

# Comparing Ocean Surface Boundary Vertical Mixing Schemes Including Langmuir Turbulence

Qing Li<sup>1\*</sup>, Brandon G. Reichl<sup>2,3</sup>, Baylor Fox-Kemper<sup>1</sup>, Alistair J. Adcroft<sup>2,3</sup>, Stephen E. Belcher<sup>4</sup>, Gokhan Danabasoglu<sup>5</sup>, Alan L. M. Grant<sup>6</sup>, Stephen M. Griffies<sup>2,3</sup>, Robert Hallberg<sup>2,3</sup>, Tetsu Hara<sup>7</sup>, Ramsey R. Harcourt<sup>8</sup>, Tobias Kukulka<sup>9</sup>, William G. Large<sup>5</sup>, James C. McWilliams<sup>10</sup>, Brodie Pearson<sup>1,11</sup>, Peter P. Sullivan<sup>5</sup>, Luke Van Roekel<sup>12</sup>, Peng Wang<sup>10</sup>, Zhihua Zheng<sup>8</sup>

<sup>1</sup>Department of Earth, Environmental and Planetary Sciences, Brown University, Providence, RI, USA

<sup>2</sup>NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA

<sup>3</sup>Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, USA

<sup>4</sup>Met Office Hadley Centre, Exeter, UK

<sup>5</sup>National Center for Atmospheric Research, Boulder, CO, USA

<sup>6</sup>Department of Meteorology, University of Reading, Reading, UK

<sup>7</sup>Graduate School of Oceanography, University of Rhode Island, Narragansett, RI, USA

<sup>8</sup>Applied Physics Laboratory, University of Washington, Seattle, WA, USA

<sup>9</sup>School of Marine Science and Policy, University of Delaware, Newark, DE, USA

<sup>10</sup>Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, CA, USA

<sup>11</sup>College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR, USA

<sup>12</sup>Fluid Dynamics and Solid Mechanics, Los Alamos National Laboratory, Los Alamos, NM, USA

## Key Points:

- Six Langmuir turbulence parameterization schemes and five non-Langmuir schemes are compared in a common single column modeling framework.
- A suite of test cases of various scenarios are used, including typical global ocean conditions using JRA55-do.
- Significant discrepancies among schemes are found, and sorted by locations, seasons and forcing regimes.

\*Current Affiliation: Fluid Dynamics and Solid Mechanics, Los Alamos National Laboratory, Los Alamos, NM, USA.  
LA-UR-19-24938.

Corresponding author: Qing Li, qingli@lanl.gov

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination, and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2019MS001810

27 **Abstract**

28 Six recent Langmuir turbulence parameterization schemes and five traditional schemes  
 29 are implemented in a common single column modeling framework and consistently com-  
 30 pared. These schemes are tested in scenarios versus matched large eddy simulations (LES),  
 31 across the globe with realistic forcing (JRA55-do, WAVEWATCH-III simulated waves)  
 32 and initial conditions (Argo), and under realistic conditions as observed at ocean moor-  
 33 ings. Traditional non-Langmuir schemes systematically under-predict LES vertical mixing  
 34 under weak convective forcing, while Langmuir schemes vary in accuracy. Under global,  
 35 realistic forcing Langmuir schemes produce 6% ( $-1\%$  to 14% for 90% confidence) or  
 36 5.2 m ( $-0.2$  m to 17.4 m for 90% confidence) deeper monthly mean mixed layer depths  
 37 (MLD) than their non-Langmuir counterparts, with the greatest differences in extratropical  
 38 regions, especially the Southern Ocean in austral summer. Discrepancies among Langmuir  
 39 schemes are large (15% in MLD standard deviation over the mean): largest under wave-  
 40 driven turbulence with stabilizing buoyancy forcing, next largest under strongly wave-  
 41 driven conditions with weak buoyancy forcing, and agreeing during strong convective  
 42 forcing. Non-Langmuir schemes disagree with each other to a lesser extent, with a sim-  
 43 ilar ordering. Langmuir discrepancies obscure a cross-scheme estimate of the Langmuir  
 44 effect magnitude under realistic forcing, highlighting limited understanding and numerical  
 45 deficiencies. Maps of the regions and seasons where the greatest discrepancies occur are  
 46 provided to guide further studies and observations.

47 **1 Introduction**

48 Langmuir turbulence is a physical process that affects the turbulent mixing in the  
 49 ocean surface boundary layer (OSBL) and requires parameterization in ocean general cir-  
 50 culation models (OGCM). Various ocean boundary vertical mixing schemes with Lang-  
 51 muir turbulence have been proposed in the recent literature, some of which have already  
 52 been implemented and tested in OGCMs. Yet in what circumstance and to what extent do  
 53 different Langmuir turbulence parameterization schemes agree or disagree? Furthermore,  
 54 can the additional mixing due to Langmuir effects be detected beyond the uncertainty in  
 55 OSBL modeling?

56 The OSBL is a turbulent region acting as a buffer between the rapid variations of  
 57 the atmosphere and the slowly varying ocean interior. The competition between turbu-  
 58 lent mixing in the OSBL and restratification mechanisms (buoyant and dynamical) deter-  
 59 mines the depth of this layer, which modulates the exchange of heat, momentum and trace  
 60 gases (such as CO<sub>2</sub>) between the atmosphere and ocean, thus affecting the workings of the  
 61 broader Earth climate system. For example, seasonal variability of the OSBL affects mode  
 62 water formation and thereby water properties in the ocean interior [e.g., Stommel, 1979;  
 63 Sallée *et al.*, 2013a,b], and biological activities in the surface ocean and thereby global  
 64 biogeochemical cycles [e.g., Moore *et al.*, 2013; Rodgers *et al.*, 2014]. On a shorter time  
 65 scale, diurnal variability of the OSBL is important for the air-sea heat flux by affecting  
 66 the diurnal cycling of sea surface temperature [Price *et al.*, 1986; Large and Caron, 2015].  
 67 Due to their small horizontal spatial scale (10–100 m), short temporal scale (10<sup>3</sup>–10<sup>5</sup> s)  
 68 and non-hydrostatic nature, turbulent boundary layer mixing processes are not resolved  
 69 in OGCMs. Therefore, such models require parameterizations (a.k.a. closures or subgrid  
 70 schemes) that are both physically accurate and numerically efficient and stable.

71 Direct observational assessments of the quality of upper ocean mixing scaling re-  
 72 lationships are rare [e.g., Sutherland *et al.*, 2014; D'Asaro *et al.*, 2014]. The present pa-  
 73 per takes a process-based, multiple scheme approach: contrasting parameterizations in an  
 74 identical numerical and forcing framework, specifically constructed to avoid influences  
 75 from biases of the driving model, and considering a range of realistic conditions. This  
 76 approach is particularly good at identifying disagreement among parameterizations, but as-  
 77 sessing parameterization skill is challenged by the lack of an appropriate “truth” for eval-

uation. Here Large Eddy Simulations (LES) serve as truth, but these relatively expensive simulations cover only a limited set of forcing regimes. So, a broader study of classes of behavior under realistic forcing (e.g., deeper or shallower, more or less variable) will extend the analysis here to guide scheme selection for climate models and highlight uncertainties.

Turbulence within a free-surface boundary layer differs significantly from that in a wall-bounded layer [e.g., *D'Asaro*, 2014; *Harcourt*, 2015], primarily due to the presence of breaking surface waves and Langmuir turbulence [*McWilliams et al.*, 1997]. Langmuir turbulence is a disordered form of Langmuir circulation [*Langmuir*, 1938; *Thorpe*, 2004] resulting from the interactions of Stokes drift with Eulerian flow [*Craik and Leibovich*, 1976; *Teixeira and Belcher*, 2010]. While breaking waves are responsible for the elevated near-surface turbulent dissipation in field measurements [*Agrawal et al.*, 1992; *Terray et al.*, 1996; *Sullivan et al.*, 2007], which can be subsequently amplified due to the distortion of turbulence by Stokes drift [*Grant and Belcher*, 2009; *Teixeira*, 2012], Langmuir turbulence is largely responsible for a greater vertical velocity contribution to turbulent kinetic energy (TKE) within the wavy OSBL than is found in a wall-bounded layer [*D'Asaro*, 2001; *Tseng and D'Asaro*, 2004; *D'Asaro et al.*, 2014; *Sutherland et al.*, 2014]. Langmuir turbulence also contributes significantly to the entrainment at the base of the OSBL and thereby the mixed layer deepening [e.g., *Kukulka et al.*, 2009; *Li and Fox-Kemper*, 2017], probably due to its coherent structures including deeply penetrating jets [*Polton and Belcher*, 2007].

Among many wave-related physical processes that affect the global climate, Langmuir turbulence is prominent [*Cavaleri et al.*, 2012; *Belcher et al.*, 2012; *D'Asaro et al.*, 2014; *D'Asaro*, 2014]. The lack of explicit representations of Langmuir turbulence in all of the Coupled Model Intercomparison Project Phase 5 (CMIP5) models may have contributed significantly to many of the persistent biases in these models [*Belcher et al.*, 2012; *Li et al.*, 2016], such as the shallow mixed layer depth bias in the Southern Ocean [*Salée et al.*, 2013b]. Non-breaking wave turbulence [e.g., *Qiao et al.*, 2004; *Babanin and Haus*, 2009] is another theory for wave-turbulence interactions akin to Langmuir turbulence in that stronger non-breaking waves cause more mixing and which has stimulated parameterization studies [e.g., *Fan and Griffies*, 2014; *Qiao et al.*, 2016; *Chen et al.*, 2018]. One model (FIO) among the CMIP5 ensemble did include non-breaking wave turbulence. However, non-breaking wave schemes are not evaluated here as it is presently unclear whether these effects should be added to Langmuir turbulence or are instead a different theoretical framing of mostly the same set of phenomena.

In this study, we focus on six Langmuir turbulence parameterization schemes: KPPLT-VR12, KPPLT-LF17, KPPLT-R16, ePBL-LT, SMCLT-H15 and OSMOSIS (Table 1), and systematically assess their behavior in a number of controlled settings. For those unfamiliar with the aspects of Langmuir turbulence parameterizations and how they differ from non-Langmuir schemes, a brief review of common elements of Langmuir parameterizations is provided in section 2. The schemes considered here have different approaches for treating the various elements of Langmuir turbulence, but none currently treat the non-local flux described in section 2.4. In Table 1 and Appendix A: we summarize the six schemes to highlight their key features and implementation. To identify the effects of Langmuir turbulence, the non-Langmuir counterparts of these six schemes (KPP-CVMix, ePBL and SMC-KC94), as well as some popular variants (KPP-ROMS and SMC-C01A), are also included in our comparison suite. See Table 2 for a summary of these five schemes without Langmuir turbulence.

Given that the scope of this paper is confined to vertical mixing schemes in the OSBL, the term *ocean boundary vertical mixing scheme* will be shortened to *scheme* throughout the remainder of this paper for brevity. We will use *Langmuir schemes* to refer to the six ocean boundary vertical mixing schemes with Langmuir turbulence (Table 1), and *non-Langmuir schemes* to refer to the five ocean boundary vertical mixing schemes without Langmuir turbulence (Table 2).

# Accepted Article

**Table 1.** A summary of Langmuir turbulence schemes compared in this study. The 2nd to 5th columns correspond to elements from sections 2.1, 2.2, 2.5 and 2.7 in section 2, showing parameters for Langmuir turbulence (LT), modifications to the turbulent diffusivity  $K_L$  in equation (2) via velocity scale  $q$ , length scale  $l$  or coefficient  $C_L$ , down-Stokes drift shear momentum flux and enhanced entrainment for each scheme. Down  $\partial_z \mathbf{u}^S$  momentum flux is *implicit* for schemes that don't explicitly write out the Coriolis-Stokes force in the momentum equations, effectively assuming the simulated velocity being Lagrangian (see more discussion in Appendix A.1). The 6th column shows the non-Langmuir turbulence counterpart for each scheme (see Table 2 for a summary of these schemes). SP=Stokes production; ARSM=Algebraic Reynolds stress model.

Name	LT Parameters	$K_L$	Down $\partial_z \mathbf{u}^S$ Momentum Flux	Entrainment	Non-LT Counterpart	References
KPPLTVR12	LaSLP	$\mathcal{E}$ on $q$ (A.4)	Implicit	$\mathcal{E}$ on $q$ (A.4)	KPP-CVMix	McWilliams and Sullivan [2000]; Van Reekel et al. [2012]; Li et al. [2016]
KPPLT-LF17	LaSLP, LaSL	$\mathcal{E}$ on $q$ (A.4)	Implicit	LT-dependent $\overline{w'b'e}$ in $U_{TL}^2$ (A.5)	KPP-CVMix	Li and Fox-Kemper [2017]
KPPLTR16	LaSLP, $\mathbf{u}^S(z)$	$\mathcal{E}_K$ on $K_L$ (A.6)	$-K_L^L \partial_z \mathbf{u}^L$	$\mathcal{E}_{U_t^2}$ on $U_t^2$ (A.6)	KPP-CVMix + retuned $Ric$ (A.7)	Reichl et al. [2016]
ePBL-LT	LaSL	$q$ via (A.11)	Implicit	Extra energy due to LT: $m_{*LT} u_*^3$ (A.9)	ePBL	Reichl and Li [2019]
OSMOSIS	$L_{3t}, \delta^S$	$K_L$ via (A.16)	-	Prognostic Eqn. of $h_b$	-	Appendix A.3
SMCLTH15	$\mathbf{u}^S(z)$	SP in $q$ & $l$ Eqns., SP in ARSM for $C_L$	$-K_u \partial_z \mathbf{u} - K_u^S \partial_z \mathbf{u}^S$ (A.18)	-	SMC-KC94	Kantha and Clayson [2004]; Harcourt [2013, 2015]

**Table 2.** A summary of schemes without Langmuir turbulence compared in this study. The 2nd and 3rd columns show the different choices of the turbulent diffusivity  $K_\lambda$  and non-local flux  $\Gamma_\lambda$  in equation (2) in each scheme. ARSM=Algebraic Reynolds stress model.

Name	$K_\lambda$	$\Gamma_\lambda$	References
KPP-CVMix	Empirical $q$ & $C_\lambda$ , $l = h_b$ determined by discretized $Ri_b$ (A.2)	Non-zero $\Gamma_\theta$	Large et al. [1994]; Danabasoglu et al. [2006]; Griffies et al. [2015]
KPP-ROMS	Empirical $q$ & $C_\lambda$ , $l = h_b$ determined by integrated $Ri_b$ (A.3)	Non-zero $\Gamma_\theta$	Large et al. [1994]; McWilliams et al. [2009]
ePBL	Empirical $q$ , $l$ & $C_\lambda$ constraining mechanical energy	-	Reichl and Hallberg [2018]
SMC-C01A	Dynamic $q$ & $l$ Eqns. ( $k-\epsilon$ ), $C_\lambda$ from ARSM (weak-equilibrium)	-	Rodi [1987]; Canuto et al. [2001]; Warner et al. [2005]
SMC-KC94	Dynamic $q$ & $l$ Eqns. ( $k-kl$ ), $C_\lambda$ from ARSM (quasi-equilibrium)	-	Mellor and Yamada [1982]; Kantha and Clayson [1994]

A typical approach for examining OSBL schemes compares different schemes in a common OGCM [e.g., Large et al., 1997; Fan and Griffies, 2014; Li et al., 2016; Li and Fox-Kemper, 2017; Ali et al., 2019], often seeking to reduce model bias versus observations. However, this approach has results that are influenced strongly by other biases in the OGCM. For example, behavior of the Smyth et al. [2002] variant of KPP differs strongly in Fan and Griffies [2014] and Li et al. [2016]. Model intercomparison programs, e.g., CMIP5, are a way to find consistent behaviors and biases across OGCMs or climate models that include different OSBL schemes [e.g., Griffies et al., 2009; Sallée et al., 2013b; Griffies et al., 2016]. However, the differences among these models result from more than just their OSBL parameterizations; differences in numerics, parameterizations for other processes, and feedbacks between all of the processes and numerics contribute. The differences among Langmuir schemes are not made plain in both of these approaches. In addition, substantial coding efforts to get every scheme in every model and high computational cost are required for thorough evaluation. Cross-comparison among multiple schemes in multiple models with distinct numerics is impractical.

