the data format received). Correlations between simulated  $N^2_{\max}$  in the pycnocline and simulated bulk gradient of  $DT/Z_t$ , are very high for both regions,  $R(N^2_{\max}, DT/Z_t) = 0.85, 0.9$  in the North and Celtic Seas, respectively. This exceeds that found in the observations ( $R = 0.6$  in the North Sea and  $R = 0.85$  in the Celtic Sea), which



**Figure 6.** (a, b) Examples of SCANFISH section observations versus the model in the Celtic Sea during July–September 1998. Model data are daily averaged, and each datum is taken at the specific date of observation at each location. Temperature across section 121 in observations and in the GA model. (c–e) Examples of SCANFISH section observations versus model in the North Sea in June–August 2001. Temperature across section N96 in the North Sea for (c) observations, (d) GA, and (e) CAB runs; (f, g) scatterplots of benthic temperature at both runs in the North Sea (all transects). Red arrows in the figures (c, d) show the width of thermal fronts. SST = sea surface temperature; SBT = near-bed temperature.

are nonetheless still high. This shows the dominance of vertical heat transport in the simulated and observed stratification. Smaller correlations in the North Sea reflect the importance of larger-scale horizontal advection of heat and salt here. In both seas, the variance in the simulated  $N_{\max}^2$  is 2–3 times smaller than that observed, which is  $\sigma_T^2 = 3.6 \times 10^{-3} \text{ s}^{-4}$  for the Celtic Sea, based on 3,140 “profiles” and  $3.4 \times 10^{-3} \text{ s}^{-4}$ , based on 1,398 profiles in the North Sea. Modeled variances are  $\sigma_T^2 = 1.6 \times 10^{-3} \text{ s}^{-4}$  and  $1.1 \times 10^{-3} \text{ s}^{-4}$ , respectively. Differences in  $N_{\max}^2$  variance can partly be explained by filtering of diurnal and tidal variability by the 25-hr averaging of the model data. The variability of  $N_{\max}^2$  is approximately the same in the Celtic and North Seas but is determined by different processes: by eddies and internal waves in the Celtic Sea and by the presence of small-scale tidal fronts in the southern part of the North Sea (Figures 6a and 6c).

Considering the statistics of the differences between observations and each model (Figure 7), the SST variability and biases ( $-0.1:0.3 \text{ }^\circ\text{C}$ ) are the best simulated of all variables considered (i–v; see Figures 7a–7c). Given uncertainties of the heat flux from ERA-Interim (resolution about  $0.7^\circ \sim 60\text{--}80 \text{ km}$ ) and the strong variability of hydrophysical variables in shelf seas, we conclude that SSTs are very well predicted by all models. The most poorly modeled variables are pycnocline thickness and  $N_{\max}^2$  with errors of about 80% in the North Sea and 90–100% in the Celtic Sea (Figures 7a and 7b). This can be attributed to high-frequency variability in the data, which is smoothed or missing in the models' process representation, and also to inadequate horizontal resolution.

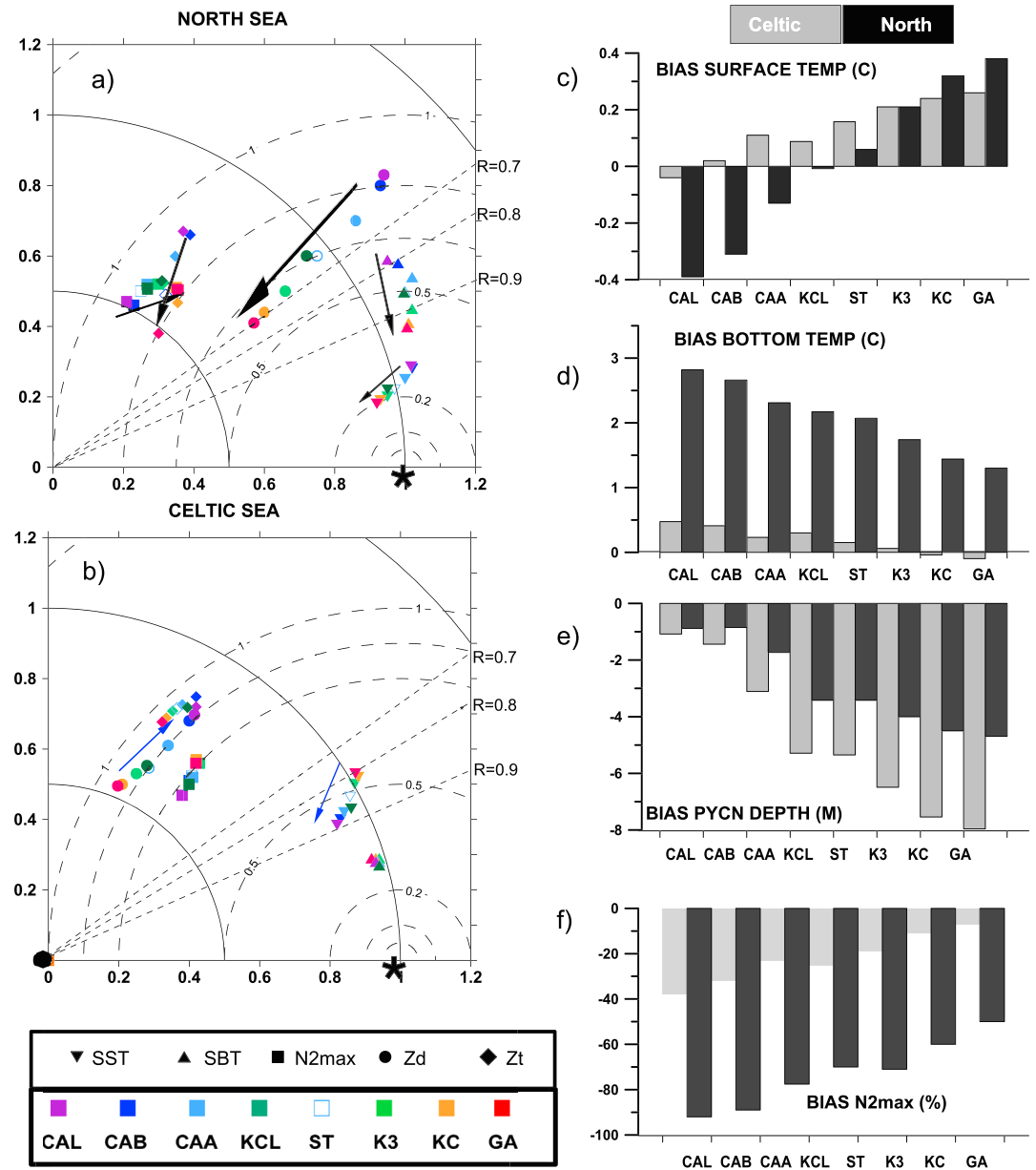
For both SCANFISH surveys, the models' skills in simulating variability of SST, SBT,  $Z_d$ ,  $Z_t$ , and  $N_{\max}^2$  correlate with the diffusivity gradation of the TMSs (18) (see Taylor diagrams [Figures 7a and 7b] and bias bar charts [Figures 7c–7f]). However, the direction of “improvement” of model skill is different for the two seas.

In the North Sea, the models' ability to reproduce variability improves from “more diffusive” (CAL, CAB) to “less diffusive” TMSs (KC, GA) for all variables, resulting in a decrease of errors by 10–25% (Figure 7a). Less diffusive TMS runs show nearly half the negative biases in  $N_{\max}^2$  (Figure 7f,  $-40\%$  to  $-50\%$  of bias for GA-KC) compared with ( $-80\%$  to  $-90\%$ ) for the CAL-CAB case. However, CAB-CAL show reduced bias in pycnocline depth in this location ( $-1$  to  $-2 \text{ m}$  for CAB-CAL versus  $-4$  to  $-5 \text{ m}$  for K3-GA). All models predict the variability of benthic temperature reasonably well, with high correlations (0.8–0.9) and variability errors (equation (15)) reducing from (0.6 to 0.4) from CAL to GA TMS. Benthic temperature is positively biased (1.4–2.9  $^\circ\text{C}$ ), with the worst SBT simulated by more “diffusive” TMS. Thus, excessive diffusiveness of CAL-CAA TMS runs results in a deeper pycnocline with weaker stratification, which allows stronger diffusive flux of heat from the surface to the bottom layer and overheats it. In shallow locations, corresponding to the areas of maximum PEA differences between the models (Figure 2), the surface ML is highly coupled with the bottom ML, resulting in a shift of tidal mixing fronts and mixing from surface to the bottom in CAB-CAL runs compared with GA (see Figures 5c–5e).