The present study takes a process-based approach by comparing different Langmuir schemes in a common single column modeling framework [GOTM5: Umlauf and Burchard, 2005; Umlauf et al., 2014] using identical calling model time-stepping and discretization (although variations in evaluating some common parameters are retained as part of the schemes). The comparison here is designed to minimize the complications associated with the use of OGCMs or climate models, and to compare each Langmuir scheme on an equal numerical footing with minimal coupled system feedbacks. Keeping a single calling model allows for some exploration of numerical robustness as well [following Van Roekel et al., 2018; Reichl and Hallberg, 2018]. CVMix [Griffies et al., 2015] is also used as an external library, where many KPP-based Langmuir schemes have been implemented [e.g., Reichl et al., 2016; Li and Fox-Kemper, 2017]. Since GOTM5 and CVMix are already incorporated in many OGCMs, the development effort for this comparison directly speeds adoption of any of these schemes in those models. One major goal of this comparison is to identify regimes, sorted by locations, seasons and dimensionless parameters, where there is disagreement between the Langmuir schemes. Such disagree-

161 ments highlight gaps in our understanding of the OSBL physics where either more LES  
 162 studies or more ocean observations are needed. This information is therefore valuable for  
 163 both ocean modelers and observationalists.

164 Similar comparisons of popular boundary vertical mixing schemes can be found in  
 165 both atmosphere [e.g., Ayotte *et al.*, 1996] and ocean [e.g., *Burchard and Bolding*, 2001;  
 166 *Umlauf and Burchard*, 2005; *Warner et al.*, 2005; *Burchard et al.*, 2008] applications. This  
 167 study is distinguished from these previous studies by focusing on schemes with Langmuir  
 168 turbulence and comparisons using forcing regimes relevant to both regional and global  
 169 climate models. In particular, some of the main questions to be addressed in this paper  
 170 include:

- 171 • To what extent do different Langmuir schemes agree and disagree?
- 172 • To what extent do the Langmuir schemes differ from the non-Langmuir schemes?
- 173 • When, where, and under what conditions do different Langmuir schemes agree or  
     disagree with each other and with non-Langmuir schemes?
- 174 • What future research directions will improve our understanding of Langmuir turbu-  
     lence and Langmuir schemes?

175 The remainder of this paper is organized as follows. In the next section, the ele-  
 176 ments of Langmuir turbulence parameterization schemes that conceptually distinguish  
 177 them from non-Langmuir schemes are reviewed for reference in later sections. In sec-  
 178 tion 3 the data and experimental setup of three test cases (ocean station moorings, global,  
 179 and versus LES) for the comparison are detailed. The results of the three test cases are  
 180 presented in section 4. Generally, the Langmuir schemes mix more strongly and deeply  
 181 than their non-Langmuir partners, but they are also in less agreement than non-Langmuir  
 182 schemes. Both numerical and physical constructions are involved in the discrepancies. A  
 183 comparison of the regimes of forcing realized in the realistic and idealized test cases in  
 184 section 4.1 highlights when the schemes most disagree. Discussions of the comparison re-  
 185 sults, highlighting remaining issues, key regimes for evaluation, and future research direc-  
 186 tions are presented in section 5. A brief summary and major conclusions of the compari-  
 187 son are presented in section 6. This paper includes four appendices, including a summary  
 188 of the key components of each Langmuir scheme and the implementation methods in Ap-  
 189 pendix A: , sensitivity tests for vertical resolutions and time steps in Appendix B: , notes  
 190 on the regimes of forcing (section 4.1) in Appendix C: , and a description of source code  
 191 and data accessibility in Appendix D: .

## 194 2 Elements of Langmuir Turbulence Parameterization Schemes

195 The aim of an OSBL vertical mixing scheme is to accurately and efficiently approxi-  
 196 mate the vertical turbulent flux of an arbitrary field  $\lambda$  in terms of a parametric dependence  
 197 on known (model resolved) properties and parameters. Here  $\lambda$  can be either a tracer or a  
 198 component of the momentum. A general form of the parameterized vertical turbulent flux  
 199 can be written as

$$200 \overline{w' \lambda'} \approx \mathcal{F}(\bar{\lambda}(z), \bar{S}(z), N(z), B_0, La_t, \dots), \quad (1)$$

201 where  $w'$  is the vertical component of the velocity  $\mathbf{u} = (u, v, w)$ ,  $(\bar{\ })$  is an ensemble mean,  
 202 often approximated by a horizontal and time mean (e.g., over a grid cell and time step  
 203 as commonly implemented in OGCMs), and  $(\ )'$  is the fluctuations from this mean. The  
 204 known variables in brackets can include mean fields  $\bar{\lambda}$ , the vertical shear of the mean hor-  
 205 izontal velocity  $\bar{S} = (\partial_z \bar{u}, \partial_z \bar{v}, 0)$ , stratification or the Brunt-Västälä frequency  $N$ , surface  
 206 boundary conditions (e.g., the surface buoyancy flux  $B_0$ ), and non-dimensional parame-  
 207 ters (e.g., the turbulent Langmuir number  $La_t$  to be introduced in equation (4)). Some  
 208 *second-moment closure* schemes evolve a system of covariances or second-moments, in-  
 209 cluding TKE and turbulent fluxes. Other schemes are based on *similarity* theory (e.g.,  
 KPP), which exploits empirical relationships in a variety of conditions.

To demonstrate the various effects of Langmuir turbulence that have been included in vertical mixing schemes, we use a specific form of equation (1) commonly adopted in first-moment closure and algebraic second-moment closure schemes. The concept of *turbulent diffusion* is commonly applied as

$$\overline{w' \lambda'} = -K_\lambda \partial_z \bar{\lambda} + \Gamma_\lambda. \quad (2)$$

The eddy diffusivity ( $K_\lambda > 0$ ) is commonly written as the product of a velocity scale  $q$ , a length scale  $l$  and a non-dimensional coefficient  $C_\lambda$  [Prandtl, 1925], all can be functions of depth  $z$ ,

$$K_\lambda = C_\lambda q l. \quad (3)$$

The last term in equation (2),  $\Gamma_\lambda$ , represents any flux not proportional to the local gradient of  $\bar{\lambda}$  and is thus known as the *non-local flux*. Various effects of Langmuir turbulence are addressed in Langmuir schemes by modifying components of equations (2) and (3), as detailed in the following sections.

## 2.1 Langmuir Turbulence Scales with Stokes Drift

Stokes drift [Stokes, 1847; Kenyon, 1969; Webb and Fox-Kemper, 2011, 2015; Myrhaug et al., 2014; van den Bremer and Breivik, 2018] is the key forcing component of the theory of Langmuir circulation and Langmuir turbulence [Craik and Leibovich, 1976; McWilliams et al., 1997]. The full Stokes drift profile is required for some schemes [Kantha and Clayson, 2004; Harcourt, 2013, 2015; Noh et al., 2016; Sinha et al., 2015; Reichl et al., 2016]. KPPLT-R16 and SMCLT-H15 represent this group in our comparison.

While the full profile is readily calculated from the wave spectrum predicted by third-generation wave models [Webb and Fox-Kemper, 2015], accurately transcribing the full profile across multiple components of a climate modeling system significantly increases the required information exchange with the wave model. Therefore, parametric profiles governed by parameters such as the Langmuir number are preferred [e.g., Sullivan et al., 2012; Webb and Fox-Kemper, 2015; Breivik et al., 2016; Li et al., 2017].

The most popular turbulent Langmuir number commonly used for different variants of KPP is given by

$$\text{La}_t = \left( u_*/u_0^S \right)^{1/2} \quad (4)$$

[McWilliams et al., 1997], which compares the magnitude of the friction velocity  $u_*$  and the surface magnitude of Stokes drift,  $u_0^S$ . These two velocity scales quantify the level of wind-driven shear turbulence and wave-driven Langmuir turbulence, respectively. Schemes using  $\text{La}_t$  include those in McWilliams and Sullivan [2000], Smyth et al. [2002], Takaya et al. [2010], Yang et al. [2015] and OSMOSIS, though the latter also requires the Stokes drift decay scale  $\delta^S$  assuming an exponential profile. This definition in (4) is attractive for its simplicity and its utility in scaling the energy budget [McWilliams et al., 1997; Grant and Belcher, 2009]. Although many of the schemes here use variants other than  $\text{La}_t$ , this parameter is used for all regime diagnoses in this paper (sections 4.1 and 4.3.2).

Notably,  $\text{La}_t$  neglects whether the Stokes drift is a result of long waves or short waves, i.e., the influence of the Stokes drift decay scale  $\delta^S$ . The degree to which the Stokes drift extends into the OSBL is important [Kukulka and Harcourt, 2017], as is the fact that it is the vertical shear of the Stokes drift rather than the surface value that results in production of energy. To account for these effects, Harcourt and D'Asaro [2008] define a Langmuir number  $\text{La}_{\text{SL}}$  from the surface layer (a fraction of the mixed layer, upper 20% in their definition) averaged Stokes drift  $\langle u^S \rangle_{\text{SL}}$  and a reference Stokes drift  $u_{\text{ref}}^S$  near the base of the mixed layer,

$$\text{La}_{\text{SL}} = \left[ u_*/\left( \langle u^S \rangle_{\text{SL}} - u_{\text{ref}}^S \right) \right]^{1/2}, \quad (5)$$

which is used in ePBL-LT and the scaling of Langmuir-enhanced entrainment in KPPLT-LF17.

255 The projected Langmuir number takes account for the effects of misalignment be-  
 256 tween wind and waves [Hanley *et al.*, 2010; Van Roekel *et al.*, 2012; Reichl *et al.*, 2016].  
 257 One variant based on  $\text{La}_{\text{SL}}$  is written as

$$\text{La}_{\text{SLP}} = \left[ \frac{u_* \cos \theta_{\text{wl}}}{\langle u^S \rangle_{\text{SL}} \cos (\theta_{\text{ww}} - \theta_{\text{wl}})} \right]^{1/2}, \quad (6)$$

258 where  $\theta_{\text{ww}}$  is the misalignment between wind and waves, and  $\theta_{\text{wl}}$  between wind and Lang-  
 259 muir cells. In GCMs, coupling with a wave model is required to estimate these angles  
 260 [e.g., Li *et al.*, 2016]. Different variants of  $\text{La}_{\text{SLP}}$  are used in KPPLT-VR12, KPPLT-R16  
 261 and the scaling of Langmuir-enhanced diffusivity in KPPLT-LF17.

## 262 2.2 Enhanced Turbulent Diffusivity $K_\lambda$

263 There is a general consensus that the diffusivity  $K_\lambda$  appearing in equation (3) should  
 264 be elevated in the presence of Langmuir turbulence to represent the enhanced vertical  
 265 mixing found in field measurements. The question is how? We can do so by enhanc-  
 266 ing  $K_\lambda$  directly (e.g., KPPLT-R16) or enhancing the velocity scale  $q$  [e.g., McWilliams  
 267 and Sullivan, 2000, KPPLT-VR12, KPPLT-LF17] in first-moment closure schemes using  
 268 an Langmuir number-dependent enhancement factor. Figure A.1 in Appendix A.1 com-  
 269 pares different formulas of the enhancement factor applied to  $q$  adopted in some of these  
 270 schemes. Note that directly enhancing  $q$  affects other components of the scheme that de-  
 271 pend on  $q$ , which is one of the differences between KPPLT-VR12 and KPPLT-R16. In  
 272 second-moment closure schemes, this is usually done by adding the contribution of Stokes  
 273 drift to the TKE and length scale equations and  $C_\lambda$  in equation (3) [e.g., Axell, 2002; Kan-  
 274 tha and Clayson, 2004, SMCLT-H15] or enhancing the length scale  $l$  in some other vari-  
 275 ants [e.g., Noh *et al.*, 2016].

## 276 2.3 Modified Vertical Profile of $K_\lambda$

277 In addition to its magnitude, the functional form of  $K_\lambda$  versus  $z$  is also modified  
 278 by the presence of Langmuir turbulence [e.g., Sinha *et al.*, 2015; Yang *et al.*, 2015]. This  
 279 modification is a direct result of the distinctive structures of Langmuir turbulence [Teixeira  
 280 and Belcher, 2010], with enhanced vertical TKE versus its horizontal components. This  
 281 effect can be accounted for by modifying the coefficient  $C_\lambda$  in equation (3), corresponding  
 282 to the shape function  $G_\lambda$  in KPP-based schemes [e.g., Sinha *et al.*, 2015; Yang *et al.*, 2015]  
 283 and the stability function  $S_\lambda$  in algebraic second-moment closure models such as SMCLT-  
 284 H15. However, a consensus has not been reached on such modifications. The Stokes drift  
 285 decay depth seems to affect this profile shape significantly [Kukulka and Harcourt, 2017].

## 286 2.4 Modified Non-Local Flux $\Gamma_\lambda$

287 The non-local flux  $\Gamma_\lambda$  in equation (2) is usually attributed to coherent structures in  
 288 the OSBL such as convective plumes [Troen and Mahrt, 1986; Large *et al.*, 1994]. Given  
 289 the coherent windrows associated with Langmuir circulation commonly seen in ocean sur-  
 290 face observations [Thorpe, 2004] and the deeply penetrating jets seen especially in LES  
 291 of Langmuir turbulence [Polton and Belcher, 2007], it is expected that such structures of  
 292 Langmuir turbulence will contribute to  $\Gamma_\lambda$ . The non-local fluxes of scalar fields and mo-  
 293 mentum are discussed extensively in the context of KPP [McWilliams and Sullivan, 2000;  
 294 Smyth *et al.*, 2002; Griffies *et al.*, 2015; Sinha *et al.*, 2015; Van Roekel *et al.*, 2018], but an  
 295 agreement on the effects of Langmuir turbulence has not been reached, perhaps because  
 296 non-local fluxes are typically diagnosed as the residual of local fluxes, causing errors in  
 297 local fluxes to appear in non-local flux diagnoses. Therefore, none of the KPP schemes  
 298 compared here alter this term.

## 299      2.5 Down-Stokes Drift Shear Momentum Flux

300      A down-Stokes drift shear momentum flux is ‘non-local’ to the Eulerian mean flow,  
 301      and can be treated as a non-local momentum flux in KPP-based schemes [e.g., *McWilliams*  
 302      and *Sullivan*, 2000; *Sinha et al.*, 2015]. The necessity of this term is also justified by the  
 303      ‘anti-Stokes’ effect [*Uchiyama et al.*, 2010; *McWilliams and Fox-Kemper*, 2013; *Haney*  
 304      *et al.*, 2015; *Suzuki and Fox-Kemper*, 2016; *Pearson*, 2018], which significantly modifies  
 305      the Eulerian mean velocity profile. The idea of down-Lagrangian shear mixing is exten-  
 306      sively examined in *McWilliams et al.* [2012] and adopted in Langmuir turbulence param-  
 307      eterization schemes such as the KPPLT-R16, where the same turbulent viscosity is used  
 308      for both the down-Eulerian shear and the down-Stokes drift shear components. Other  
 309      schemes, such as SMCLT-H15, require a generally different turbulent viscosity for the  
 310      down-Stokes drift shear component, due to the distinct dynamical effects of the Stokes  
 311      drift and the Eulerian shear.

## 312      2.6 Momentum Flux Distinct from Shear Orientation

313      Another issue of parameterizing the momentum flux is whether it is sufficient to use  
 314      just a scalar turbulent viscosity  $K_u$  in equation (2), especially in the presence of ocean  
 315      surface waves. It has been shown that the turbulent momentum flux in LES exhibits sig-  
 316      nificant misalignment from both Eulerian and Lagrangian velocity shear, though the mis-  
 317      alignment angle from the latter is much smaller [*McWilliams et al.*, 2012]. Thus, the diag-  
 318      nosed  $K_u$  in equation (2) assuming  $\Gamma_u = 0$  (or  $K_u^L$  relative to the Lagrangian mean flow)  
 319      is a tensor rather than a scalar. In addition, this misalignment depends on the decay scale  
 320      and direction of Stokes drift, the direction of the winds, and the relative strength of waves  
 321      and winds [*Van Roekel et al.*, 2012; *McWilliams et al.*, 2014; *Yang et al.*, 2015]. Different  
 322      Stokes and Eulerian turbulent viscosity in SMCLT-H15 will generate such a misalignment,  
 323      but none of the other schemes compared here using a  $K_u^L$  formulation represent this effect.

## 324      2.7 Enhanced Entrainment Due to Unresolved Shear

325      Entrainment at the base of the OSBL is important for the vertical fluxes of heat and  
 326      other tracers that have the source at the surface or below the OSBL [*McWilliams et al.*,  
 327      2014]. LES show enhanced entrainment in the presence of Langmuir turbulence [e.g.,  
 328      *McWilliams et al.*, 1997; *Grant and Belcher*, 2009; *McWilliams et al.*, 2014; *Li and Fox-*  
*Kemper*, 2017], partially due to plumes reminiscent of convective plumes [*Li*, 2018], as-  
 329      sociated with the penetrating jets of Langmuir turbulence [*Polton and Belcher*, 2007]. To  
 330      account for this effect, the unresolved shear term in KPP can be enhanced either by apply-  
 331      ing an enhancement factor (e.g., KPPLT-VR12 and KPPLT-R16), or by explicitly allowing  
 332      the entrainment buoyancy flux  $w'b'_e$  to depend on Langmuir number in the derivation of  
 333      this term [e.g., *McWilliams et al.*, 2014, KPPLT-LF17]. The ePBL-LT scheme accounts  
 334      for this effect by including the Langmuir turbulence contribution in the mechanical energy  
 335      budget.

## 337      2.8 Enhanced Entrainment due to Resolved Shear

338      Along with other mechanisms including wave breaking [*Kudryavtsev et al.*, 2008]  
 339      and a partition of the momentum flux between wave-induced and shear-induced compo-  
 340      nents [*Teixeira*, 2018], enhanced vertical mixing of momentum by Langmuir turbulence  
 341      also contributes to the weaker velocity shear near the surface than in a wall-bounded layer,  
 342      as seen in both observations [e.g., *Schudlich and Price*, 1998; *Gargett et al.*, 2004] and  
 343      LES [e.g., *McWilliams et al.*, 1997; *Kukulka et al.*, 2010]. As a result, the mean shear flow  
 344      at the base of the OSBL is elevated and contributes significantly to the mixed layer deep-  
 345      ening [*Kukulka et al.*, 2010]. In addition, the interaction between Langmuir turbulence  
 346      and inertial currents also needs to be considered to correctly simulate the shear-induced  
 347      mixing at the base of the OSBL, especially after a tropical cyclone passage [*Wang et al.*,

348 2018]. With sufficient vertical and temporal resolutions to resolve the shear near the base  
 349 of the OSBL and inertial currents, this effect is included in most of the Langmuir schemes  
 350 compared here, though it depends heavily on the parameterized momentum flux.

### 351 3 Data and Comparison Methods

352 The General Ocean Turbulence Model version 5 [GOTM5, *Umlauf et al.*, 2014, see  
 353 updated version on [gotm.net](http://gotm.net)] is a one-dimensional water column model with flexible con-  
 354 figurations of vertical mixing schemes. We make use of GOTM5 as a common driver for  
 355 all vertical mixing schemes. All Langmuir schemes to be compared were implemented di-  
 356 rectly in GOTM5 by adapting their current implementations in popular OGCMs or single  
 357 column models, or in the external CVMix package [*Griffies et al.*, 2015, see updated ver-  
 358 sion on [github.com/CVMix](https://github.com/CVMix)] and called from GOTM5. Implementations of these Langmuir  
 359 schemes in GOTM5 were verified separately and in various degrees of detail by compar-  
 360 ing with their original implementations under scenarios that were used for development.  
 361 To provide the necessary wave information for all Langmuir schemes and, where applica-  
 362 ble, Coriolis-Stokes force additions (see Table 1 and Appendix A.1), we implemented in  
 363 GOTM5 an interface that allows the input of Stokes drift via various methods. A list of  
 364 all these schemes can be found in Table 1 and Table 2. Brief descriptions of key compo-  
 365 nents of each scheme and the implementation methods in GOTM5 can be found in Ap-  
 366 pendix A: .

367 To reveal the differences among schemes in different scenarios, single column simu-  
 368 lations with GOTM5 were conducted in three test cases that serve distinct purposes. The  
 369 initial conditions and surface forcing of these simulations range from idealized scenar-  
 370 ios, as commonly used in LES studies, to more realistic scenarios, either as observed at  
 371 Ocean Stations and by Argo floats, or as covered in the global atmospheric surface forcing  
 372 dataset JRA55-do [*Tsujino et al.*, 2018] designed to drive global ocean/sea-ice mod-  
 373 els. JRA55-do is derived from a blend of observations and reanalysis, and represents an  
 374 estimate of the global atmospheric state that we use to derive ocean surface forcing condi-  
 375 tions.

376 All code and data are available online. See Appendix D: for more detail.

#### 377 3.1 Test Case 1: Simulations at Ocean Stations

378 Observed temperature and salinity profiles and surface fluxes from two ocean sta-  
 379 tions: (1) the Ocean Surface Mixing, Ocean Submesoscale Interaction Study in the north-  
 380 east Atlantic [OSMOSIS, 48.7°N, 16.2°W, *Damerell et al.*, 2016], and (2) Ocean Climate  
 381 Station Papa in the northern Pacific (OCS-Papa, 50.1°N, 144.9°W, [www.pmel.noaa.gov/OCS/Papa](http://www.pmel.noaa.gov/OCS/Papa))  
 382 and waverider buoy data (J. Thomson, APL-UW), were used to initialize and force GOTM5  
 383 simulations in this test case. Three sets of simulations were conducted, as listed in Ta-  
 384 ble 3, focusing on the deepening of mixed layer in winter, shoaling of mixed layer in  
 385 spring and a full annual cycle, respectively. Each set consists of 11 runs with different  
 386 schemes.

387 All simulations are run with  $\Delta z = 1$  m vertical grid spacing and  $\Delta t = 1$  min time  
 388 step, and with 3-hourly instantaneous output for analysis. It is previously demonstrated  
 389 that similar simulations are sensitive to numerical choices such as model time step and  
 390 vertical resolution [e.g., *Van Roekel et al.*, 2018; *Reichl and Hallberg*, 2018]. We there-  
 391 fore choose this numerical configuration similar to their fine resolution cases, where the  
 392 dependence of the simulation on model time step and vertical resolution is expected to  
 393 converge. In addition, combinations of three different vertical grids,  $\Delta z = [1, 5, 10]$  m, and  
 394 four different time steps,  $\Delta t = [1, 10, 30, 60]$  min, were tested to assess the sensitivity of  
 395 each scheme to the vertical resolution and time steps that are more typical of contempo-  
 396 rary climate modeling practice (See Appendix B: for more details). For simplicity, a lin-

**Table 3.** A summary of simulations at ocean stations. The domain refers to the vertical extent of the water column considered.

Name	Domain (m)	Simulation time	Description
OSMOSIS-Winter	200	09/22/2012 - 12/05/2012	Winter mixed layer deepening
OSMOSIS-Spring	480	12/25/2012 - 09/10/2013	Spring restratification
OCS-Papa	200	03/21/2012 - 03/20/2013	Full annual cycle

ear equation of state with thermal expansion coefficient,  $\alpha_T = 1.66 \times 10^{-4} \text{ }^{\circ}\text{C}^{-1}$ , and saline contraction coefficient,  $\beta_S = 7.6 \times 10^{-4} \text{ kg g}^{-1}$ , was used in all simulations. This choice also allows for easy diagnosis of the active (buoyancy  $b$ ) and passive (spice  $s$ ) tracer transports from the simulated temperature  $T$  and salinity  $S$  fields by

$$b = g [\alpha_T (T - T_{\text{ref}}) - \beta_S (S - S_{\text{ref}})], \quad (7)$$

$$s = g [\alpha_T (T - T_{\text{ref}}) + \beta_S (S - S_{\text{ref}})], \quad (8)$$

with  $g$  the gravitational acceleration, and  $T_{\text{ref}} = 10 \text{ }^{\circ}\text{C}$  and  $S_{\text{ref}} = 35 \text{ g kg}^{-1}$ . In order to suppress the strong inertial oscillations that tend to develop in the single column simulations, damping of horizontal velocity was applied by relaxation to zero with a 5 day decay time. This arbitrary damping mimics the transfer of energy from the inertial currents to the ocean interior by, e.g., exciting internal waves. Similar decay time has been used in a simple damped slab model of the mixed layer to obtain the inertial currents that best match the measurements [D'Asaro, 1985]. Varying this decay time by a factor of 2 does not strongly affect the results. We note, however, that the strong inertial oscillation generated by resonant winds is an important mechanism for pycnocline erosion [e.g., Skillingstad *et al.*, 2000], though it is uncertain how well this process is represented in our one-dimensional model.

Although we have the “truth” of evolving temperature, salinity, and velocity measured at the ocean stations, direct comparison with these observations requires additional terms in the single column simulations to close the heat and salt budgets, e.g., by prescribing horizontal and vertical heat and salt transport [Large, 1996]. Without these additional terms, which are not precisely known, to account for the effects of lateral processes and vertical advection, the simulated temperature and salinity drift away from the observation due to unbalanced surface fluxes. We therefore choose not to compare the simulation results with observations. Instead, the comparison is focused on quantifying the distinct behaviors of different schemes under realistic forcing. These differences will be discussed in a much broader context with the help of global JRA55-do forced single column simulations introduced next. The data at ocean stations is a valuable complement to the global JRA55-do results because of the high accuracy of the observed fluxes at these ocean stations when compared to the less accurate, but global view provided by JRA55-do.

### 3.2 Test Case 2: JRA55-do Forced Simulations

To provide a global view of the differences among the Langmuir schemes over an annual cycle, we conduct one set of GOTM5 simulations within each  $4^{\circ} \times 4^{\circ}$  box over the global ocean south of  $72^{\circ}\text{N}$  using realistic surface forcing and ocean surface wave information for each month. The polar regions are not covered by these simulations as a sea ice model will be required to compute the fluxes at the ice-sea interface and interactions with sea ice are not included in any of the schemes compared here. A full nonlinear equation of state would also be required in the polar regions. In total, we conducted around 24000 sets of such one-month single column simulations ( $\sim 2000$  [grid points]  $\times$  12 [months]). Each set of simulations contains 11 runs with different schemes, all initial-

435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
ized from the same Argo temperature and salinity profiles with zero velocity. Argo profile  
data from years 2004 to 2013 [WOD13, Boyer *et al.*, 2013] were used to generate the ini-  
tial conditions for these simulations. To select the most representative Argo profiles at a  
given location and time, all profiles within the domain and time window of  $\pm 20$  days rela-  
tive to the targeted day of the year were first collected in a pool. The actual Argo profiles  
nearest the median surface temperature of all the available profiles in the pool were se-  
lected as the initial conditions for that  $4^\circ \times 4^\circ$  box. The year of the Argo data was ignored  
to maximize the number of Argo profiles in use for better representability, noting that the  
focus here is comparison across the schemes under realistic conditions rather than repro-  
ducing the observations on any particular day or location. Argo profiles with too much  
missing data (more than 1/3 of the total within the upper 100 m) were removed.

446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
The JRA55 [Japan Meteorological Agency/Japan, 2013] based surface dataset for  
driving ocean-sea ice models [JRA55-do, version 1.3, Tsujino *et al.*, 2018] was used here  
to provide the surface forcing, including 3-hourly mean surface wind, air temperature, hu-  
midity, sea level pressure, precipitation and short wave radiation. The values of these vari-  
ables at the targeted location were taken from the nearest grid in the JRA55-do gridded  
data. Air-sea fluxes were calculated from these variables internally in GOTM5 following  
*Fairall et al.* [1996], using the air-sea temperature difference and relative wind velocity  
from the simulated sea surface temperature and currents. Ocean surface wave information  
was obtained from a WAVEWATCH III simulation forced by the same JRA55-do surface  
dataset for consistency. Instead of saving the full wave spectrum from WAVEWATCH III,  
the directional partitioned surface Stokes drift  $\mathbf{u}_{0i}^S$  with three frequency bins ( $n = 3$ ) were  
saved and input into GOTM5 to approximately reconstruct the full profile of Stokes drift  
from

$$\mathbf{u}^S(z) = \sum_{i=1}^n \mathbf{u}_{0i}^S \exp(2k_i z), \quad (9)$$

459  
460  
461  
462  
463  
464  
465  
466  
where  $\mathbf{u}_{0i}^S$  is determined by integrating the full wave spectrum over each frequency bin  
with a mean wavenumber  $k_i$ . This approach is found to be a suitable compromise in terms  
of accuracy of the profile and data storage considerations compared to using the full wave  
model wave spectra ( $n = 40$ ). The JRA55-do surface forcing data from years 2008-2009  
was used here, because it allows direct comparison with similar simulations forced by the  
Coordinated Ocean-ice Reference Experiments phase II [CORE-II, updated to the year  
2009, *Large and Yeager*, 2009]. The comparison of these forcing products in terms of  
their effects on mixing will be described in a separate paper.

467  
468  
469  
470  
471  
472  
The length of each simulation was set to one month, starting from the first day of  
a month. This approach is more favorable than running each simulation for a whole year  
(the approach taken in our initial tests). In this way, the simulation time is relatively short  
so that the drift away from initial conditions inherent in a single column model is limited,  
yet the simulations are still long enough to allow the different schemes to diverge from the  
initial condition and from one another.

473  
474  
475  
476  
477  
478  
All simulations are run on a uniform vertical grid extending to 500 m with  $\Delta z =$   
1 m grid spacing. Time step of  $\Delta t = 10$  min was used to reduce computational cost with-  
out significant degradation in the quality of the results (see Appendix B: ). Three-hourly  
instantaneous output were saved for analysis. We used the same linear equation of state as  
simulations at the ocean stations (Section 3.1). Again, nudging of the horizontal velocity  
to zero with a relaxation time scale of 5 days was applied to suppress inertial oscillations.

### 479    3.3 Test Case 3: Idealized Simulations and Comparison with LES

480  
481  
482  
483  
LES with various constant wind, waves, and destabilizing surface buoyancy flux  
forcing [*Li and Fox-Kemper*, 2017] and in tropical cyclone conditions [*Reichl et al.*, 2016]  
were used to tune parameters in KPPLT-LF17, KPPLT-R16 and ePBL-LT. The initial con-  
ditions and surface forcing in those studies are idealized, though the parameter space cov-

ered is intended to be realistic. The same initial temperature and salinity profiles and surface forcing conditions are used here to setup GOTM5 simulations with different schemes. It provides a suitable way to evaluate these schemes, without the complications of the effects of horizontal processes, instrumental error, etc. However, our LES experiments are missing many real world processes, such as breaking surface waves, bubbles, spray, and interactions with submesoscale fronts, all of which could be empirically represented within some of our GOTM5 schemes. Therefore, the most accurate GOTM5 scheme might fail to match precisely the corresponding LES because the LES is missing processes. However, most of the parameterization schemes were developed in a similar theoretical framing as the LES so this outcome is unlikely. Some of the schemes were designed with some of these LES results as a target; it is expected that these schemes will have an unfair advantage in performing most like their training LES. Nevertheless, it is our aspiration to explain some of the discrepancies among schemes found in the realistic forcing cases making use of these more controlled circumstances.

Two sets of GOTM5 simulations were conducted to compare with LES results of *Li and Fox-Kemper* [2017] and *Reichl et al.* [2016], respectively. The first set of GOTM5 simulations uses the initial conditions and surface forcing conditions of Case S-L1 and Case S-B in *Li and Fox-Kemper* [2017] (see their Table 1). The parameters in these simulations were set to be identical to the LES wherever possible. All the simulations run for 48 hours with time step of  $\Delta t = 1$  min and instantaneous output was saved every 15 min for analysis. This set of simulations focuses on the weak to moderate entrainment processes under constant wind, wave and destabilizing surface buoyancy conditions. The second set of GOTM5 simulations explores scenarios under an idealized hurricane with  $5 \text{ m s}^{-1}$  translation speed and maximum wind speed of  $65 \text{ m s}^{-1}$  [*Reichl et al.*, 2016]. In particular, the initial condition and surface forcing conditions were taken at the location of maximum wind speed, 50 km to the right of the path of the hurricane center. Consistent with the LES, all the GOTM5 simulations run on a vertical grid with maximum depth of 240 m and grid spacing of  $\Delta z = 1$  m. The total simulation time was 72 hours (wind peaked at around hour 37), with the same 1 min time step and 15 min output interval. This set of simulations focuses on the strong entrainment process driven by the transient high wind and waves during a hurricane, with significant deepening of mixed layer within a few hours. Unlike the simulations at ocean stations and the JRA55-do forced simulations, no damping to the velocity was applied here for both sets of simulations to be consistent with the LES. Note that momentum equations in these LES also include the Coriolis-Stokes force, which is accounted for in the GOTM5 simulations depending on the Langmuir schemes (see Table 1 and Appendix A.1).

## 4 Results

### 4.1 Regimes of Langmuir Turbulence: Realistic Parameter Space

Langmuir turbulence parameterizations are often developed through training and tuning against a particular set of LES. Consequently, different schemes emphasize distinct regions of the parameter space. One such example is the different formulas for the enhancement factor implied from different LES studies of Langmuir turbulence as compared in Figure A.1 (see more discussion in Appendix A.1). Therefore, we briefly review the parameter space explored in previous LES studies of Langmuir turbulence versus the parameter space covered in the JRA55-do test case here.