The situation is different in the Celtic Sea, where the simulation of variability of some characteristics slightly improves from less diffusive to more diffusive TMSs: for pycnocline depth by 5% and for SST by 10% (Figure 7b). Biases are small for all variables, except the pycnocline depth, for which biases are negative everywhere and better predicted by the CAA and CAB models. K3-GA models strongly underestimate  $Z_d$  with a deficit of 4–7 m. We hypothesize that sporadic mixing, produced by mesoscale and submesoscale events, shear instability, or breaking of internal tidal waves deepen and broaden the pycnocline in the Celtic Sea, and their effects are better represented by the more diffusive closures. In other words, use of more diffusive turbulence closures is probably compensating for the lack of unresolved processes or their parameterizations in moderate-resolution 3-D models.

Considering the skill scores for the comparison with SCANFISH data (Figure 8), all models have high skills in the prediction of surface and benthic temperatures, with slightly better properties of CAB-CAA models in the Celtic Sea and GA-KC in the North Sea (Figures 8a and 8b), consistent with Figure 7. Similar trends are valid for pycnocline characteristic (Figures 8c–8e), with much stronger preference for KC-GA in the North Sea. Despite stronger biases in the prediction of the pycnocline depth, GA-KC TMSs have greater skill than CAA-CAB for this particular characteristic. Integral model skill (Figure 8f) shows the same tendency: slightly better skills of CAL-CAB runs in the Celtic Sea and much better prediction by K3, KC, and GA TMSs in the North Sea.

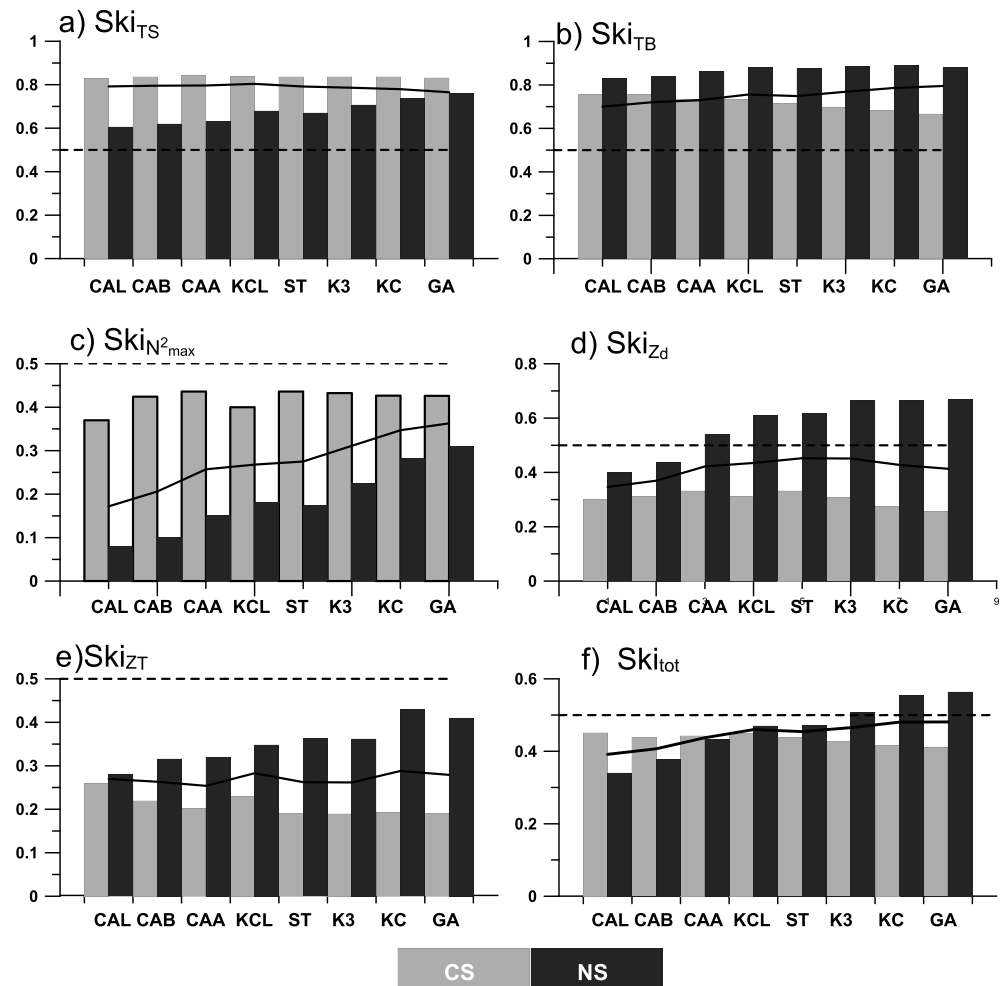




**Figure 7.** (a, b) Taylor diagrams for simulated pycnocline characteristics versus SCANFISH surveys (i–v). Arrows show improvement from model to model. (c–f) Biases for SST, SBT,  $Z_d$ , and  $N^2_{max}$ , with the latter being calculated as a percentage. Gray and black bars denote the Celtic Sea and North Sea, respectively.

In general, model skill to predict pycnocline characteristics on the synoptic scale is satisfactory ( $>0.5$ ) only for GA–KC models in the North Sea. Combining skill metrics in both regions as  $Ski^i = \{Ski^i_{NS} \cdot Ski^i_{CS}\}^{1/2}$  for each variable and for total skills, the preference for low diffusive (GA and KC) TMSs for SBT and  $N^2_{max}$  becomes evident. Equilibrium TMSs (ST, K3, and KCL) win in the prediction of pycnocline depth only. Resulting skills (Figure 8e) grow from 0.39 to 0.48 from high to low diffusive TMS.

KCL and ST models show very similar results despite the differences in the formulation of the closure, the critical Richardson numbers in steady state, and the limiting length-scale parameter. Increasing length-scale limiter and type of limiting (from KC to KCL and from CAA to CAL) results in a reduction of skills in some parameters (bottom temperature and  $N^2_{max}$ ) or nearly neutral in others. Similarly, an increase in surface roughness and of the effects of breaking waves (from ST run to CAA) does not improve model skill.

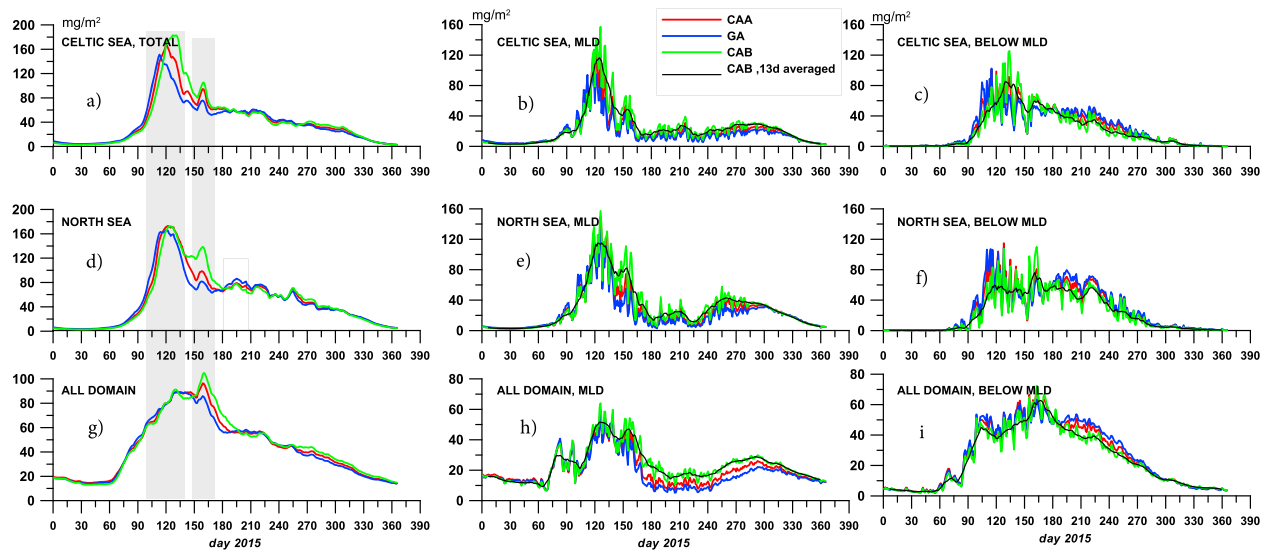


**Figure 8.** Model skill for assessment for (a) SST, (b) SBT, (c)  $N^2_{max}$ , (d)  $Z_d$ , (e)  $Z_T$ , and (f) total skill, based on these five variables. Gray = in the Celtic Sea; black = in the North Sea; contour solid line shows combined skills for each SCANFISH survey, dashed line denotes skill threshold of 0.5. SST = sea surface temperature; SBT = near-bed temperature.