Here, we generalize the regime diagram in  $\text{La}_t - h/L_L$  parameter space introduced by *Belcher et al.* [2012]. The turbulent Langmuir number  $\text{La}_t$  defined in equation (4) measures the relative importance of the wind-driven shear turbulence and Langmuir turbulence in the TKE budget [*McWilliams et al.*, 1997]. The parameter  $h/L_L$  measures the relative importance of convection versus Langmuir turbulence [*Belcher et al.*, 2012], with

534  $h$  a length scale of the OSBL thickness and

$$535 \quad L_L = -\frac{u_*^2 u_0^S}{B_0} \quad (10)$$

536 the Langmuir stability length. The surface buoyancy flux  $B_0$  (dimensions of squared length  
 537 per cubed time) is defined such that a positive value corresponds to stabilizing (e.g., ocean  
 538 freshening or warming) conditions. Note that this is opposite to the definition used in  
 539 *Belcher et al.* [2012] so that a positive value of  $L_L$  corresponds to convection here. How-  
 540 ever, since we also define  $h$  to be positive, the parameter  $h/L_L$  is consistent with the defi-  
 541 nition in *Belcher et al.* [2012].

542 *Belcher et al.* [2012] categorize the surface forcing regime under destabilizing sur-  
 543 face buoyancy flux conditions ( $B_0 < 0$ ) into shear turbulence, Langmuir turbulence and  
 544 convective turbulence regimes (see their Figure 3). They do so according to the relative  
 545 contribution of each kind of turbulence to the turbulent dissipation rate  $\epsilon > 0$  (dimensions  
 squared length per cubed time) at the center of the OSBL,

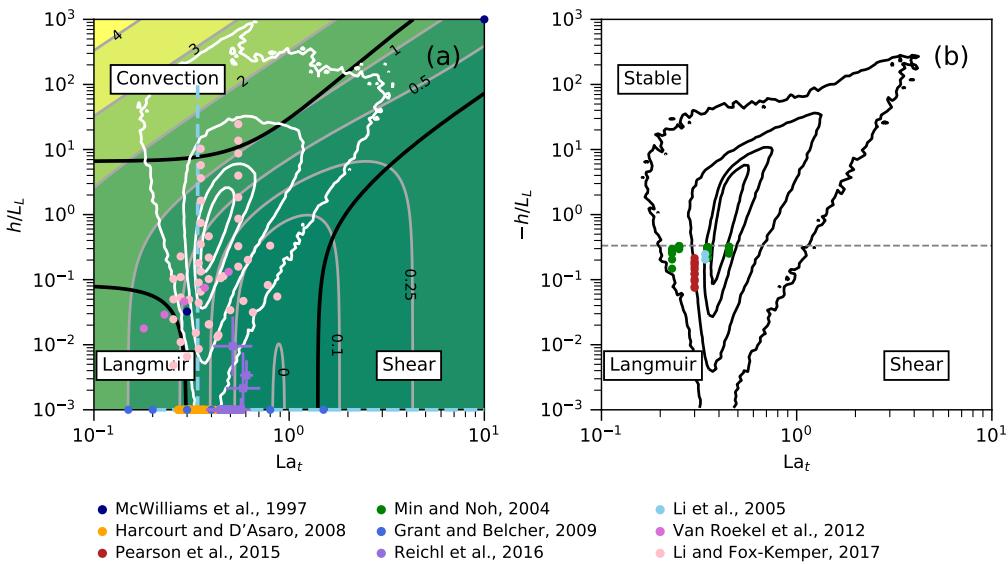
$$546 \quad \frac{\epsilon(z = -0.5h)}{u_*^3/h} = \underbrace{A_S}_{\text{Shear}} + \underbrace{A_L \text{La}_t^{-2}}_{\text{Langmuir}} + \underbrace{A_C \text{La}_t^{-2} \frac{h}{L_L}}_{\text{Convection}}, \quad (11)$$

547 with dimensionless coefficients ( $A_S = 2[1 - \exp(-0.5\text{La}_t)]$ ,  $A_L = 0.22$  and  $A_C = 0.3$ )  
 548 estimated from LES studies. Equation (11) is derived by scaling the three TKE produc-  
 549 tion mechanisms and assuming the dissipation is proportional to the production for each  
 550 mechanism [see the derivation and discussions in *Belcher et al.*, 2012]. This scaling serves  
 551 as an example of many possible ways to divide the parameter space into different regimes  
 552 in an asymptotic sense. Controversies exist on both the exact values of the dimensionless  
 553 coefficients and the functional form of this scaling, and its applicability to accurately de-  
 554 scribe the sea surface conditions in field measurements [e.g., *Large et al.*, 2019]. For the  
 555 purpose here, however, this regime diagram is helpful to contrast the parameter space al-  
 556 ready explored in LES of Langmuir turbulence with that in realistic global ocean datasets.  
 557 More discussion on the scaling of TKE dissipation and a theoretical model based on rapid  
 558 distortion theory with some other assumptions, particularly focusing on the effect of Lang-  
 559 muir turbulence, can be found in *Teixeira* [2012].

560 Note that *Belcher et al.* [2012] use the mixed layer depth (i.e.,  $h = h_m$ ) in equation  
 561 (11) and their Figure 3, focusing on destabilizing surface forcing conditions. Under stabi-  
 562 lizing surface forcing conditions, the active mixing layer, or the boundary layer, is much  
 563 shallower than the mixed layer [*Pearson et al.*, 2015]. Therefore, here we use the boundary  
 564 layer depth diagnosed in KPP-CVMix (i.e.,  $h = h_b$ ), which is generally consistent with  $h_m$   
 565 under destabilizing surface forcing conditions, but allows us to extend the regime diagram  
 566 to stable surface forcing regimes. Diagnosing  $h_b$  in KPP-CVMix or other schemes using a  
 567 diffusivity threshold [e.g., *Noh and Lee*, 2008] yield similar results.

568 Figure 1a reproduces Figure 3 of *Belcher et al.* [2012]. The background shows the  
 569 dimensionless turbulent dissipation rate,  $\log_{10}(\epsilon h/u_*^3)$ , under destabilizing buoyancy forc-  
 570 ing according to equation (11). Thick contours delineate regions where each of the three  
 571 kinds of turbulence dominates the energy. The overlaid white contours show contours of  
 572 the joint probability distribution function (PDF) of  $\log_{10}(\text{La}_t)$  and  $\log_{10}(h/L_L)$  that en-  
 573 close 30%, 60%, 90% and 99% of all  $B_0 < 0$  instances centered at the highest PDF. These  
 574 contours are estimated from a set of single column and global wave model simulations  
 575 using JRA55-do over 12 months described in section 3.2. It is important to note that this  
 576 figure has logarithmic axes, consistent with the wide range and asymptotic scaling under  
 577 study, but some care is required in interpreting probability densities in log-log figures (see  
 578 more discussion in Appendix C: ).

579 The probability distribution in Figure 1a represents about 62% of all the 3-hourly  
 JRA55-do data over 12 months, which is the likelihood of destabilizing buoyancy forcing.



**Figure 1.** Regime diagrams in parameter spaces (a)  $\text{La}_t - h/L_L$  following *Belcher et al.* [2012], and (b)  $\text{La}_t - -h/L_L$ , which describe situations with destabilizing and stabilizing surface buoyancy fluxes, respectively. Both panels are plotted in log-log space. Filled contours in the background of panel (a) show the normalized turbulent dissipation rate  $\log_{10}(\epsilon h/u_*^3)$  according to equation (11). Thick contours in black divide the parameter space into regions where each of the three kinds of turbulence forcing dominates (greater than 90%). White contours in panel (a) and black contours in panel (b) show the contours of the joint probability distribution function (PDF) of  $\log_{10}(\text{La}_t)$  and  $\log_{10}(h/L_L)$  ( $\log_{10}(-h/L_L)$  for panel (b)) that enclose 30%, 60%, 90% and 99% of all instances centered at the highest PDF. These distributions are estimated from the 3-hourly output of the GOTM5 simulations forced by JRA55-do over 12 months, representing about 62% and 38% of all the data, respectively. Overlaid dots in color mark the regimes explored in some representative LES studies of Langmuir turbulence. Squares with error bars show the mean and range (mean  $\pm$  standard deviation, calculated in  $\log_{10}$  space) of parameters represented in LES with transient forcing, in contrast to LES with constant forcing as shown in dots. Dashed lines in color show the approximate range of parameter space covered in those studies where the exact parameters were not explicitly reported. The gray dashed line in panel (b) shows the maximum equilibrium  $-h/L_L$  value suggested by *Pearson et al.* [2015].

580 This distribution indicates that when  $B_0 < 0$  the typical scale of  $h/L_L$  over the global  
 581 ocean places surface forcing conditions from the 3-hourly JRA55-do data between Lang-  
 582 muir turbulence and convection. The other 38% of forcing has stabilizing surface buoy-  
 583 ancy fluxes ( $B_0 > 0$ ), shown in the  $\text{La}_t - -h/L_L$  parameter space in Figure 1b. Under both  
 584 destabilizing and stabilizing surface conditions, wind-driven shear turbulence seldom dom-  
 585 inates the dissipation by this measure, though its role on the vertical mixing in the OSBL  
 586 is likely underrepresented in this regime diagram partially due to its strong coupling with  
 587 Langmuir turbulence.

588 Scattered dot and square (with error bars) symbols and dashed lines in color mark  
 589 the regimes explored in some representative LES studies of Langmuir turbulence in the  
 590 literature, some of which were used in training the schemes here. In particular, orange  
 591 dots mark the regimes of realistic wind and waves in pure wind sea explored in *Harcourt*  
 592 and *D'Asaro* [2008], of which the LES data was used to tune parameters in SMCLT-H15.

593 Blue dots mark the regimes explored in *Grant and Belcher* [2009] with various  $\text{La}_t$ ,  $u_*$   
 594 and Stokes drift decay depth assuming an exponential profile, and red the regimes ex-  
 595 plored in *Pearson et al.* [2015] with various surface heating conditions and rotation effects.  
 596 Both sets of LES were used to tune parameters in OSMOSIS. Neutral surface forcing con-  
 597 ditions in both *Harcourt and D'Asaro* [2008] and *Grant and Belcher* [2009] correspond to  
 598  $h/L_L = 0$ , represented in this diagram by symbols on the bottom edge for convenience.  
 599 Magenta dots mark the regimes of misaligned wind and waves explored in *Van Roekel*  
 600 *et al.* [2012], of which the scaling for vertical velocity variance was used in KPPLT-VR12  
 601 to estimate the enhancement factor. Pink dots mark the regimes explored in LES of mixed  
 602 Langmuir turbulence and convection by *Li and Fox-Kemper* [2017], which was used in  
 603 both KPPLT-LF17 and ePBL-LT to constrain the entrainment buoyancy flux and inte-  
 604 grated buoyancy flux over the entrainment zone, respectively. Purple squares with error  
 605 bars (mean  $\pm$  standard deviation) mark the regimes spanned in the transient conditions of  
 606 idealized hurricanes explored in *Reichl et al.* [2016], which were used to tune parameters  
 607 in KPPLT-R16.

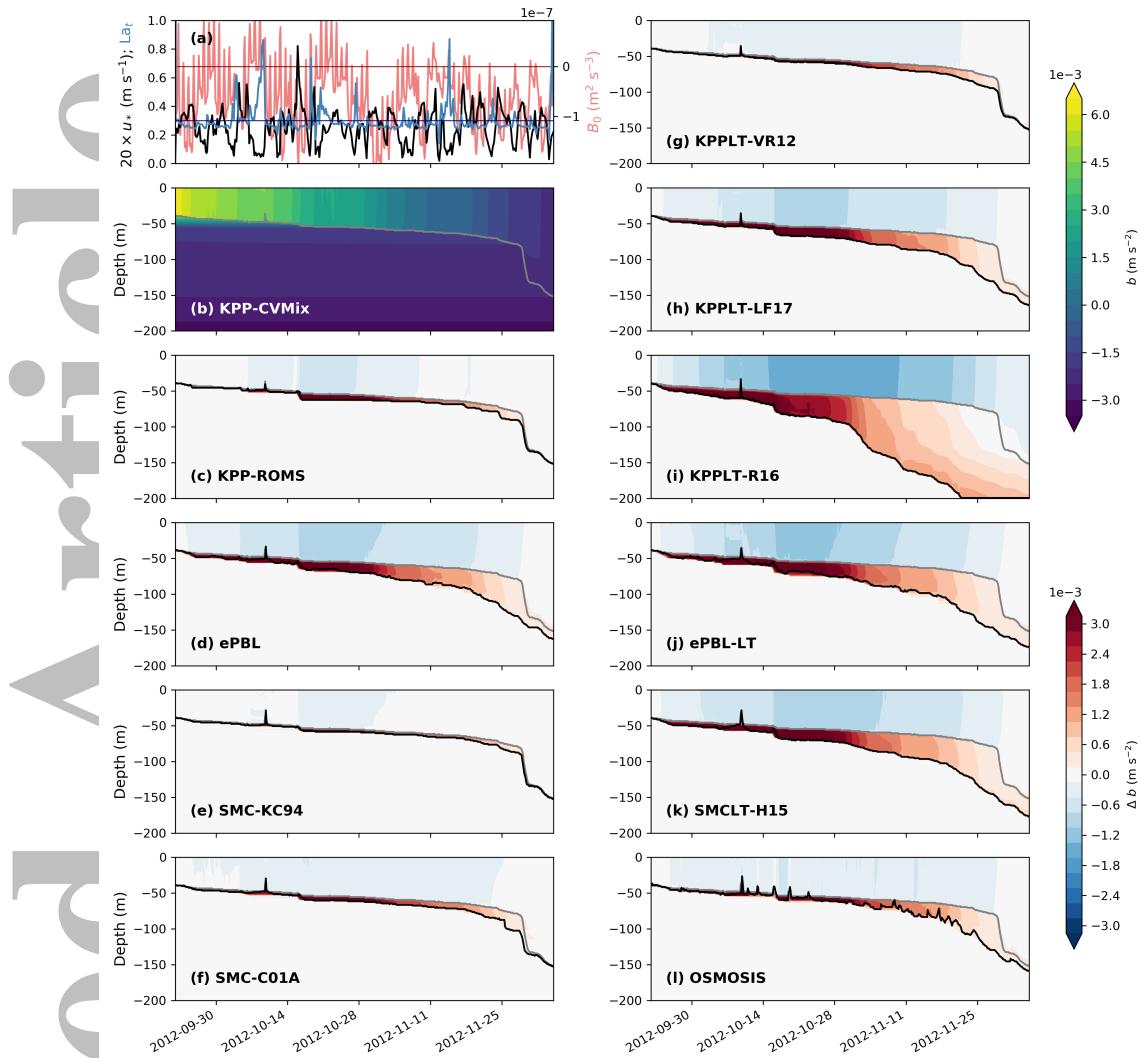
608 For LES under constant forcing in stable regimes (colored dots in Figure 1b),  $h/L_L$   
 609 was estimated from  $h_i/L_L$  with  $h_i$  the initial mixed layer depth using equation (4) of *Pear-*  
*610 son et al.* [2015]. This equation suggests a maximum equilibrium value of  $-h/L_L = 1/3$  in  
*611* the limit of strong surface heating, which is shown by the gray dashed line. However, un-  
*612* der transient forcing  $-h/L_L$  can be significantly larger as illustrated by the black contours.

613 It should be noted that the symbols and dashed lines here only represent the projec-  
 614 tions of the actual regimes explored in each study onto the  $\text{La}_t - h/L_L$  parameter space.  
 615 There are other dimensions of the full parameter space explored in these studies that are  
 616 not represented in this regime diagram, such as different penetration depths of Stokes  
 617 drift in *Harcourt and D'Asaro* [2008], different wind-wave misalignment conditions in  
 618 *Van Roekel et al.* [2012], and different combinations of surface and penetrative (radiative)  
 619 heat fluxes in *Pearson et al.* [2015]. Even so, this regime diagram reveals that many parts  
 620 of the parameter space are still not well explored, especially those of Langmuir turbulence  
 621 under transient stabilizing surface conditions (Figure 1b). Likewise, were the probability  
 622 distributions for the station observations in Test Case 1 (section 3.1, Table 3) shown, it  
 623 would be evident that the LES differ greatly from the forcing regimes seen at the OCS-  
 624 Papa and OSMOSIS sites. For this reason, it is insufficient to evaluate schemes here ver-  
 625 sus only LES.

## 626 4.2 Test Case 1: An Overview of the Comparison

627 This section presents results of Test Case 1, which provides an overview of the com-  
 628 parison among 11 schemes under realistic forcing conditions. Time evolution of the tem-  
 629 perature and salinity profiles simulated by GOTM5 for OSMOSIS-Winter, OSMOSIS-  
 630 Spring and OCS-Papa were examined, as well as buoyancy and spice assuming a linear  
 631 equation of state. No significant distinction was found among temperature, salinity, buoy-  
 632 ance, and spice behaviors of different schemes. Therefore, we only present the evolution  
 633 of buoyancy profiles for the three cases here as shown in Figures 2, 3 and 4, respectively.

634 OSMOSIS-Winter and OSMOSIS-Spring together represent a full annual cycle of  
 635 the mixed layer at a mid-latitude region in the North Atlantic, with strong variations in the  
 636 mixed layer depth. The winter mixed layer deepening is rapid due to a strong destabiliz-  
 637 ing surface buoyancy flux, leading to a mixed layer depth of a few hundred meters starting  
 638 from the rather shallow summer mixed layer depth of less than 50 m. Intermittent deepen-  
 639 ing events occur during the spring shoaling of the mixed layer. The water properties in the  
 640 mixed layer are significantly influenced by horizontal advection of different water masses  
 641 and submesoscale processes [Damerell et al., 2016; Thompson et al., 2016], which are ig-  
 642 nored here in the single column simulations. The turbulent Langmuir number is around

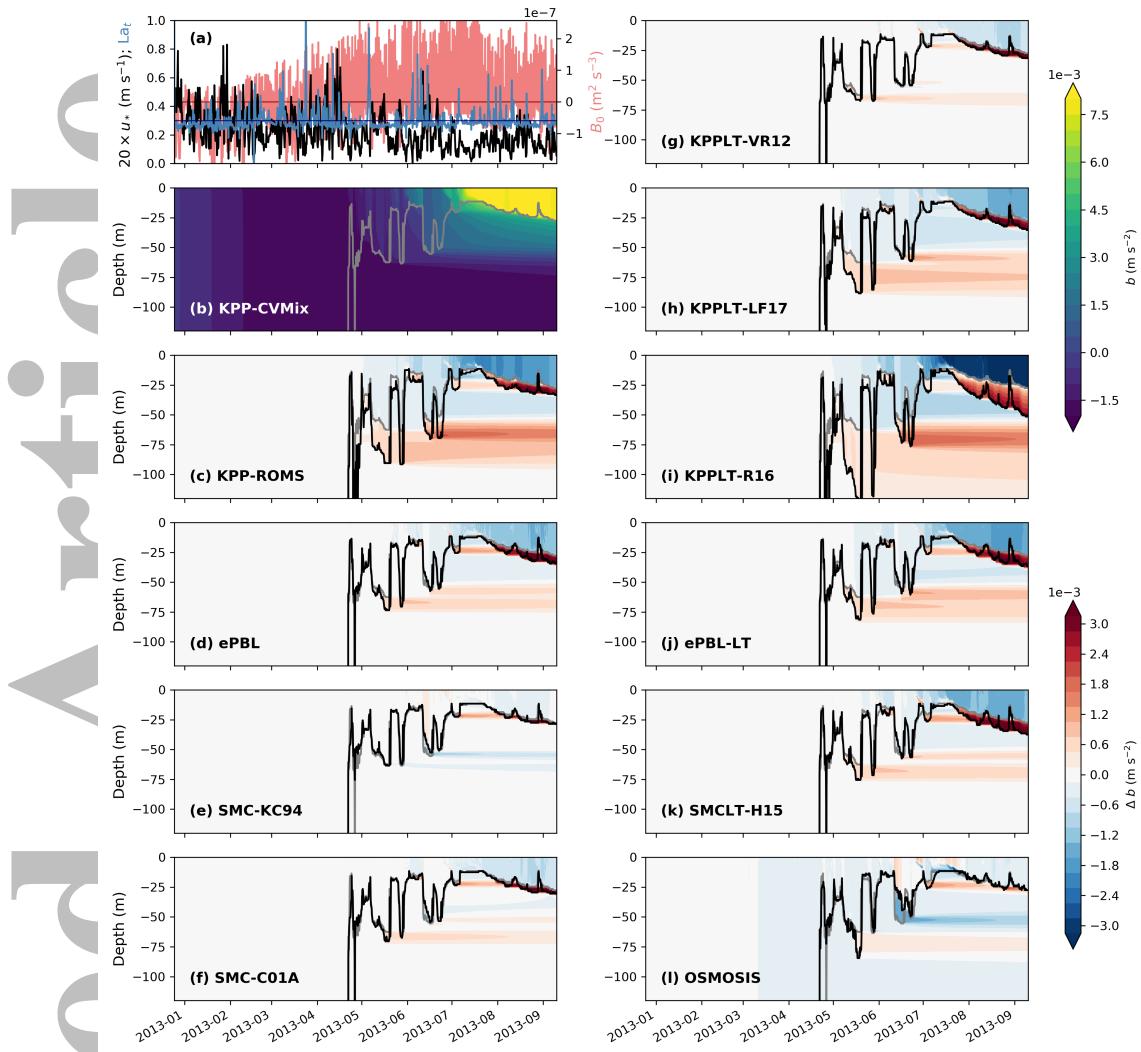


**Figure 2.** Comparison of simulated buoyancy with different schemes for the OSMOSIS-Winter case.

Panel (a) shows time series of water-side surface friction velocity  $u_*$  ( $\text{m s}^{-1}$ ) in black, turbulent Langmuir number  $\text{La}_t$  in blue ( $\text{La}_t = 0.3$  in dark blue for reference), and surface buoyancy flux  $B_0$  ( $\text{m}^2 \text{s}^{-3}$ ) in red ( $B_0 = 0 \text{ m}^2 \text{s}^{-3}$  in dark red for reference). For clarity,  $20 \times u_*$  is shown, sharing the same vertical axis with  $\text{La}_t$  on the left, whereas  $B_0$  uses the vertical axis on the right. Panel (b) shows Hovmöller diagram of the simulated buoyancy  $b$  ( $\text{m s}^{-2}$ ) for KPP-CVMix, with the mixed layer depth (MLD) defined by the  $0.03 \text{ kg m}^{-3}$  density criterion [*de Boyer Montégut et al., 2004*] marked in gray. For reference, a buoyancy anomaly of  $3.0 \times 10^{-3} \text{ m s}^{-2}$  is the equivalent of  $1.8^\circ\text{C}$  temperature or  $0.40 \text{ g kg}^{-1}$  salinity anomaly. Panels (c)-(l) show the differences in simulated buoyancy from KPP-CVMix for all other schemes, with the MLD marked in black. For comparison, the MLD for KPP-CVMix is also shown in (c)-(l) in gray. Note that the order of the panels is chosen so that panels on the left show the results of non-Langmuir schemes, and panels on the right show results of Langmuir schemes. KPP-CVMix is arbitrarily used as the reference here only for demonstration.

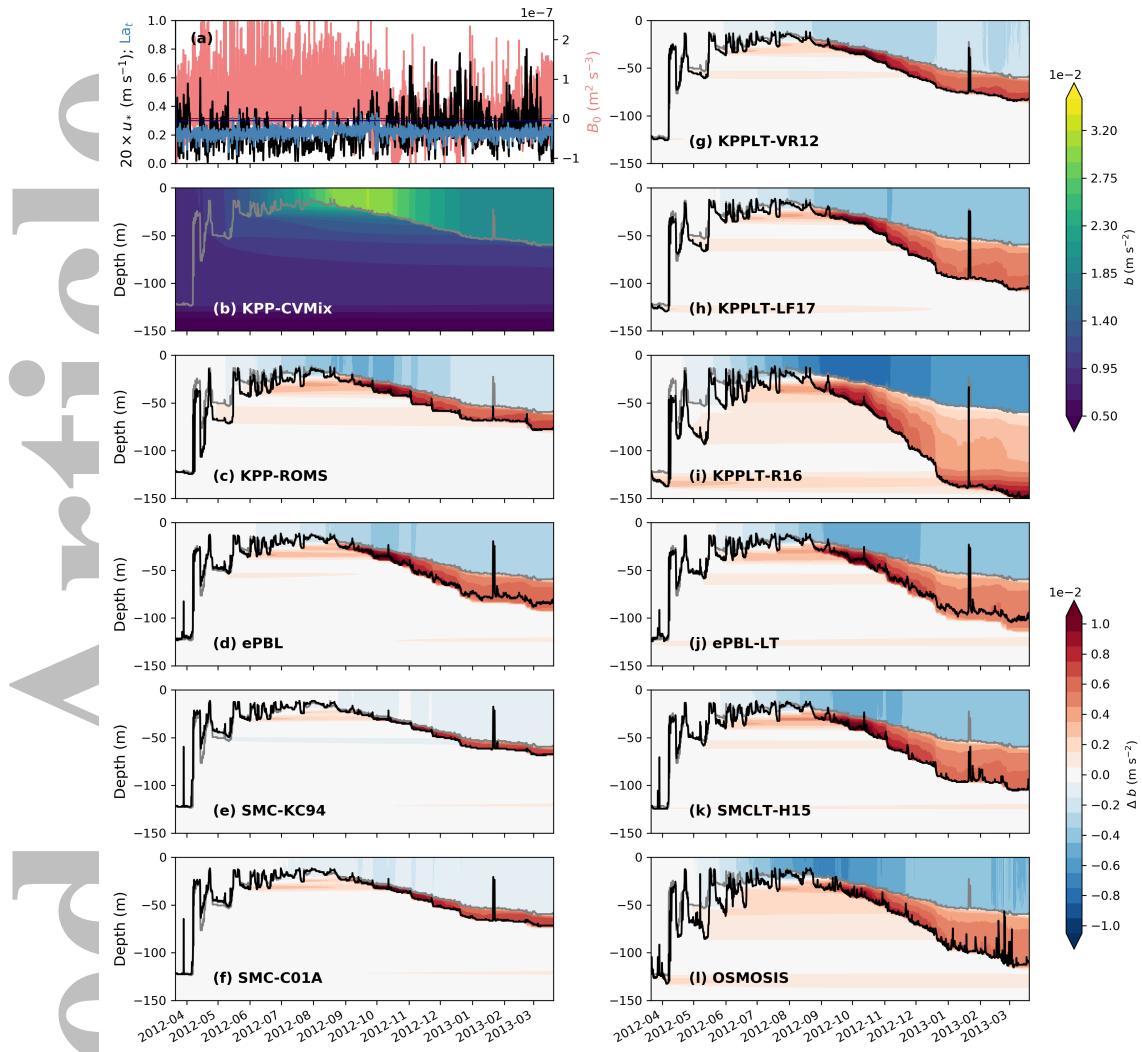
643  
644

0.3 throughout the year (Figures 2a and 3a), suggesting significant influences of Langmuir turbulence on the vertical mixing.



**Figure 3.** Same as Figure 2 but for the OSMOSIS-Spring case. Panel (a) shows water-side surface friction velocity  $u_*$  ( $\text{m s}^{-1}$ ) in black, turbulent Langmuir number  $\text{La}_t$  in blue, and surface buoyancy flux  $B_0$  ( $\text{m}^2 \text{s}^{-3}$ ) in red. Panel (b) shows the simulated buoyancy  $b$  ( $\text{m s}^{-2}$ ) for KPP-CVMix, with MLD defined by the  $0.03 \text{ kg m}^{-3}$  density criterion in gray. Panels (c)-(l) show the differences in simulated buoyancy from KPP-CVMix for all other schemes, with their  $0.03 \text{ kg m}^{-3}$  MLD in black and the KPP-CVMix MLD in gray. For reference, a buoyancy anomaly of  $3.0 \times 10^{-3} \text{ m s}^{-2}$  is the equivalent of  $1.8^\circ\text{C}$  temperature or  $0.40 \text{ g kg}^{-1}$  salinity anomaly. Only the upper 150 m of the domain is shown for clarity.

OCS-Papa represents a full annual cycle of the mixed layer at mid-latitude in the North Pacific. As a result of the much milder winter destabilizing surface buoyancy flux, the seasonal variations of mixed layer depth at this site is much weaker than that at the OSMOSIS site, with a summer mixed layer depth of about 20 m and a winter mixed layer depth of about 100 m. However, the wind is strong in winter and surface gravity waves are intense with a turbulent Langmuir number of around 0.2 (Figure 4a). Therefore, considering the much shallower mixed layer depth than at the OSMOSIS site, we expect much greater influences of Langmuir turbulence on the vertical mixing at the OCS-Papa site.



**Figure 4.** Same as Figure 2 but for the OCS-Papa case. Panel (a) shows water-side surface friction velocity  $u_*$  ( $\text{m s}^{-1}$ ) in black, turbulent Langmuir number  $\text{La}_t$  in blue, and surface buoyancy flux  $B_0$  ( $\text{m}^2 \text{s}^{-3}$ ) in red. Panel (b) shows the simulated buoyancy  $b$  ( $\text{m s}^{-2}$ ) for KPP-CVMix, with MLD defined by the  $0.03 \text{ kg m}^{-3}$  density criterion in gray. Panels (c)-(l) show the differences in simulated buoyancy from KPP-CVMix for all other schemes, with their  $0.03 \text{ kg m}^{-3}$  MLD in black and the KPP-CVMix MLD in gray. For reference, a buoyancy anomaly of  $3.0 \times 10^{-3} \text{ m s}^{-2}$  is the equivalent of  $1.8^\circ\text{C}$  temperature or  $0.40 \text{ g kg}^{-1}$  salinity anomaly. Only the upper 150 m of the domain is shown for clarity.

653 Figures 2-4 indicate that the different schemes give quite distinct results. Signifi-  
 654 cant discrepancies are observed both among the Langmuir schemes and among the non-  
 655 Langmuir schemes, with substantial overlaps between the two groups.

656 A closer examination of these figures show that most of the discrepancies appear to  
 657 be in the mixed layer deepening phase, e.g., evident in Figure 2 and Figure 4 (September  
 658 to January). Also note the discrepancies among schemes in sporadic mixed layer deep-  
 659 ening events shown by the buoyancy anomaly extending in time in Figures 3 and 4. In  
 660 particular, of the five non-Langmuir schemes, ePBL gives the most rapid deepening of

661 mixed layer under destabilizing surface conditions, whereas KPP-CVMix gives the weakest,  
 662 with the result of SMC-C01A closer to ePBL and SMC-KC94 closer to KPP-CVMix.  
 663 Note, however, that the differences between the results of SMC-C01A and SMC-KC94  
 664 are relatively small in all three cases. KPP-ROMS seems to agree with KPP-CVMix in  
 665 OSMOSIS-Winter case where mixed layer depth is relatively deep and destabilizing sur-  
 666 face buoyancy flux is strong, but less so in the other two cases where mixed layer depth  
 667 is relatively shallow, destabilizing surface buoyancy flux is relatively weak and wind is  
 668 strong. Sensitivity tests show that the disagreements between the two versions of KPP  
 669 decreases with increased damping of the horizontal velocity, mostly resulting from the  
 670 changes in KPP-ROMS (not shown). This result suggests that KPP-ROMS is more suscep-  
 671 tible to strong velocity shear at the base of the OSBL associated with inertial oscillations,  
 672 which may also contribute to its larger sensitivity to vertical resolution than KPP-CVMix  
 673 (Appendix B: ). Consistent with *Van Roekel et al.* [2018], we also notice quite strong sen-  
 674 sitivity of KPP-CVMix to numerical details in determining the boundary layer depth from  
 675 a bulk Richardson number criterion (see Appedix A.1 for more details). All these results  
 676 suggest that the algorithm and other numerical details in the formulation and implemen-  
 677 tation of a particular scheme are also important in addition to the underlying physics.