## 5. Discussion

### 5.1. Comparison With Other Sensitivity Studies in the Northwest European Shelf

Here we demonstrate that the differences induced by changing the stability functions, as discussed in this work, produce model biases of the same order of magnitude ( $\sim 0.2\text{--}0.4\text{ }^\circ\text{C}$ ) as differences arising from changes in heat flux or from improvements in horizontal resolution, as reported in other studies in this region. For example, O’Dea et al. (2017) used different light attenuation schemes in 30-year simulations of AMM7 and found summer month differences in SST and bottom temperature to be  $0.2\text{--}0.3\text{ }^\circ\text{C}$  in the seasonally stratified part of the shelf (their Figure 15a). Similarly, Graham et al. (2018) compared 20-year runs of AMM15 (1.5-km resolution) and an AMM7 run, which corresponds to the ST run in this study. They report that the summer months (June–August, JJA) bottom temperature biases (model minus EN4) varied from  $0.47 \pm 1.19\text{ }^\circ\text{C}$  to  $0.50 \pm 1.31\text{ }^\circ\text{C}$  for the different configurations. An SST comparison with EN4 gives the following biases and standard deviations:  $0.18 \pm 0.97\text{ }^\circ\text{C}$  and  $0.12 \pm 0.99\text{ }^\circ\text{C}$  for AMM15 and AMM7 correspondingly. These values are not identically comparable with our study, as the simulations length differed (20 and 14 years correspondingly) and analyses were performed for different periods of time (JJA versus March–October). However, the model appears to be more sensitive to differences in TMSs (Figure 5b) than to resolution: a noteworthy result. SST biases increase from  $0.07 \pm 0.74\text{ }^\circ\text{C}$  in CAB run to  $0.18 \pm 0.78\text{ }^\circ\text{C}$  in GA run, while SBT biases decrease from  $0.33 \pm 1.01\text{ }^\circ\text{C}$  in CAB to  $0.01 \pm 0.84\text{ }^\circ\text{C}$  in GA run. ST run has  $0.09 \pm 0.74\text{ }^\circ\text{C}$  in SST and  $0.18 \pm 0.92\text{ }^\circ\text{C}$  in SBT. The smaller biases in SBT compared with the Graham



**Figure 9.** Depth-integrated and area mean chlorophyll in 2015. (a–c) the Celtic Sea integrated over water column, surface mixed layer only, and below mixed layer depth; (d–f) as in (a)–(c) but for the North Sea; (g–i) for basin mean. Gray color marks periods of time averaging of chlorophyll fields, shown in Figure 10. Blue = GA; red = ST; green = CAB runs. Black curves in central and right panels are 13-day running means for the CAB run.

et al. (2018) paper is mostly due to different averaging over seasons, as SBT differences are much smaller during March–May (not shown here).

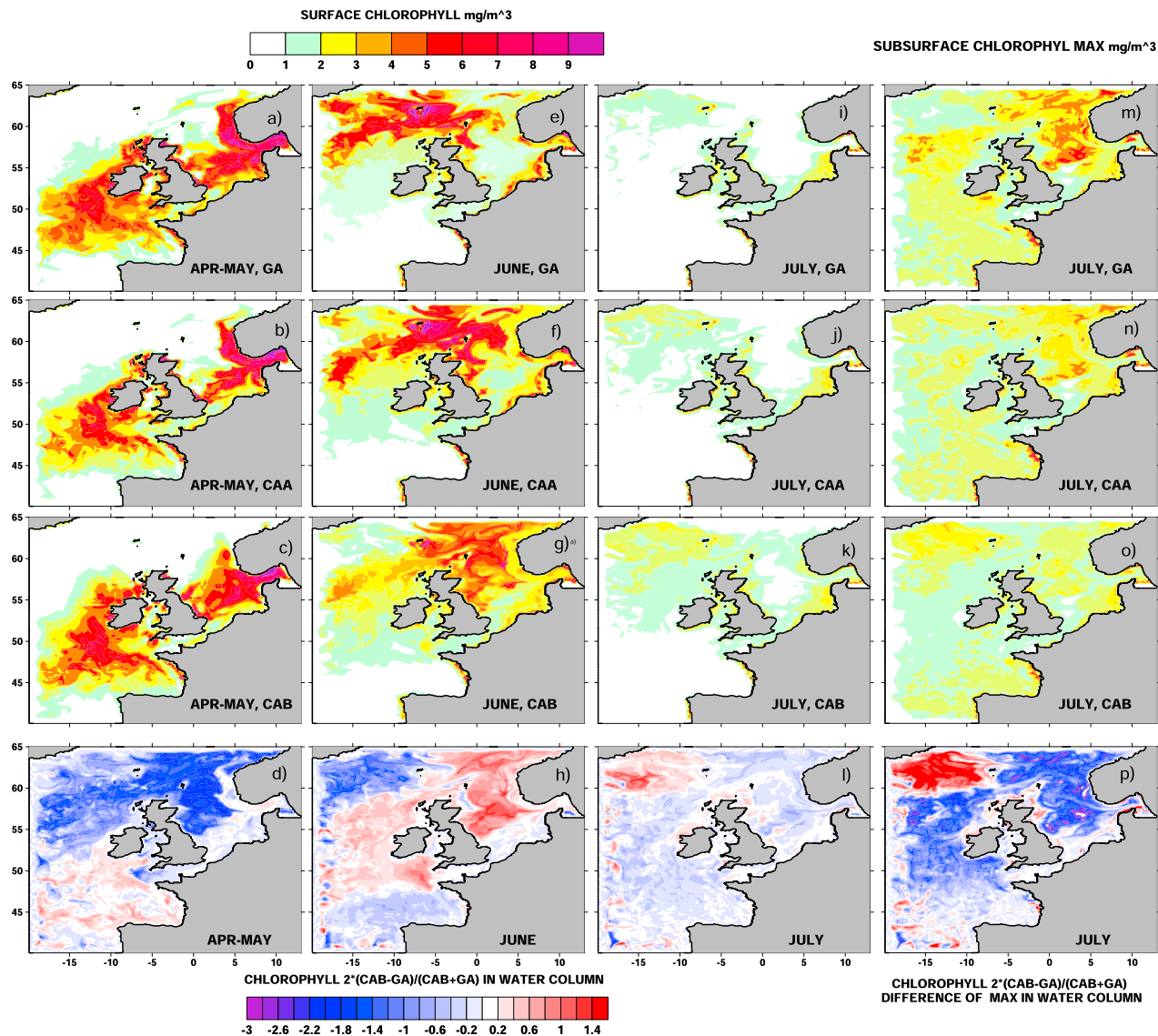
### 5.2. Implications for Ecosystem Modeling

Here we consider the degree to which the differences in stratification and mixing, caused by the different mixing schemes, impact on the simulation of the shelf sea ecosystem. We performed three runs with the most (CAB), the least (GA), and moderately (ST) diffusive TMSs (but with standard limiter parameters) using the ERSEM ecosystem model. These runs start from the same initial condition in 1 January 2014 from the multidecadal standard AMM7-ERSEM run, which corresponds to ST case.

Ecosystem validation is not a trivial matter and requires a multivariable and multimetric approach (Allen et al., 2007; Edwards et al., 2012). We leave a more detailed analysis to future work and for brevity focus on chlorophyll as a combined indicator of phytoplankton biomass and primary production. Figures 9a, 9d, and 9g show the water column integrated and area mean chlorophyll concentration in 2015 for the Celtic and North Seas and for the whole domain. These shelf sea regions have two distinct phytoplankton bloom peaks in April–May and June, whereas the whole domain has a peak in June (Figure 9g). The initial bloom starts earlier in the GA run, with about a 20-day delay in the timing of the bloom peak in the CAB run, which is later and stronger. We relate the earlier bloom in the GA run to an earlier onset of stratification in spring, which retains phytoplankton in the euphotic zone allowing earlier bloom initiation. The secondary peak is much stronger in the North Sea in the CAB run.

After a strong bloom, caused by the excess of nutrients in the CAB run built up before July, waters are then depleted of nutrients with a consequent strong reduction in chlorophyll compared to the GA run. The earlier bloom initiation in lower seasonal light conditions in the GA case results in a weaker and longer bloom (Figure 9d). Over the whole domain, depth-integrated temporal differences in chlorophyll reach 20–30%. Depth-integrated area and annual mean chlorophyll increase with the diffusiveness of the model (GA:  $39.3 \pm 33.6$  mg/m<sup>2</sup>, ST:  $41.5 \pm 37.9$  mg/m<sup>2</sup>, and CAB:  $43.4 \pm 42.7$  mg/m<sup>2</sup>).

The first strong bloom peak is associated with processes in the ML (Figures 9b, 9e, and 9h). Chlorophyll, integrated both over the ML and below the pycnocline, exhibits very strong temporal synoptic scale variability. This is related to changes in the MLD with wind (or night convection), destratification, and restratification due to solar radiation and has a variability of 20–40 mg/m<sup>2</sup> compared to 80–120 mg/m<sup>2</sup> mean peak magnitudes. Time mean chlorophyll is almost equally distributed between the upper ML and subsurface, pycnocline GA:  $17.5 \pm 17.3$  mg/m<sup>2</sup> in ML versus  $22.8 \pm 22.5$  mg/m<sup>2</sup> below ML (pycnocline), ST:  $19.9 \pm$



**Figure 10.** (a–c, e–g, i–k) Surface chlorophyll at the time periods shown in Figure 8: April–May, June, and July; (d, h, l) depth-integrated chlorophyll differences between CAB and GA for the same period; (m–o) maximum of chlorophyll below mixed layer in July; (p) relative difference between maximums of subsurface (below mixed layer) chlorophyll in CAB and GA runs.