678 Of the six Langmuir schemes, KPPLT-R16 gives the most rapid deepening of mixed  
 679 layer under destabilizing surface conditions. This behavior is expected as KPPLT-R16  
 680 was targeted to study strong wind and waves regimes under hurricane conditions with-  
 681 out an explicit constraint on effects of destabilizing surface buoyancy flux [*Reichl et al.*,  
 682 2016] and will likely over-predict the entrainment by convection [see also *Li and Fox-*  
 683 *Kemper*, 2017]. KPPLT-VR12 gives the weakest mixed layer deepening, especially when  
 684 the destabilizing surface buoyancy flux is strong. The results of KPPLT-LF17, ePBL-LT  
 685 and SMCLT-H15 are similar in most respects. However, the net effects of Langmuir tur-  
 686 bulence, as reflected by comparing each Langmuir scheme to their non-Langmuir coun-  
 687 terpart, appear to agree better in the OCS-Papa and OSMOSIS-Spring cases but less so in  
 688 the OSMOSIS-Winter case. The OSMOSIS scheme behaves similarly to other Langmuir  
 689 schemes in mixed layer deepening events, but differs significantly in mixed layer shoal-  
 690 ing and restratification events. This is probably because it solves prognostic equations for  
 691 boundary layer depth under both deepening and shoaling conditions (see more details in  
 692 Appendix A.3).

### 693 4.3 Test Case 2: A Global Perspective

694 In this section we present results of Test Case 2, which compares different schemes  
 695 from a global perspective. Unlike the previous section, here we use the monthly mean  
 696 mixed layer depth (MLD) defined by the  $0.03 \text{ kg m}^{-3}$  density criterion [*de Boyer Mon-*  
 697 *tégut et al.*, 2004] as a diagnostic quantity to quantify the different behaviors of all the 11  
 698 schemes. Time series of MLD for each one-month simulation is first diagnosed from the  
 699 density field, then averaged to get the monthly mean. Note that due to the monthly aver-  
 700 age and the identical initial conditions for all schemes at the beginning of every month,  
 701 this measure tends to underestimate the differences among difference schemes. Yet sig-  
 702 nificant discrepancies among schemes are found by this measure. This measure is also  
 703 directly relevant to the application in climate models. Other diagnostics, such as alterna-  
 704 tive definitions of MLD, the rate of change in potential energy and sea surface tempera-  
 705 ture (SST), were also examined and they generally tell the same story, though the results  
 706 are quantitatively different, reflecting the complex nature of the turbulent mixing problem  
 707 that is not fully captured by the MLD. While we acknowledge the necessity of using mul-  
 708 tiple diagnostics to fully describe the different behaviors of these schemes, for brevity full  
 709 documentation of all diagnostics is left out, though the reader is encouraged to access the  
 710 simulation code and full suite of diagnostic results online: [github.com/qingli411/gotm](https://github.com/qingli411/gotm) and  
 711 [github.com/qingli411/gotmwork](https://github.com/qingli411/gotmwork).

712      **4.3.1 Comparisons by Region and Season**

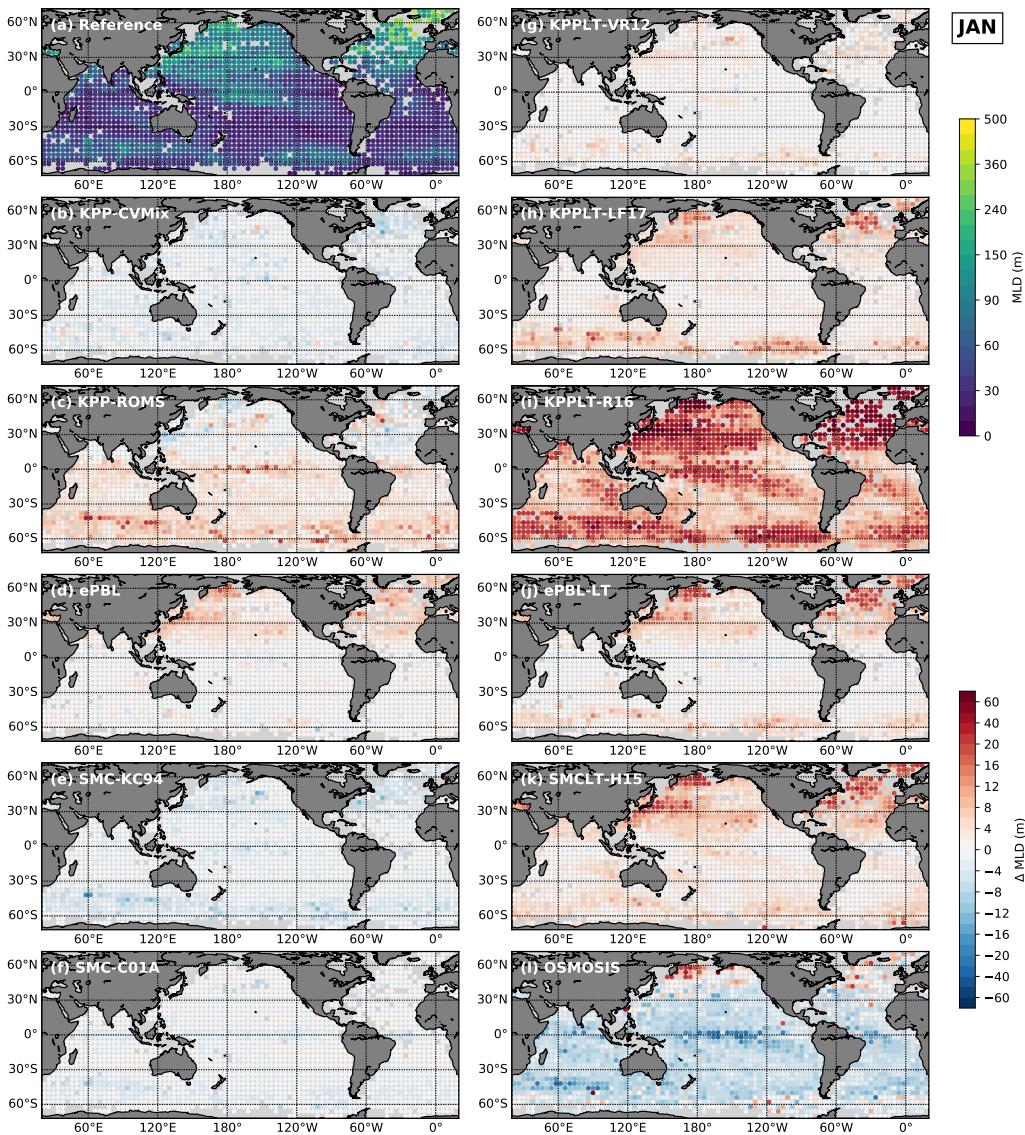
713      Figure 5 and Figure 6 show maps of monthly mean MLD simulated by the different  
 714      schemes in January and July. In the absence of “truth”, MLD of all schemes are compared  
 715      with the mean MLD of the five non-Langmuir schemes (Panel (a)), and this mean is taken  
 716      as the *reference MLD* throughout sections 4.3.1 and 4.3.2 unless otherwise noted. The  
 717      anomaly from this reference MLD for each scheme are shown in other panels. In each  
 718      panel, as well as those in Figure 10 to be introduced later, results are shown at locations  
 719      where reasonable Argo profiles are found as initial conditions (section 3.2) and the sim-  
 720      ulated MLD is within the 500 m domain throughout the month. No explicit sea-ice mask  
 721      was applied. But in practice the above two conditions preclude locations at high-latitude  
 722      where sea-ice tends to grow.

723      As in the ocean station comparison, significant discrepancies are seen among the  
 724      Langmuir schemes, as well as among the non-Langmuir schemes. Of the five non-Langmuir  
 725      schemes, SMC-C01A is closest to the reference MLD. ePBL gives systematically deeper  
 726      MLD and SMC-KC94 systematically shallower, both with the differences from the mean  
 727      approximately proportional to the mean MLD, i.e., greater differences where the mean  
 728      MLD is deep. Interestingly, the two non-Langmuir versions of KPP are at the opposite  
 729      ends of the distribution, with KPP-CVMix having the shallower MLD and KPP-ROMS  
 730      having deeper. The differences between these two appear smaller in winter hemispheres  
 731      but substantial in summer hemispheres. This result suggests that both the physical and nu-  
 732      matical formulations of each scheme have a nontrivial impact on a global scale.

733      Among the Langmuir schemes, KPPLT-LF17, ePBL-LT and SMCLT-H15 give more  
 734      similar results than the other three Langmuir schemes, showing significantly deeper MLD  
 735      than the reference MLD in the Northern Hemisphere extratropical regions in winter and  
 736      Southern Ocean in both winter and summer. KPPLT-VR12 gives deeper MLD than its  
 737      non-Langmuir turbulence counterpart, KPP-CVMix, but shallower than some other non-  
 738      Langmuir schemes, such as ePBL. As expected, KPPLT-R16 predicts the deepest MLD of  
 739      all schemes, especially in winter when strong convection occurs. OSMOSIS results differ  
 740      from all other schemes, with the simulated MLD deeper in the high-latitudes of the winter  
 741      hemisphere as consistent with Test Case 1, but shallower than the reference non-Langmuir  
 742      MLD in the mid- to low-latitudes.

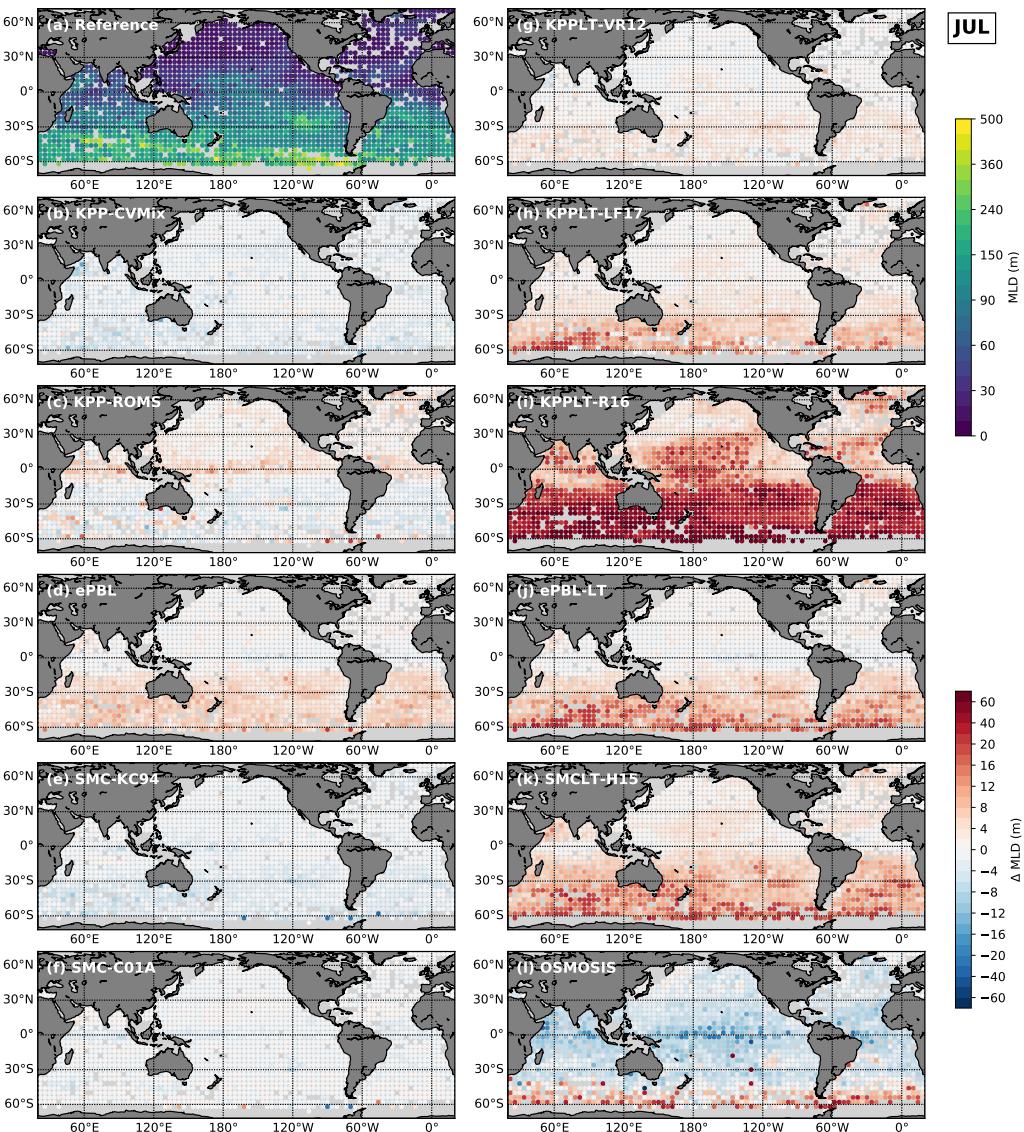
743      These discrepancies among schemes are consistent throughout the annual cycle, as  
 744      shown by the zonal mean MLD for each month in Figure 7 and their root-mean-square  
 745      (RMS) differences from the reference MLD in Figure 8. KPPLT-R16 gives the deepest  
 746      MLD at all latitudes for all months. OSMOSIS gives the shallowest MLD except in the  
 747      high latitudes, especially in winter. All other schemes seem to agree in simulated MLD in  
 748      low latitudes for all months and in Northern Hemisphere mid- to high-latitudes in summer,  
 749      while they disagree significantly in the Southern Ocean all year and Northern Hemisphere  
 750      mid- to high-latitudes from October to May. Interestingly, KPP-ROMS gives up to 10 m  
 751      deeper MLD than other schemes in this group near the equator (Figure 7). This behavior  
 752      might be related to the strong vertical shear of near surface current in the equatorial Pa-  
 753      cific picked up by the integral of the shear magnitude in KPP-ROMS (see equation A.3).

754      A summary of the differences of simulated MLD among all the 11 schemes is pre-  
 755      sented in Figure 9. Panel (a) shows the distribution of the differences of monthly mean  
 756      MLD in each scheme from the reference MLD, whereas panel (c) shows the distribution  
 757      of the ratios of these two. Again, significant differences are seen among both the Lang-  
 758      muir schemes and non-Langmuir schemes. Generally, we find greater spread in the Lang-  
 759      muir schemes than in the non-Langmuir schemes (see the distributions shown by the box  
 760      and whisker in light and dark gray), with significant overlap. In addition, different cate-  
 761      gories of schemes show slightly different effects of Langmuir turbulence, as seen by com-  
 762      paring boxes with similar colors.



**Figure 5.** Comparison of January 2009 mean mixed layer depth (MLD) simulated by different schemes with atmospheric forcing from JRA55-do. Panel (a) shows the map of the reference MLD taken as the mean of the five non-Langmuir schemes (upper color bar in m). Panels (b)-(l) show the differences of MLD from that mean for each of the 11 schemes (lower color bar in m; positive for deeper MLD). Note that the color scale in both color bars are nonlinear.