20.7  $\text{mg/m}^2$  in ML versus  $22.7 \pm 24.2 \text{ mg/m}^2$  below, and CAB:  $23.6 \pm 25.2 \text{ mg/m}^2$  in ML versus  $21.0 \pm 25.1 \text{ mg/m}^2$  in pycnocline. Thus, the portion of subsurface chlorophyll reduces from 62% to 44% with increasing diffusiveness of the model. The nearly equal distribution of annual mean chlorophyll between ML and pycnocline is consistent with Rippeth et al.'s (2014) estimate of up to 50% of annual carbon fixation in the northwest European shelf being related to pycnocline mixing.

After the strong bloom in spring (80–120  $\text{mg/m}^2$ ) ML chlorophyll reduces to 10–20  $\text{mg/m}^2$  in strongly stratified conditions in summer (days 160–240). On the contrary, sub-ML level of chlorophyll is stabilized at the much higher level of 40–60  $\text{mg/m}^2$  (Figures 9c and 9f). On the shelf, in winter, chlorophyll levels nearly vanish due to strong cooling in well-mixed seas and reduced light levels.

Edwards et al. (2012) performed a detailed chlorophyll assessment of the NEMO-ERSEM operational model for the northwest European shelf versus satellite-derived chlorophyll and four “Smart Buoys,” located in relatively shallow regions. Seasonal mean variability (their Figure 9) in NEMO-ERSEM and satellite-

derived surface chlorophyll is  $\sim 1\text{--}10\text{ mg/m}^3$  during spring and summer, of the same order as the model results presented here (Figure 10).

Despite similar patterns in surface chlorophyll, depth-integrated relative differences between CAB and GA are large (Figures 10d, 10h, 10l, and 10p), exceeding 100% in different seasons. During April–May, depth-integrated chlorophyll values are higher by as much as 80% in the GA run in the central and northern parts of the North Sea (Figure 10d) and slightly higher in the CAB run in the Celtic Sea. In June, CAB chlorophyll exceeds that of the GA run almost everywhere on the shelf by 40–60%.

In accordance with Figures 9b, 9e, and 9h in July surface chlorophyll is reduced (Figures 10i–10k), while subsurface maxima below the ML exceed surface values by 5 times in some locations (Figures 10m, 10n, and 10p). Differences between models become more prominent in the shelf regions (Figures 10m and 10p) with 3 times larger subsurface concentrations in GA than in CAB.

Thus, the ecosystem characteristics are highly sensitive indicators of the quality of the turbulence closure scheme and in future could be used for the assessment of these schemes using, for example, ocean color data from satellites.

### 5.3. Which Closure Works Best?

There is no clear answer to this question. We considered here three different turbulent closures, based on works of GKHR, KC94, and CHCD with different physics and different critical steady state Richardson numbers. Two simulations considered here (KC and K3) correspond to different constants in KC94 and Kantha (2003), and two (CAA and CAB) in CHCD. Another variation arises from the parameter value and way the Deardorff/Galperin limiters are applied, which affects the steady state critical Richardson numbers and background diffusivity (Figure 1). This variation of parameters allows the TMS to be graduated from low diffusive to high diffusive with steady state Richardson numbers in the range from 0.12 (GA) to 0.492 (CAL).

The limiting parameter is critical: The KCL case, using the KC94 closure, appears to be slightly more diffusive than the ST run, which uses the CHCD closure, but in most characteristics their skill is indistinguishable (Figures 5, 7, and 8). Application of limiters is important: The weak limiting case (CAL versus CAA, Figure 4 a) results in extra heating of the bottom temperature ( $0.4\text{ }^\circ\text{C}$  compared to  $0.2\text{ }^\circ\text{C}$ ) and reverses the sign of heat content biases, when compared with the EN4 profile data set.

In all comparisons, variability (except SST in the Celtic Sea, Figure 7b) is better predicted by less diffusive models, with differences of about 10–20%.

At multiannual time scales more diffusive models, CAL-CAA, better control the heat content but result in an excessively smoothed pycnocline, compared to both the EN4 data set and the SCANFISH surveys. Conversely, less diffusive (KC-GA) TMS predict more precisely the strength of the pycnocline and bottom temperature but have strong negative biases in pycnocline depth (up to 8 m) and in heat content (up to  $0.3\text{ }^\circ\text{C}$ ).

In these experiments the surface layer thickness largely controls the SST and so the sensible heat flux and total heat content. Indeed, having the same shortwave radiation and advective heat flux, less diffusive models simulate a shallower pycnocline, higher SST, and hence less heat content in the domain compared with observations, with average temperatures lower by  $0.2\text{--}0.3\text{ }^\circ\text{C}$ . With a shallower pycnocline, it is easier to gain and lose heat in the ML during the diurnal cycle.

One of the most diffusive closures, CAB, produces negative biases in heat content and pycnocline depth but also has negative biases in the strength of stratification. We conclude that all the models have a general deficit of mechanically driven mixing, presumably due to the absence of parameterizations of missing physical processes on the shelf (Gerkema & Shrira, 2005; Green et al., 2008; Inall et al., 2011; Palmer, Inalland, et al., 2013; Palmer, Polton, et al., 2013; Palmer et al., 2015; Rippeth et al., 2009, 2014).

### 5.4. What Processes and Parameterizations May Be Missing in the Models?

It is clear from the analysis presented here that missing physical processes play a key role in the model-observation mismatch. Some of these processes are on their way to being resolved or already included in other models (e.g., long internal tides and the effects of gravity waves); the other group of processes are

nonhydrostatic and still require reliable parameterizations. Candidate “missing processes,” worthy of further consideration, are the following:

1. Shear instability of long internal tides, which is resolved in fine-resolution kilometer-scale models (Graham et al., 2018; Guihou et al., 2018), but not in the current study. Graham et al. (2018) shows that “over the continental shelf break, there are still warm biases compared with observations, however it was reduced compared with AMM7.” Interaction of increased tidal shear with NIOs, induced by wind, or wind-driven Ekman shear stresses (Burchard & Rippeth, 2009; Lenn et al., 2011), is a good candidate, which would lead to deepening of the ML and pycnocline.
2. Effects of gravity waves. The strongest impact on turbulence and shear stresses is produced by young waves (Brown & Wolf, 2009) that are sharper and steeper and so have a surface roughness parameter 3 times greater than for well-developed, old waves (Janssen et al., 2002) with stronger wave dissipation. A coupled ocean-wave-atmosphere system for the northwest European shelf region (Lewis et al., 2019) includes feedbacks on the turbulence due to shear stresses between ocean waves and the atmosphere and the effects of waves on ocean roughness. Those studies show a reduction of SST biases versus observations compared with the ocean-only model only, presumably due to an increase in ML and pycnocline depth.
3. LCs are another feature, induced by waves that is not present in the current parameterization. LCs are organized structures, similar to convective rolls, induced by the combined effects of current shear and wave Stoke's drift. Despite recent progress in the extension of turbulence closures to include LCs (Harcourt, 2015; Kantha & Clayson, 2004; Kukulka & Harcourt, 2017), the role of interaction between LCs, wave breaking, and strong stratification is not well understood. Most parameterizations are based on LESs that do not have the effects of wave breaking, and only a few experiments account for strong stratification. Craik-Leibovich instability theories (Craik & Leibovich, 1976; Leibovich, 1983) do not predict the development of instabilities, generating of LCs at the conditions of strong stratification or a high viscosity environment (the latter is the case of convection or when strong wave breaking happens). Another restriction from Craik-Leibovich theory is that LC instability is unlikely under conditions of short-wavelength dominant waves (Leibovich, 1983). The importance of swell waves was highlighted in MacWilliams et al. (2014) in a LES study, which demonstrated a strong effect of swells/wind alignment on the depth and intensity of LC. Thus, the presence of swell waves aligned with wind shear in the weakly stratified ocean is the preferred niche of LCs. LCs turbulent closures (Harcourt, 2015; Kantha & Clayson, 2004) developed for the  $(q^2, q^2l)$  equations could in principle be adapted to the GLS scheme (Kantha, 2004).
4. Direct effects of the horizontal (northward) component of the Coriolis force are usually neglected in turbulence closures. These terms act in a similar way to stratification, depending on the sign of the zonal velocity shear. If this sign is positive, rotation effects enhance mixing and vice versa. Galperin et al. (1989), using their extension of the GKHR model, found that in an unstratified flow these effects can enhance/suppress the depth of the Ekman mixing layer by 50%, with weaker (5–10%) effects in a stably stratified environment. Zikanov et al. (2003) examined the structure of the Ekman layer in unstratified flows in low and moderate latitudes and found a similar estimate of 50% in enhanced/suppressed mixing by the effects of the horizontal component of the Coriolis force.
5. Near inertial oscillations (NIOs) associated with the horizontal component of the Coriolis force (Shrira & Forget, 2015) provide another source of mixing. NIOs of this type cannot penetrate through the seasonal pycnocline and concentrate strong shear at the top of the pycnocline and so deepen it. Extra, unresolved mechanical mixing, if present, will deepen the summer pycnocline, decrease night and seasonal autumn heat loss from the sea surface, and reduce water column heat content loss.
6. Another source of turbulence in the tidally active seas is from sharp and steep nonlinear internal (solitary) waves. These are widely present in observations (Holt & Thorpe, 1997; Inall et al., 2000; Palmer et al., 2015; Rippeth & Inall, 2002) but absent in the AMM models, due to both resolution and the use of the hydrostatic assumption. Observations in the South China Sea (Liu and Lozovatsky, 2012) demonstrated that the averaged dissipation rate of TKE in the upper pycnocline is approximately linearly related with the available potential energy of the internal waves. Simulation of these waves requires both a horizontal resolution as fine as 50 m (Vlasenko et al., 2014) and the lifting of the hydrostatic assumption and so will not be realizable for a least the next 10 years at a shelf scale.

Thus, models, used for forecasting and hindcasting require new parameterizations of these processes based on in situ and satellite observations, theory, and LESs. MacKinnon et al. (2017) review in detail the results of the program “Climate Process Team on internal wave-driven ocean mixing,” aiming to improve parameterizations of internal wave driven ocean mixing in the deep ocean; however, shelf-sea processes are not included in that study.

## 6. Summary

In decadal scale simulations of the dynamics of the tidally active northwest European shelf seas, we assess the properties of eight vertical (diapycnal) TMS, present in the NEMO model. Each scheme (Table 1 and Figure 1) is a combination of a turbulent closure (Canuto et al., 2001; Galperin et al., 1988; Kantha, 2003; Kantha & Clayson, 1994), different sets of constants used in these closures, and the application of the Deandorf/Galperin limiter  $c_t$ .

Ozmidov (Deardorff/Galperin) limiting condition, derived from a scaling analysis, implies the upper limit on the turbulent length scale: At strong stratification, eddies exceeding the Ozmidov scale are unable to overturn. If limiting conditions are applied, critical gradient Richardson number in steady state becomes dependent on the limiter value and type of closure, summarized in Figure 1a. The formulation of background mixing is part of the mixing scheme (10a) and inversely proportional to stratification, with limiter-dependent coefficients (Figures 1b–1d).

Using the PEA as a measure of total stratification, TMSs are graduated from less diffusive to more diffusive (equation (15)): GA (Galperin et al., 1988) → KC (Kantha & Clayson, 1994) → K3 (Kantha, 2003) → CAA, CAB (Canuto et al., 2001), in agreement with the steady state critical Richardson numbers,  $Ri_{cl}$ , which varies from 0.12 to 0.492 in the different schemes (see Figure 1a and Table 1).

We characterize the density profiles in the shelf seas by five parameters (Figure 3a): surface and benthic temperatures, pycnocline depth, thickness, and strength (Figure 2d). To assess model results versus observation, we propose a measure, “model skills”, which combines errors in variability in biases for all considered characteristics in a single parameter (equations (16a), (16b) and (17)). As smaller model's errors, the closer “model skills” to unity.

For the chosen pycnocline characteristics we find model skill to be moderate to high (>0.6) for moderate and low diffusive TMSs (ST, K3, KC, and GA) on decadal time scales (Figure 5e).

In the simulations of “sea weather,” comparing seasonal simulations with SCANFISH data, only the low diffusive models (KC and GA) have satisfactory skills (>0.6) in the North Sea (Figure 8f). There is a strong area dependence in the model biases (Figure 7): In the synoptic scale assessment of the North Sea simulations bottom temperature is too warm (+3 °C) and  $N^2_{max}$  in the pycnocline is too low (−100%) in the more diffusive models, with biases twice as large as in less diffusive TMSs (+1.44 °C and −50%). In the tidally energetic Celtic Sea, low diffusive TMSs fail to simulate sufficiently deep pycnoclines, with biases reaching 8 m, twice as large as biases in the North Sea. Thus, there is not a “winner” between these TMSs.

In all simulations, the strength of pycnocline is too weak and pycnocline depth is too shallow. Analyzing the behavior of different metrics, we conclude that some processes responsible for the mechanical mixing are not included into the models (section 5.4). These might be the effects of young waves under conditions of growing winds, LCs caused by winds and swells (old waves), nontraditional NIOs generated as winds drop, strong shallow seasonal pycnocline, and others. Each of the missing processes occupies their own niche in physical parameter space.

We found that the sensitivity of models to mixing schemes and choice of limiters is comparable to the sensitivity to horizontal resolution (Graham et al., 2018) and surface flux formulation (O'Dea et al., 2017).

Finally, we find that simulations with the coupled ecosystem model show a strong effect of the choice of mixing scheme on the timing of phytoplankton blooms, on the spatial patterns of chlorophyll (Figures 9 and 10), and partitioning of chlorophyll production between ML and pycnocline. We found that, in accordance with previous observation-based estimates, nearly half of the yearly chlorophyll production occurs in the pycnocline, and this dominates in the strongly stratified period of the year (JJA). In less diffusive TMS (GA) ML spring blooming starts earlier and peak is weaker, compared with “moderate” ST and diffusive CAB runs.

The pycnocline bloom is longer lasting, with 20–30% greater intensity in the GA run. The portion of subsurface chlorophyll reduces from 62% to 44% with increasing diffusiveness of the TMS. This sensitivity motivates an assessment against high-frequency/resolution ecosystem and biogeochemical observations, but this is left for future work.

## 7. Code and Data Availability

AMM7 is a regional configuration of NEMO, at version 3.6 stable (Madec & NEMO Team, 2016). The model code is available from the NEMO website ([www.nemo-ocean.eu](http://www.nemo-ocean.eu)). ERSEM code (ERSEM16.06 commit e5272b7f) is available for download from the website ([https://www.pml.ac.uk/Modelling\\_at\\_PML/Models/ERSEM](https://www.pml.ac.uk/Modelling_at_PML/Models/ERSEM)). The EN4 database (version EN.4.2.1) is available from the Met Office website (<https://www.metoffice.gov.uk/hadobs/en4/>). SCANFISH data and model outputs are available by request to authors (mane1@noc.ac.uk). NEMO-GLS section of namelist and compilation keys is given in Appendix B.

## Appendix A: Critical Richardson Number in Equilibrium Regimes and the Deardorff/Galperin Limiter

The critical gradient Richardson number (Kantha, 2003, Cheng et al., 2003) is defined by the quadratic equation

$$(c_1 g_h^2 + c_4 g_h + 2) Ri^2 + (c_5 g_h - c_2 g_h^2) Ri + c_3 g_h^2 = 0; \quad (A1)$$

where  $g_h = (c_l^2 B_1^2)$  is the upper limit of  $g_h = G_h B_1^2$ . The list of constants,  $c_1, c_2, c_3, c_4, c_5, B_1$  for GA, K3 case can be found in Kantha (2003). For CHCD cases we calculate relevant constants using equations (18)d and (23c) from Cheng et al. (2002):

$$(c_1 \times 10^{-3}, c_2 \times 10^{-3}, c_3, \times 10^{-5}, c_4, c_5 \times 10^{-5}, B_1) = \left\{ \begin{array}{l} CHCD(A) : 1.60, 1.40, 1.09, 0.184, -3.91, 19.3 \\ CHCD(B) : 1.30, 1.30, 1.50, 0.158, -4.77, 16.6 \\ K3 : 2.87, 2.48, -0.18, 0.259, -4.73, 16.6 \\ GKHR : 6.89, 1.33, 3.55, 0.0, -4.72, 16.6 \end{array} \right\} \quad (A2)$$

The KC94 case has been taken from Umlauf & Burchard, 2005, Figure 3), as constants (A2) are not found in the published literature.

$$Ri_{cL} = \left\{ \begin{array}{l} c_l : 0.267, 0.53, +\infty \\ GA : 0.124, 0.165, 0.194 \\ KC : 0.143, 0.193, 0.2416 \\ K3 : 0.231, 0.479, 0.894 \\ CAA : 0.2497, 0.492, 0.849 \\ CAB : 0.2543, 0.521, 1.023 \end{array} \right\} \quad (A3)$$

## Appendix B: Details of NEMO Setup

The model has been compiled using the following fppkeys:

key\_dynspg\_ts key\_ldfslp key\_zdfgls key\_mpp\_mpi key\_netcdf4 key\_nosignedzero key\_traldf\_c2d key\_dynldf\_c2d key\_bdy key\_tide key\_vvl key\_iomput.