Panel (b) and (d) of Figure 9 highlight three Langmuir schemes of different categories, KPPLT-LF17, ePBL-LT and SMCLT-H15, by comparing them with their non-Langmuir turbulence counterparts, KPP-CVMix, ePBL and SMC-KC94, respectively. On average the Langmuir schemes predict 6% deeper MLD than their non-Langmuir turbulence counterparts (-2% to 14% for 90% confidence range). Although the simulated MLD by SMCLT-H15 and ePBL-LT seem to agree (Figures 5, 6, 9a,c), results of SMC-94 and ePBL disagree remarkably and are among the shallowest and deepest in the non-Langmuir turbulence results. Therefore, the strongest effects of Langmuir turbulence are seen in SMCLT-H15 (around 10%) and weakest in ePBL-LT (around 2%) as compared with their non-Langmuir counterparts. Maps of the percentage change in the simulated

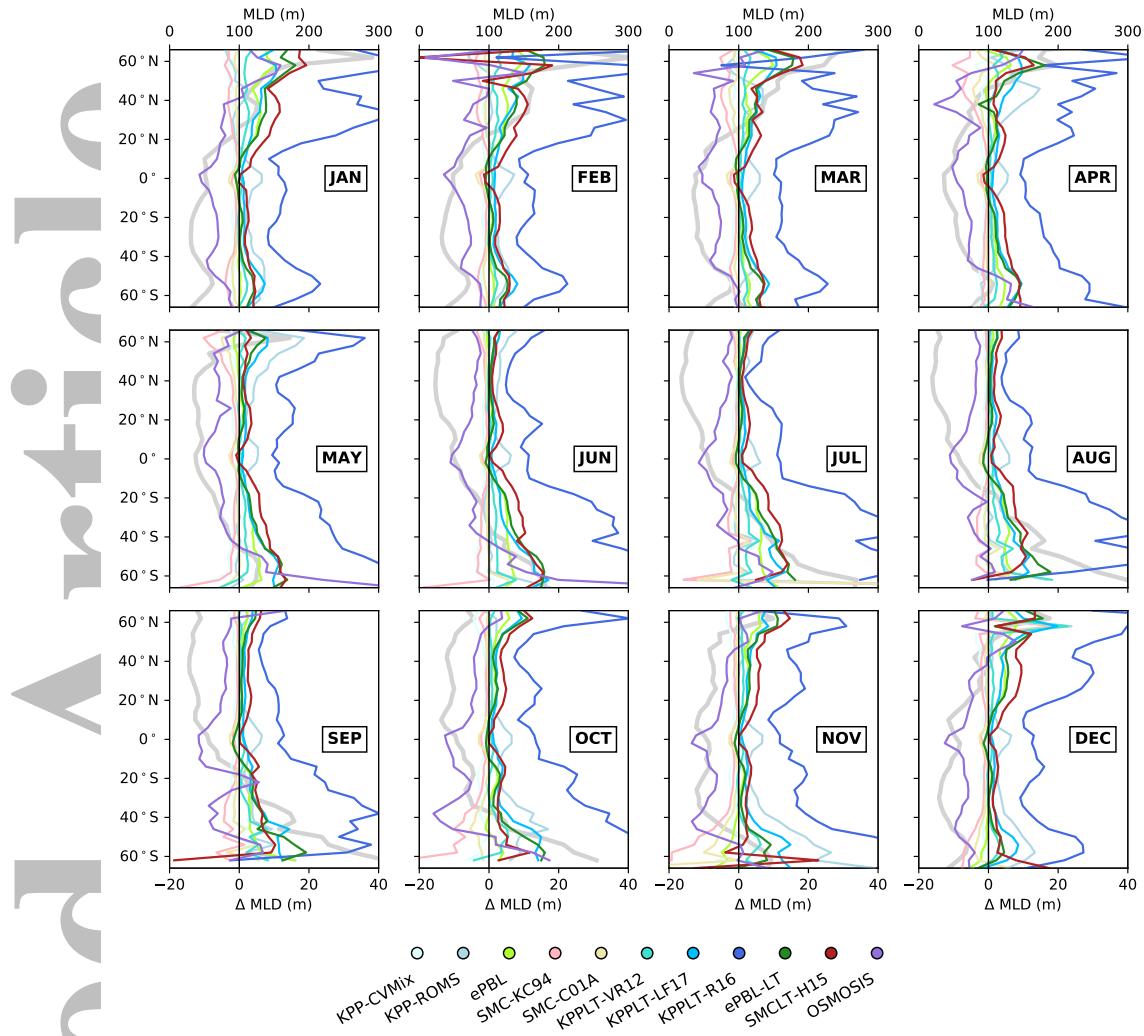


**Figure 6.** Same as Figure 5, but comparison of July 2008 mean mixed layer depth (MLD). (a) Mean of five non-Langmuir schemes, (b)-(l) show the differences from mean in (a) for each of the 11 schemes.

January and July mean MLD in KPPLT-LF17, ePBL-LT and SMCLT-H15, as compared to KPP-CVMix, ePBL and SMC-94, respectively, are shown in Figure 10. Consistent with previous studies [e.g., Belcher *et al.*, 2012; D'Asaro *et al.*, 2014; Li *et al.*, 2016], the effects of Langmuir turbulence appear to be most important in the extratropical regions, especially the Southern Ocean in austral summer.

#### 4.3.2 Comparison by Forcing Parameters

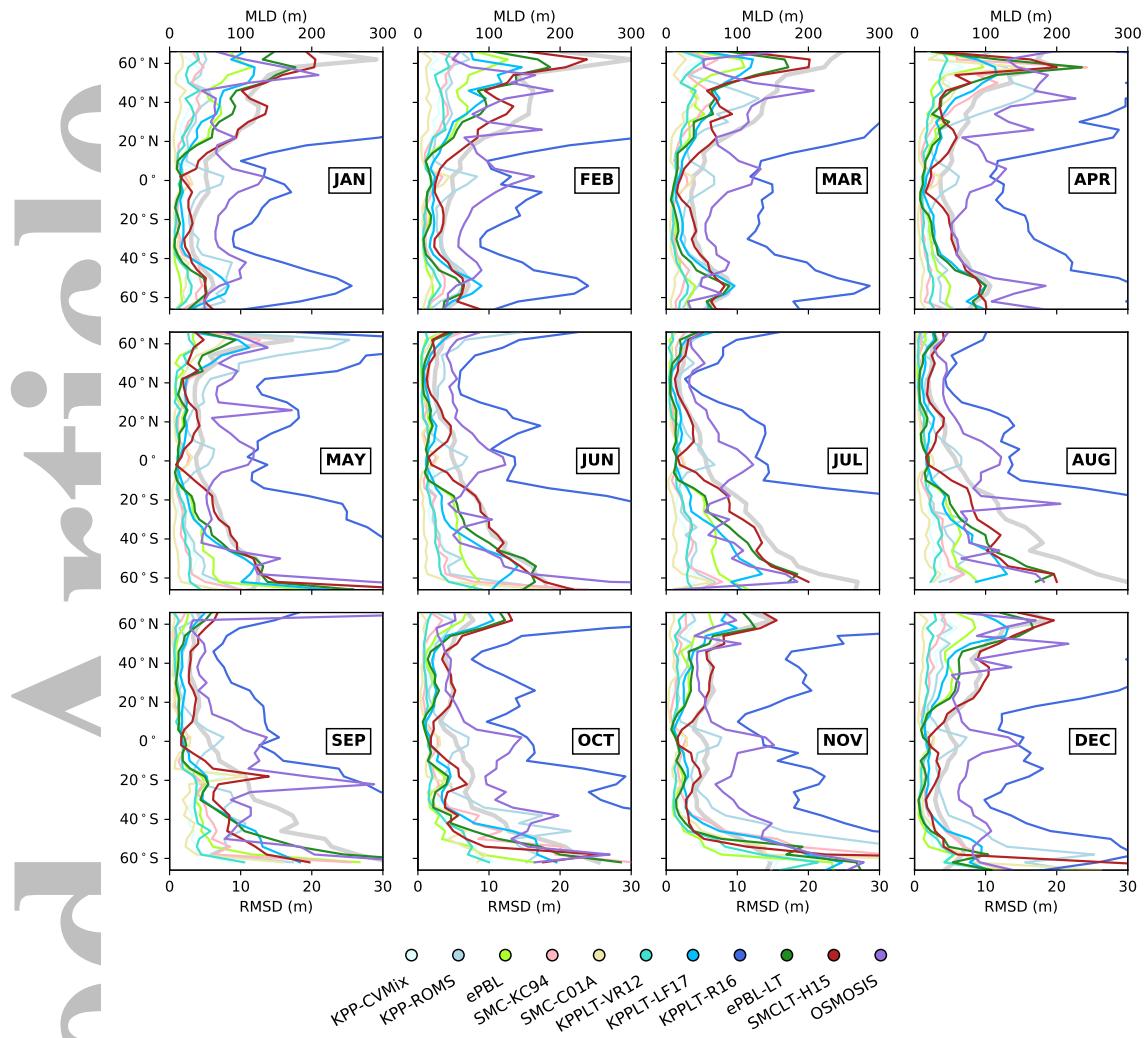
To provide more insight into the differences among schemes, the above results are further sorted by the surface forcing regimes described in section 4.1. The surface forcing conditions are categorized into seven regimes in the  $\text{La}_t - h/L_L$  parameter space of Belcher *et al.* [2012], depending on the relative importance of wind-driven shear turbulence, Langmuir turbulence and convective turbulence as measured by the turbulent dis-



**Figure 7.** Comparison of zonal mean mixed layer depth (MLD; m) for each month. Thick line in gray shows the zonal mean MLD averaged over all the non-Langmuir cases (reference MLD; upper x-axis). Thin lines in color show the difference of zonal mean MLD from the reference ( $\Delta$  MLD; m) for each of the 11 schemes (lower x-axis).

784 sipation scaling in equation (11). As shown by colors in Figure 11, the seven forcing  
 785 regimes include three single forcing regimes for shear turbulence (S), Langmuir turbulence  
 786 (L) and convective turbulence (C), respectively, where the contributions from the other  
 787 two kinds of turbulence are both less than 25%; three combined forcing regimes (SL, SC,  
 788 LC), where both kinds of turbulence contribute more than 25% while the other contributes  
 789 less than 25%; and one mixed forcing regime (SLC), where contributions of three kinds  
 790 of turbulence are all greater than 25%. It is clear from the overlaid probability distribution  
 791 that convection, Langmuir turbulence and their combination are the dominant regimes in  
 792 the 3-hourly JRA55-do data.

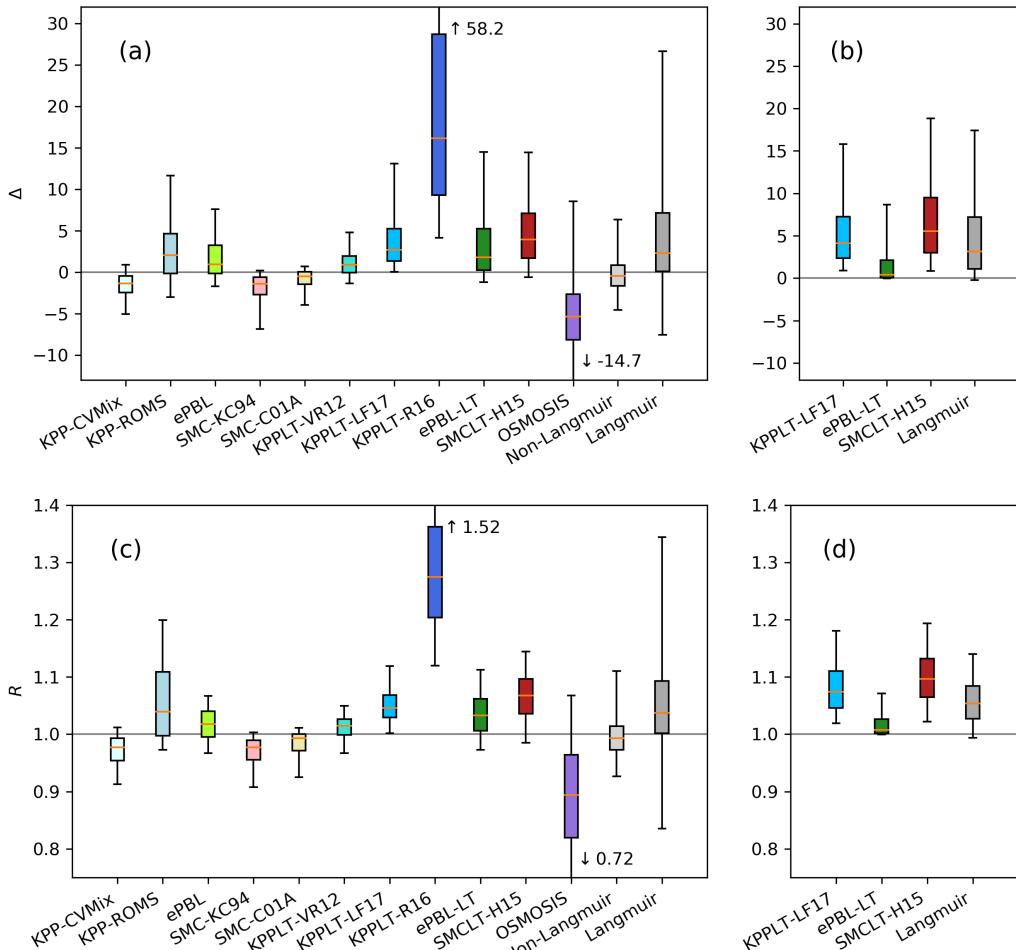
793 We compute the contribution fractions of each kind of turbulence at each output  
 794 time step (3-hourly) at each location when the surface buoyancy flux is destabilizing ( $B_0 <$   
 795 0), and take the monthly average. The mean forcing regime at each location for each month  
 796 was then determined based on the above categorization. The resulting maps of the mean



**Figure 8.** Same as Figure 7, but the thin lines show the root-mean-square differences (RMSD; m; lower x-axis) from the reference MLD (gray thick line, upper x-axis).

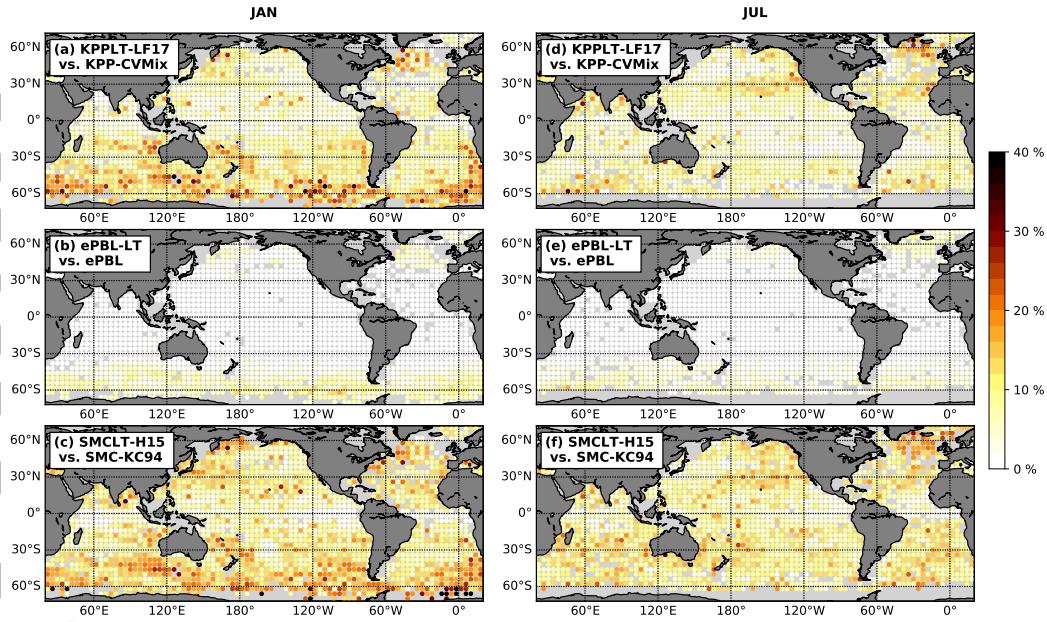
forcing regime for each month are presented in Figure 12. If the number of occurrence of  $B_0 < 0$  is less than 1/3 of the total number of snapshots over a month, the mean forcing regime is marked as not applicable (NA). Since the above forcing regime categorization only covers instances when  $B_0 < 0$ , additional information of the surface stability condition is provided in the map by marking locations with stabilizing mean surface buoyancy flux ( $B_0 > 0$ ) by black crosses in Figure 12. To distinguish from the aforementioned seven forcing regimes under destabilizing surface conditions, a star is added to the name of the forcing regime when referring to these cases (e.g., L\* denotes L regime with  $\overline{B_0} > 0$ ).