Full namelist\_ref and namelist\_cfg could be found in the supporting information. We varied the following options in namelist (shown in bold)



```

!-----
&namzdf_gls      !  GLS vertical diffusion      ("key_zdfgls")
!-----
ln_length_lim  = .true. (for all cases except CAL, KCL)
ln_sigpsi       = .true.
nn_bc_bot       = 1
nn_bc_surf      = 0
nn_clos         = 1
nn_stab_func   = 0 (varied from 0- to 4)
! 0=GKHR, 1=KC94, 2=CanutoA, 3=Canuto B, 4=K3, added option
nn_z0_met     = 1 (for ST case and =2 for others)
rn_charn        = 100000.0
rn_clim_galp    = 0.267
rn_crban        = 100.0
rn_emin         = 1.0e-6
rn_epsmin       = 1.0e-9
rn_frac_hs      = 1.3
rn_hsro         = 0.003
/

```

#### Acknowledgments

The authors acknowledge the National Capability funding in Ocean Modelling provided by the UK Natural Environment Research Council (NERC); NERC funded research grants PycnMix (grant NE/L003325/1) and Shelf Seas Biogeochemistry (NE/K001698/1). I. K. and E. Z. acknowledge the support by RFBR under grant 16-35-60072 mol\_a\_dk and by the Russian Science Foundation under grant 17-77-30019. We are grateful to two anonymous reviewers, whose valuable comments allowed us to significantly improve the manuscript. We also acknowledge use of the UK National High Performance Computing resource (<http://www.archer.ac.uk>).

#### References

- Abarbanel, H. D., Holm, D. D., Marsden, J. E., & Ratiu, T. (1984). Richardson number criterion for the non-linear stability of 3D stratified flow. *Physics Research Letters*, 52(26), 2352–2355. <https://doi.org/10.1103/PhysRevLett.52.2352>
- Allen, J. I., Holt, J. T., Blackford, J., & Proctor, R. (2007). Error quantification of a high-resolution coupled hydrodynamic ecosystem coastal-ocean model: Part 2. Chlorophyll-a, nutrients and SPM. *Journal of Marine Systems*, 68(3-4), 381–404. <https://doi.org/10.1016/j.jmarsys.2007.01.005>
- Andren, A., & Moeng, C.-H. (1993). Single point closures in a neutrally stratified boundary layer. *Journal of the Atmospheric Sciences*, 50(20), 3366–3379. [https://doi.org/10.1175/1520-0469\(1993\)050<3366:SPCIAN>2.0.CO;2](https://doi.org/10.1175/1520-0469(1993)050<3366:SPCIAN>2.0.CO;2)
- Badin, G., Williams, R. G., Holt, J. T., & Fernand, L. J. (2009). Are mesoscale eddies in shelf seas formed by baroclinic instability of tidal fronts? *Journal of Geophysical Research, Oceans*, 114, C10021. <https://doi.org/10.1029/2009JC005340>
- Blackford, J. C., Allen, J. I., & Gilbert, F. J. (2004). Ecosystem dynamics at six contrasting sites: A generic modelling study. *Journal of Marine Systems*, 52(1-4), 191–215. <https://doi.org/10.1016/j.jmarsys.2004.02.004>
- Brown, J., Carrillo, L., Fernand, L., Horsburgh, K. J., Hill, A. E., Young, E. F., & Medler, K. J. (2003). Observations of the physical structure and seasonal jet-like circulation of the Celtic Sea and St. George's Channel of the Irish Sea. *Continental Shelf Research*, 23(6), 533–561. [https://doi.org/10.1016/S0278-4343\(03\)00008-6](https://doi.org/10.1016/S0278-4343(03)00008-6)
- Brown, J., & Wolf, J. (2009). Coupled wave and surge modelling for the eastern Irish Sea and implications for model wind stress. *Continental Shelf Research*, 29, 1329–1342.
- 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. [https://doi.org/10.1175/1520-0485\(2001\)031<1943:CAOFSM>2.0.CO;2](https://doi.org/10.1175/1520-0485(2001)031<1943:CAOFSM>2.0.CO;2)
- Burchard, H., Craig, P. D., Gemmrich, J. R., van Haren, H., Mathieu, P.-P., Meier, H. E. M., et al. (2008). Observational and numerical modeling methods for quantifying coastal ocean turbulence and mixing. *Progress in Oceanography*, 76, 399–442. <https://doi.org/10.1016/j.pocean.2007.09.005>
- Burchard, H., & Rippeth, T. P. (2009). Generation of bulk shear spikes in a shallow stratified tidal seas. *Journal of Physical Oceanography*, 39, 969–985. <https://doi.org/10.1175/2008JPO4074.1>
- Butenschön, M., Clark, J., Aldridge, J. N., Allen, J. I., Artioli, Y., Blackford, J., et al. (2016). ERSEM 15.06: A generic model for marine biogeochemistry and the ecosystem dynamics of the lower trophic levels. *Geoscientific Model Development*, 9, 1293–1339. <https://doi.org/10.5194/gmd-9-1293-2016>
- Canuto, V. M., Howard, A., Cheng, J., & 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. [https://doi.org/10.1175/1520-0485\(2001\)031<1413:OTPIOP>2.0.CO;2](https://doi.org/10.1175/1520-0485(2001)031<1413:OTPIOP>2.0.CO;2)
- Canuto, V. M., Howard, A. M., Cheng, J., Muller, C. J., Lebbosier, A., & Jayne, S. R. (2010). Ocean turbulence, III: New GISS vertical mixing scheme. *Ocean Modelling*, 34, 70–91. <https://doi.org/10.1016/j.ocemod.2010.04.006>
- Carniel, S., Warner, J. C., & Chiggiato, J. M. (2009). Investigating the impact of surface wave breaking on modeling the trajectories of drifters in the northern Adriatic Sea during a wind-storm event. *Ocean Modelling*, 30, 225–239. <https://doi.org/10.1016/j.ocemod.2009.07.001>
- Charnok, H. (1955). Wind stress on a water surface. *Quarterly Journal of the Royal Meteorological Society*, 81(350), 639–640. <https://doi.org/10.1002/qj.49708135027>
- Cheng, Y., Canuto, V. M., & Howard, A. M. (2002). An improved model for the turbulent PBL. *Journal of the Atmospheric Sciences*, 59(9), 1550–1565. [https://doi.org/10.1175/1520-0469\(2002\)059<1550:AIMFTT>2.0.CO;2](https://doi.org/10.1175/1520-0469(2002)059<1550:AIMFTT>2.0.CO;2)

- Cheng, Y., Canuto, V. M., & Howard, A. M. (2003). Comments on "On an improved model for the turbulent PBL". *Journal of the Atmospheric Sciences*, 60(24), 3043–3046. [https://doi.org/10.1175/1520-0469\(2003\)060<3043:coaaim>2.0.co;2](https://doi.org/10.1175/1520-0469(2003)060<3043:coaaim>2.0.co;2)
- Cheng, Y., Canuto, V. M., & Howard, A. M. (2005). Non-local convective PBL model based on new third and fourth order moments. *Journal of the Atmospheric Sciences*, 62, 2189–2204. <https://doi.org/10.1175/JAS3474.1>
- Costa, A., Doglioli, A. M., Marsaleix, P., & Petrenko, A. A. (2017). Comparison of in situ microstructure measurements to different turbulence closure schemes in a 3-D numerical circulation ocean model. *Ocean Modelling*, 120, 1–17. <https://doi.org/10.1016/j.oceanmod.2017.10.002>
- Craik, A. D. D., & Leibovich, S. (1976). A rational model for Langmuir circulations. *Journal of Fluid Mechanics*, 73(3), 401–426. <https://doi.org/10.1017/S0022112076001420>
- Davies, A. M., & Kwong, S. C. M. (2002). Tidal energy fluxes and dissipation on the European continental shelf. *Journal of Geophysical Research*, 105(C9), 21,969–21,989. <https://doi.org/10.1029/2000JC900078>
- Deardorff, J. W. (1973). The use of subgrid transport equation in a three-dimensional model of atmospheric turbulence. ASME paper N73-FE21.
- Dee, D., Uppala, S., & Simmons, A. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597. <https://doi.org/10.1002/qj.828>
- Edwards, K. P., Barciela, R., & Butenson, M. (2012). Validation of NEMO-ERSEM operational ecosystem model for the north west European continental shelf. *Ocean Science*, 8, 983–1000. [www.ocean-sci.net/8/983/2012/](http://www.ocean-sci.net/8/983/2012/), <https://doi.org/10.5194/os-8-983-2012>
- Flather, R. A. (1976). A tidal model of the northwest European continental shelf. *Memoires de la Societe Royale des Sciences de Liege*, 10(6), 141–164.
- Flather, R. A. (1981). Results from a model of northeast Atlantic relating to the Norwegian coastal current. In R. Saetre & M. Mork (Eds.), *The Norwegian Coastal Current* (pp. 427–458). Bergen, Norway: University of Bergen.
- Galperin, B., Kantha, L. H., Hassid, S., & Rosati, A. (1988). A quasi-equilibrium turbulent energy model for geophysical flows. *Journal of the Atmospheric Sciences*, 45(1), 55–62. [https://doi.org/10.1175/1520-0469\(1988\)045<0055:AQETEM>2.0.CO;2](https://doi.org/10.1175/1520-0469(1988)045<0055:AQETEM>2.0.CO;2)
- Galperin, B., Rosati, A., Kantha, L. H., & Mellor, G. L. (1989). Modelling rotating stratified flows with application to oceanic mixed layers. *Journal of Physical Oceanography*, 19(7), 901–916. [https://doi.org/10.1175/1520-0485\(1989\)019<0901:mrstfw>2.0.co;2](https://doi.org/10.1175/1520-0485(1989)019<0901:mrstfw>2.0.co;2)
- Gargett, A. E., & Holloway, G. (1984). Dissipation and diffusion by internal wave breaking. *Journal of Marine Science*, 42(1), 15–27. <https://doi.org/10.1357/002224084788506158>
- Gerbi, G. P., Trowbridge, J. T., Terray, E. A., Plueddemann, A. J., & Kukulka, T. (2009). Observations of turbulence in the ocean surface boundary layer: Energetics and transport. *Journal of Physical Oceanography*, 39, 1077–1096. <https://doi.org/10.1175/2008JPO4044>
- Gerkema, T., & Shrira, V. I. (2005). Near inertial waves in the ocean: Beyond the "traditional approximation". *Journal of Fluid Mechanics*, 529, 195–219. <https://doi.org/10.1017/S0022112005003411>
- Gleckler, P. J., Taylor, K. E., & Doutriaux, C. (2008). Performance metrics for climate models. *Journal of Geophysical Research*, 113, D06104. <https://doi.org/10.1029/2007JD008972>
- Good, S. A., Martin, M. J., & Rayner, N. A. (2013). EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *Journal of Geophysical Research: Oceans*, 118, 6704–6716. <https://doi.org/10.1002/2013JC009067>
- Graham, J. A., O'Dea, E., Holt, J. T., Polton, J., Hewitt, H. T., Furner, R., et al. (2018). AMM15: A new high-resolution NEMO configuration for operational simulation of the European north-west shelf. *Geoscientific Model Development*, 11, 681–696. <https://doi.org/10.5194/gmd-11-681-2018>
- Green, J. M., Simpson, J. H., Legg, S., & Palmer, M. R. (2008). Internal waves, baroclinic energy fluxes and mixing at the European shelf edge. *Continental Shelf Research*, 28(7), 937–950. <https://doi.org/10.1016/j.csr.2008.01.014>
- Guihou, K., Polton, J., Harle, J., Wakelin, S., O'Dea, E., & Holt, J. (2018). Kilometric Scale modelling of the North West European Shelf Seas: Exploring the spatial and temporal variability of Internal Tides. *Journal of Geophysical Research: Oceans*, 123, 688–707. <https://doi.org/10.1002/2017JC012960>
- Harcourt R. R. (2015). An improved second-moment closure model of Langmuir turbulence. *Journal of Physical Oceanography*, 45, 84–103. <https://doi.org/10.1175/JPO-D-14-0046.1>
- Holloway, G., Nguyen, A., & Wang, Z. (2011). Oceans and ocean models as seen by current meters. *Journal of Geophysical Research*, 116, C00D08. <https://doi.org/10.1029/2011JC007044>
- Holt, J., Hughes, S., Hopkins, J., Wakelin, S. L., Holliday, N. P., Dye, S., et al. (2012). Multi-decadal variability and trends in the temperature of the northwest European continental shelf: A model-data synthesis. *Progress in Oceanography*, 106, 96–117. <https://doi.org/10.1016/j.pocean.2012.08.001>
- Holt, J., Schrum, C., Cannaby, H., Daewel, U., Allen, I., Artioli, Y., et al. (2016). Potential impacts of climate change on the primary production of regional seas: A comparative analysis of five European seas. *Progress in Oceanography*, 140, 91–115. <https://doi.org/10.1016/j.pocean.2015.11.004>
- Holt, J., & Umlauf, L. (2008). Modelling the tidal mixing fronts of the northwest European continental shelf. *Continental Shelf Research*, 28, 887–903. <https://doi.org/10.1016/j.csr.2008.01.012>
- Holt, J. T., & James, I. D. (2006). An assessment of the fine scale-eddies in a high resolution model of the shelf seas west of Great Britain. *Ocean Modelling*, 13, 271–291. <https://doi.org/10.1016/j.oceanmod.2006.02.005>
- Holt, J. T., & Thorpe, S. A. (1997). The propagation of high frequency internal waves in the Celtic Sea. *Deep Sea Research*, 44(12), 2087–2116. [https://doi.org/10.1016/S0967-0637\(97\)00091-5](https://doi.org/10.1016/S0967-0637(97)00091-5)
- Howard, L. N. (1961). Note on a paper by J. W. Miles. *Journal of Fluid Mechanics*, 10, 509–512.
- Inall, M., Aleynik, D., Boyd, T., Palmer, M., & Sharples, J. (2011). Internal tide coherence and decay over a wide shelf sea. *Geophysical Research Letters*, 38, L23607. <https://doi.org/10.1029/2011gl049943>
- Inall, M. E., Rippeth, T. P., & Sherwin, T. J. (2000). Impact of nonlinear waves on the dissipation of internal tidal energy at a shelf break. *Journal of Geophysical Research*, 105(C4), 8687–8705. <https://doi.org/10.1029/1999JC900299>
- Janssen, P. A. E. M., Doyle, J. D., Bidlot, J., Hansen, B., Isakwen, L., & Viterbo, P. (2002). Impact and feedbacks of ocean waves to atmosphere. In W. Pierre (Ed.), *Atmosphere-Ocean interactions, Advances in Fluid Mechanics*, 33, (Vol. 1, pp. 155–197). Boston, Massachusetts: WIT press.
- Kantha, L. H. (2003). On an improved model for the turbulent PBL. *Journal of the Atmospheric Sciences*, 60(17), 2239–2246. [https://doi.org/10.1175/1520-0469\(2003\)060<2239:OAIMFT>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)060<2239:OAIMFT>2.0.CO;2)
- Kantha, L. H. (2004). The length scale equation in turbulence models. *Nonlinear Processes in Geophysics*, 11, 83–97.

- Kantha, L. H., & Clayson, C. A. (1994). An improved mixed layer model for geophysical applications. *Journal of Geophysical Research*, 99(C12), 25,235–25,266. <https://doi.org/10.1029/94JC02257>
- Kantha, L. H., & Clayson, C. A. (2004). On the effect of surface gravity waves on mixing in the oceanic mixed layer. *Ocean Modelling*, 6(2), 101–124. [https://doi.org/10.1016/S1463-5003\(02\)00062-8](https://doi.org/10.1016/S1463-5003(02)00062-8)
- Kolmogorov, A. N. (1942). Equations of turbulent motion in an incompressible fluid. *Izvestiya Akademii Nauk SSSR, Seriya Fizicheskaya*, 6, 56–58.
- Kukulka, T., & Harcourt, R. R. (2017). Influence of Stokes drift decay scale on Langmuir turbulence. *Journal of Physical Oceanography*, 45, 84–103. <https://doi.org/10.1175/JPO-D-14-0046.1>
- Large, W., & Yeager, S. (2009). The global climatology of an interannually varying air–sea flux data set. *Climate Dynamics*, 33(2-3), 341–364. <https://doi.org/10.1007/s00382-008-0441-3>
- Lauder, B. E., & Jones, W. P. (1969). Sink flow turbulent boundary layers. *Journal of Fluid Mechanics*, 38(4), 817–831. <https://doi.org/10.1017/S002211206900262X>
- Leibovich, S. (1983). The form and dynamics of Langmuir circulations. *Annual Review of Fluid Mechanics*, 15(1), 391–427. <https://doi.org/10.1146/annurev.fl.15.010183.002135>
- Lenn, Y. D., Rippeth, T. P., Old, C. P., Bacon, S., Polyakov, I., Ivanov, V., & H€olemann, J. (2011). Intermittent intense turbulent mixing under ice in the Laptev Sea continental shelf. *Journal of Physical Oceanography*, 41, 531–547. <https://doi.org/10.1175/2010JPO4425.1>
- Lewellen, W. S., & Teske, M. (1973). Prediction of the Monin-Obukhov similarity functions from an invariant model of turbulence. *Journal of the Atmospheric Sciences*, 30(7), 1340–1345. [https://doi.org/10.1175/1520-0469\(1973\)030<1340:POTMOS>2.0.CO;2](https://doi.org/10.1175/1520-0469(1973)030<1340:POTMOS>2.0.CO;2)
- Lewis, H. W., Castillo Sanchez, J. M., Arnold, A., Fallmann, J., Saulter, A., Graham, J., et al. (2019). The UKC3 regional coupled environmental prediction system. *Geoscientific Model Development*, 12, 2357–2400. <https://doi.org/10.5194/gmd-12-2357-2019>
- Liu, Z. Y., & Lozovatsky, I. (2012). Upper pycnocline turbulence in the northern South China Sea. *Chinese Science Bulletin*, 57, 2302–2306. <https://doi.org/10.1007/s11434-012-5137-8>
- MacKinnon, J. A., Zhao, Z., Whalen, C. B., Waterhouse, A. F., Trossman, D. S., Sun, O. M., et al. (2017). Climate Process Team on internal wave-driven ocean mixing. *Bulletin of the American Meteorological Society*, 98, 2429–2454. <https://doi.org/10.1175/BAMS-D-16-0030.1>
- MacLachlan, C., Arribas, A., Peterson, K. A., Maidens, A., Fereday, D., Scaife, A. A., et al. (2015). Global Seasonal forecast system version 5 (GloSea5): A high-resolution seasonal forecast system. *Quarterly Journal of the Royal Meteorological Society*, 141, 1072–1084. <https://doi.org/10.1002/qj.2396>
- MacWilliams, J. C., Hucle, E., Liang, J., & Sullivan, P. P. (2014). Langmuir turbulence in swell. *Journal of Physics Oceanography*, 44, 870–890. <https://doi.org/10.1175/JPO-D-13-0122.1>
- Madec, G., & NEMO Team (2016). NEMO reference manual 3\_6\_STABLE: “NEMO ocean engine” Note du Pole de mod\_élisation, Inst. Pierre-Simon Laplace (IPSL), France, No 27 ISSN 1288-1619.
- Megann, A., Storkey, D., Aksenov, Y., Alderson, S., Calvert, D., Graham, T., et al. (2014). GO5.0: The joint NERC–Met Office NEMO global ocean model for use in coupled and forced applications. *Geoscientific Model Development*, 7, 1069–1092. <https://doi.org/10.5194/gmd-7-1069-2014>
- Mellor, G. L., & Yamada, T. (1974). A hierarchy of turbulence closure models for planetary boundary layers. *Journal of the Atmospheric Sciences*, 31(7), 1791–1806. [https://doi.org/10.1175/1520-0469\(1974\)031<1791:AHOTCM>2.0.CO;2](https://doi.org/10.1175/1520-0469(1974)031<1791:AHOTCM>2.0.CO;2)
- Mellor, G. L., & Yamada, T. (1982). Development of a turbulence closure model for geophysical fluid problems. *Reviews of Geophysics and Space Physics*, 20(4), 851–875. <https://doi.org/10.1029/RG020i004p00851>
- Miles, J. W. (1961). On the stability of heterogeneous shear flows. *Journal of Fluid Mechanics*, 10, 496–508.
- Monin, A. S., & Ozmidov, R. V. (1985). In H. Tennekes & D. Reidel (Eds.), *Turbulence in the ocean, trans from Russian*. Dordrecht/Boston/Lancaster: Publishing Co.
- O’Dea, E., Furner, R., Wakelin, S., Siddorn, J., While, J., Sykes, P., et al. (2017). The CO5 configuration of the 7 km Atlantic Margin Model: Large-scale biases and sensitivity to forcing, physics options and vertical resolution. *Geoscientific Model Development*, 10, 2947–2969. <https://doi.org/10.5194/gmd-10-2947-2017>
- Orlanski, I., & Bryan, K. (1969). Formation of the thermohaline step structure by large-amplitude internal waves. *Journal of Geophysical Research*, 74(28), 6975–6983. <https://doi.org/10.1029/JC074i028p06975>
- Palmer, M. R., Inalland, M. E., & Sharples, J. (2013). The physical oceanography of Jones Bank: A mixing hotspot in the Celtic Sea. *Progress in Oceanography*, 117, 9–24.
- Palmer, M. R., Polton, J. A., Inall, M. E., Rippeth, T. P., Green, J. A. M., Sharples, J., & Simpson, J. H. (2013). Variable behaviour in pycnocline mixing over shelf seas. *Geophysical Research Letters*, 40, 161–166.
- Palmer, M. R., Stephenson, G. R., Inall, M. E., Balfour, C., Dusterhus, A., & Green, J. A. M. (2015). Turbulence and mixing by internal waves in the Celtic Sea determined from ocean glider microstructure measurements. *Journal of Marine Systems*, 144, 57–69. <https://doi.org/10.1016/j.jmarsys.2014.11.005>
- Prandtl, L., and K. Wieghart (1945) Yeber ein Neues Formel System für die ausgebildete Turbulenz. *Nachr. Akad. Wiss., Göttingen, Mathematik-Phys.*, Kl. 6-19.
- Raschle, N., Arduin, F., Queffelec, P., & Croizé-Fillon, D. (2008). A global wave parameter database for geophysical applications. Part 1: Wave–current–turbulence interaction parameters for the open ocean based on traditional parameterizations. *Ocean Modelling*, 25, 154–171. <https://doi.org/10.1016/j.ocemod.2008.07.006>
- Rippeth, T. P., & Inall, M. E. (2002). Observations of the internal tide and associated mixing across the Malin Shelf. *Journal of Geophysical Research*, 107(C3), 3032. <https://doi.org/10.1028/2000JC000761>
- Rippeth, T. P., Lincoln, B. J., Kennedy, H. A., Palmer, M. R., Sharples, J., & Williams, C. A. J. (2014). Impact of vertical mixing on sea surface pCO<sub>2</sub> in temperate seasonally stratified shelf seas. *Journal of Geophysical Research: Oceans*, 119, 3868–3882. <https://doi.org/10.1002/2014JC010089>
- Rippeth, T. P., Palmer, M. R., Simpson, J. H., & Fisher, N. R. (2005). Thermocline mixing in summer stratified continental shelf seas. *Geophysical Research Letters*, 32, L05602. <https://doi.org/10.1029/2004GL022104>
- Rippeth, T. P., Wiles, P., Palmer, M. R., Sharples, J., & Tweddle, J. (2009). The diapycnal nutrient flux and shear-induced diapycnal mixing in the seasonally stratified western Irish Sea. *Continental Shelf Research*, 29, 1580–1587. <https://doi.org/10.1016/j.csr.2009.04.009>
- Rotta, J. C. (1951). Statistische theorie nichthomogener Turbulenz. *Zeitschrift für Physik*, 129(6), 547–572. <https://doi.org/10.1007/BF01330059>
- Shrira, V. I., & Forget, P. (2015). On the nature of near-inertial oscillations in the uppermost part of the ocean and a possible route toward HF radar probing of stratification. *Journal of Physical Oceanography*, 45, 2660–2678. <https://doi.org/10.1175/JPO-D-14-0247.1>

- Simpson, J. H., Crawford, W. R., Rippeth, T. P., Campbell, A. R., & Cheok, J. V. (1996). The vertical structure of turbulent dissipation in shelf seas. *Journal of Physical Oceanography*, *26*(8), 1579–1590. [https://doi.org/10.1175/1520-0485\(1996\)026<1579:TVSOTD>2.0.CO;2](https://doi.org/10.1175/1520-0485(1996)026<1579:TVSOTD>2.0.CO;2)
- Simpson, J. H., & Tinker, J. P. (2009). A test of the influence of tidal stream polarity on the structure of turbulent dissipation. *Continental Shelf Research*, *29*(1), 320–332. <https://doi.org/10.1016/j.csr.2007.05.013>
- Sorbjan, Z., & Grachev, A. A. (2010). An evaluation of the flux–gradient Richardson relationship in the stable boundary layer. *Boundary Layer Meteorology*, *135*, 385–405. <https://doi.org/10.1007/s10546-010-9482-3>
- Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, *106*(D7), 7183–7192. <https://doi.org/10.1029/2000JD900719>
- 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. <https://doi.org/10.1016/j.csr.2004.08.004>
- Vlasenko, V., Stashchuk, N., Inall, M. E., & Hopkins, J. E. (2014). Tidal energy conversion in a global hot spot: On the 3D dynamics of baroclinic tides at the Celtic Sea shelf break. *Journal of Geophysical Research: Oceans*, *119*, 3249–3265. <https://doi.org/10.1002/2013JC009708>
- Williams, C. A. J., Sharples, J., Green, M., Mahaffey, C., & Rippeth, T. P. (2013). The maintenance of the subsurface chlorophyll maximum in the western Irish Sea. *Limnology and Oceanography: Methods*, *3*, 61–73. <https://doi.org/10.1215/21573689-2285100>
- Williams, C. A. J., Sharples, J., Mahaffey, C., & Rippeth, T. P. (2013). Wind driven pulses of nutrients to phytoplankton in stratified shelf seas. *Geophysical Research Letters*, *40*, 5467–5472. <https://doi.org/10.1002/2013GL058171>
- Zhou, B., & Chow, F. K. (2012). Turbulence modelling for the stable atmospheric boundary layer and implication for wind energy. *Flow, Turbulence and Combustion*, *88*, 255–277. <https://doi.org/10.1007/s10494-011-9359-7>
- Zikanov, O., Slinn, D. N., & Dhanak, M. H. (2003). Large-eddy simulations of the wind-induced turbulent Ekman layer. *Journal of Fluid Mechanics*, *495*, 343–368. <https://doi.org/10.1017/S0022112003006244>