Note that this categorization of surface forcing regimes is built on the monthly mean contribution fractions of each kind of turbulence to turbulent mixing and has surface stability conditions integrated in. A cross in Figure 12 does not suggest the dominance of surface stabilizing conditions in a region. Instead, it only denotes that over a month a region spends more time with or experiences stronger stabilizing forcing than regions without a cross. Such differentiation provides a way to show the integrated effect of stabilizing surface conditions on inhibiting turbulent mixing, which otherwise is not directly com-

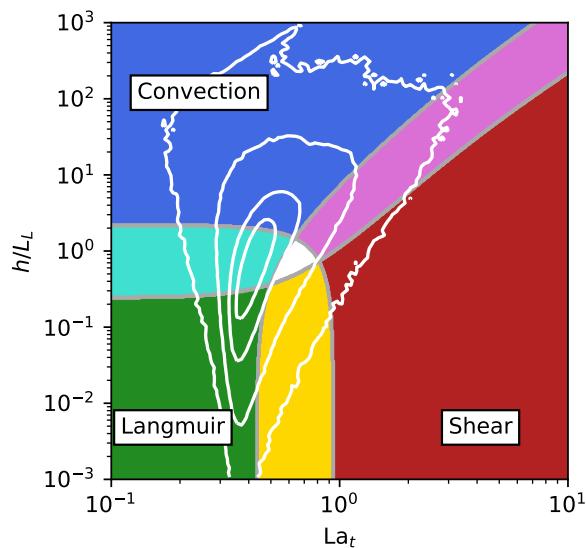


**Figure 9.** (a) Distribution of the difference of monthly mean mixed layer depth (MLD; m) from the reference MLD,  $\Delta = \text{MLD} - \text{MLD}_{\text{ref}}$ , for each scheme (colored), and all non-Langmuir turbulence and Langmuir schemes together, respectively (light and dark gray). The box marks the upper and lower quartiles, whiskers mark the 95th and 5th percentiles. (b) Same as (a), but for the difference of Langmuir schemes from their non-Langmuir counterparts rather than from the overall reference,  $\Delta = \text{MLD}_{\text{LT}} - \text{MLD}_{\text{nLT}}$ , for KPPLT-LF17, ePBL-LT, SMCLT-H15, and the mean of six Langmuir schemes. (c) Same as (a), but for the ratio of monthly mean MLD to the reference MLD,  $R = \text{MLD} / \text{MLD}_{\text{ref}}$ . (d) Same as (c), but for Langmuir schemes normalized by their non-Langmuir counterparts rather than the overall reference,  $R = \text{MLD}_{\text{LT}} / \text{MLD}_{\text{nLT}}$ , for KPPLT-LF17, ePBL-LT, SMCLT-H15, and the mean over all Langmuir schemes.

parable to the effects of the three kinds of turbulence in driving the mixing by this measure (according to their contributions to the turbulent dissipation). Hence the emphasis of this categorization is still on the mechanisms that drive turbulent mixing under destabilizing surface conditions. The physical meaning of the regime C\* is that convection dominates vertical mixing during times when the surface condition is destabilizing, yet such destabilizing surface condition is weak and short during a month. One example is some low-latitude regions in summer with weak wind and waves, where nocturnal convection drives the vertical mixing, but strong diurnal heating may result in  $\overline{B_0} > 0$  over a month. Therefore, this categorization of surface forcing regimes should be distinguished from the

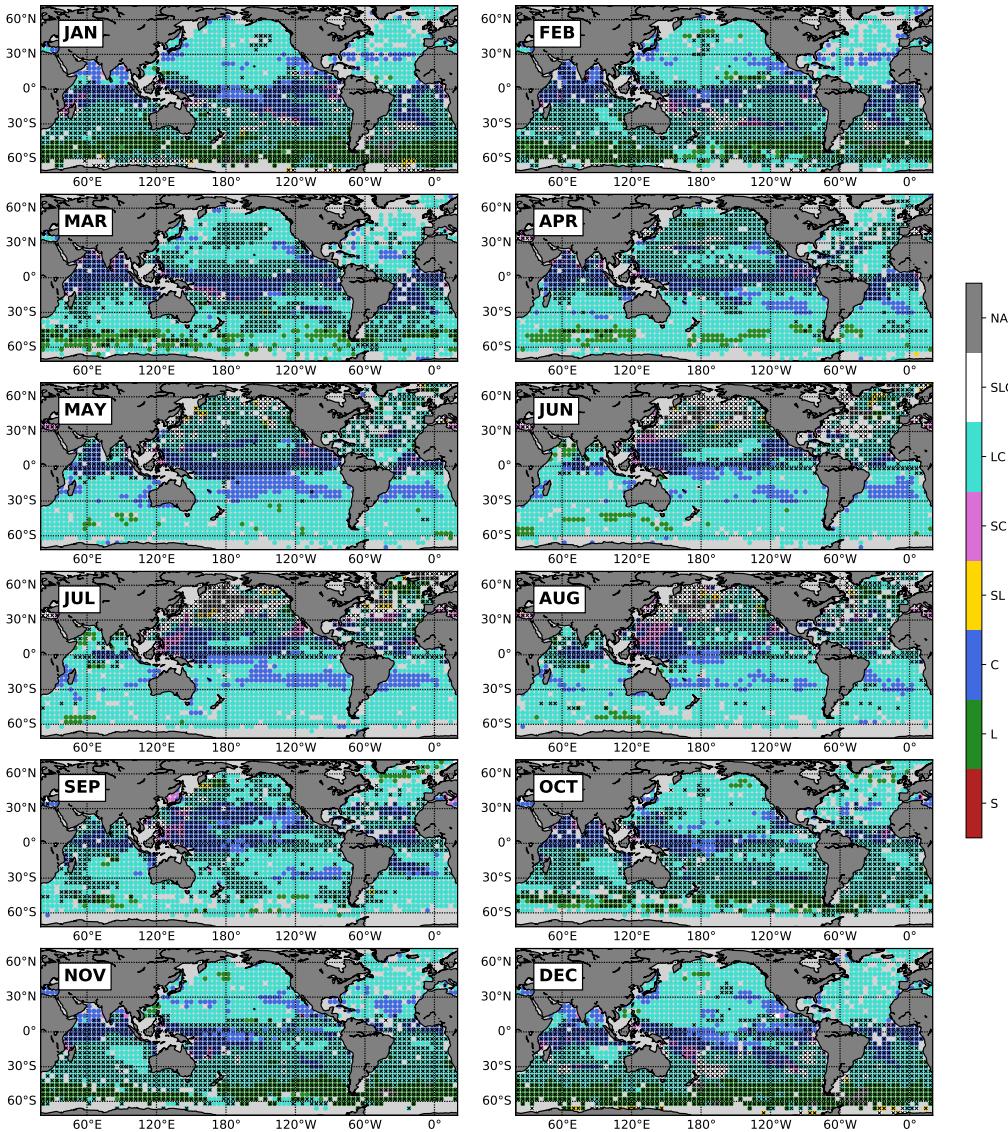


**Figure 10.** Maps of percentage change in simulated (a-c) January and (d-f) July mean mixed layer depth when including Langmuir turbulence, estimated by comparing (a, d) KPPLT-LF17 and KPP-CVMix; (b, e) ePBL-LT and ePBL; and (c, f) SMCLT-H15 and SMC-KC94.



**Figure 11.** The same  $\text{La}_t - h/L_L$  parameter space as in Figure 1a divided into seven regimes according to the contribution of each of the three forcing mechanisms to the total turbulent dissipation (Red: S; Green: L; Blue: C; Magenta: SC; Yellow: SL; Cyan: LC; White: SLC). See the text for the criterion used for the categorization. Overlaid white contours show the same probability distribution from JRA55-do forcing as in Figure 1a.

821 distributions in regime diagrams of Figures 1 and 11, which are estimated from 3-hourly  
 822 statistics.



**Figure 12.** Maps of surface forcing regimes for each month, color coded according to the regime diagram in Figure 11, estimated from the JRA55-do forced GOTM5 simulations. Seven forcing regimes, shear turbulence (S), Langmuir turbulence (L), convective turbulence (C), combined shear and Langmuir turbulence (SL), combined shear and convective turbulence (SC), combined Langmuir and convective turbulence (LC), and a combination of all (SLC), are color coded according to the color bar. Locations without enough data for the categorization are shown in gray (NA). See the text for the criterion of each forcing regime. Locations with stabilizing mean surface buoyancy flux are marked by black  $\times$ .

Figure 12 shows that most areas of the global ocean are in either LC or LC\* regimes. In total these two regimes account for about 67% of all surface forcing conditions (see the gray bars in Figure 13). During austral summer, large areas of the Southern Ocean are in either L or L\* regimes (account for 10%), supporting the hypothesis that Langmuir turbulence is an important part of shallow mixed layer biases there [Belcher *et al.*, 2012]. The C and C\* regimes dominate a significant portion of the low-latitudes all year round (account for 17%). In total these six regimes account for over 94% of the surface forcing

830 conditions in the JRA55-do forced runs. We therefore examine differences of simulated  
 831 MLD among schemes and the relative effects of Langmuir turbulence by sorting into these  
 832 six regime categories in Figure 13 and Figure 14.

833 At first glance, the distributions of the MLD ratio,  $R$ , in all of the six regimes are  
 834 very similar to that in Figure 11c. However, the differences among schemes differ in mag-  
 835 nitude between regimes. First, the differences among all the schemes, as well as the dif-  
 836 ferences between the non-Langmuir schemes and Langmuir schemes, decrease as moving  
 837 from L to C, or  $L^*$  to  $C^*$ , regimes, thus consensus is weaker in Langmuir conditions than  
 838 in convective conditions. This behavior is expected since many of the Langmuir schemes  
 839 converge back to their non-Langmuir turbulence counterpart in the convection limit, when  
 840 the relative effect of Langmuir turbulence becomes small.

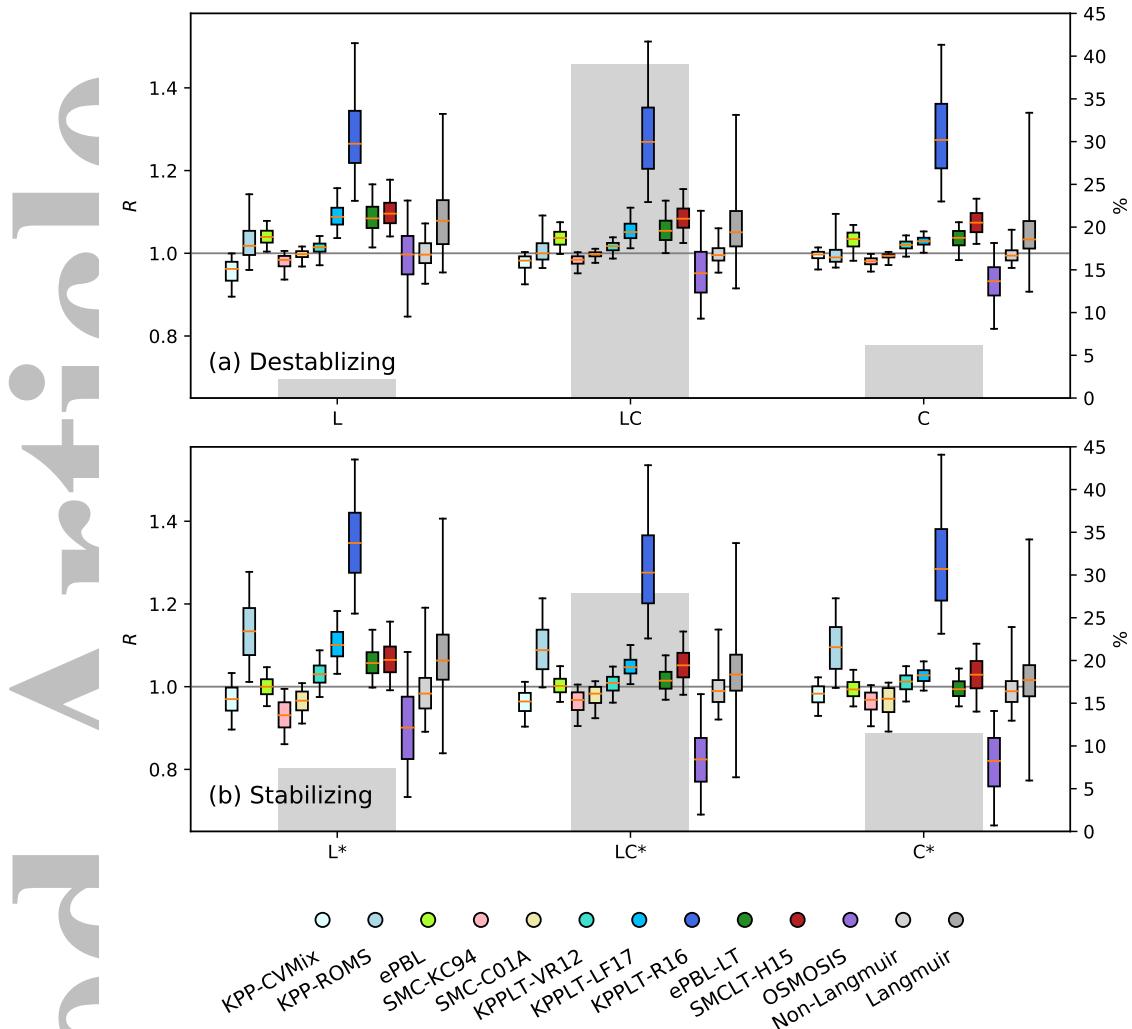
841 The significant differences among the schemes in L and  $L^*$  regimes highlight the  
 842 necessity of better constraints of Langmuir turbulence effects from theory, observations,  
 843 and/or numerical process models. Interestingly, the larger difference under L versus C or  
 844  $L^*$  versus  $C^*$  is also true of the differences among non-Langmuir schemes (to a lesser  
 845 degree), even though wave forcing does not affect the results (note that winds versus buoy-  
 846ancy also varies among these regimes). Second, the differences among all the schemes  
 847 are greater in regimes with stabilizing surface conditions, especially when comparing  $L^*$   
 848 regime to L regime. This result suggests that uncertainties exist in our understanding of  
 849 turbulence under stabilizing surface conditions in both Langmuir and non-Langmuir set-  
 850 tings, associated with the less well-covered parameter space by existing LES studies in  
 851 Figure 1b. This regime has long been a challenge in the atmosphere, because the turbu-  
 852 lence is weak and small and thus hard to simulate [e.g., Beare *et al.*, 2006; Sullivan *et al.*,  
 853 2016], although Pearson *et al.* [2015] note that this difficulty is eased in the presence of  
 854 Langmuir turbulence.

855 As in Figure 9, KPPLT-R16 gives the deepest MLD and OSMOSIS gives the shal-  
 856 lowest. These differences compared to other Langmuir schemes are greatest in C and  $C^*$   
 857 regimes and smaller in L and  $L^*$  regimes. Interestingly, KPP-ROMS exhibits much greater  
 858 sensitivity to stabilizing surface conditions than KPP-CVMix, predicting much deeper  
 859 MLD than the reference MLD under stabilizing surface conditions than under destabilizing  
 860 surface conditions.

861 Consistently, the relative effects of Langmuir turbulence revealed by comparing indi-  
 862 vidual Langmuir schemes to their non-Langmuir counterparts (Figure 14) are most promi-  
 863 nent in L and  $L^*$  regimes, with MLD deepening of 9% on average and 2% to 17% for  
 864 90% confidence range. It is also interesting to note that, in C and  $C^*$  regimes, ePBL-LT  
 865 converges to ePBL as expected; KPPLT-LF17 still gives deeper MLD than KPP-CVMix,  
 866 but significantly less so than in other regimes; whereas SMCLT-H15 shows the least sen-  
 867 sitivity to the transition from L and  $L^*$  regimes to C and  $C^*$  regimes (consistent with Fig-  
 868 ure 10).

#### 869 4.4 Test Case 3: Evaluation Against Idealized LES

870 While the single column simulations used here are not directly comparable to ob-  
 871 servations due to the lack of horizontal processes and vertical advection, they are directly  
 872 comparable to many of the idealized LES results in which horizontally homogeneous tur-  
 873 bulence is assumed. Here we choose two recent idealized LES studies of Langmuir tur-  
 874 bulence, Li and Fox-Kemper [2017] and Reichl *et al.* [2016], to compare the behaviors of dif-  
 875 ferent schemes when simulating the entrainment process under (1) constant surface forcing  
 876 of mixed Langmuir turbulence and convection, and (2) transient surface forcing of typi-  
 877 cal wind and Langmuir turbulence during an idealized hurricane, respectively. Other LES  
 878 studies of Langmuir turbulence under more realistic transient forcing, such as a mixed  
 879 layer deepening event in Kukulka *et al.* [2009], diurnal restratification events in Kukulka  
 880 *et al.* [2013], and at the Southern Ocean Flux Station in Large *et al.* [2019], are also ex-



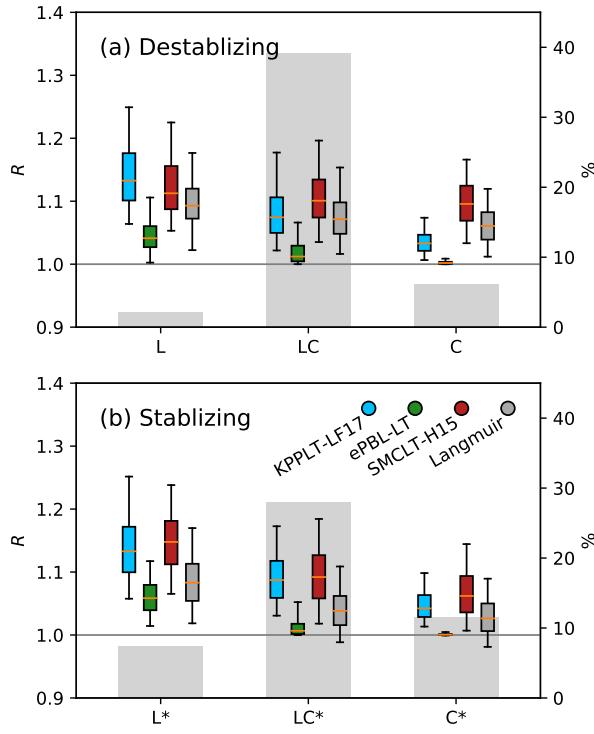
**Figure 13.** Distribution of the ratio of monthly mean MLD to the reference MLD,  $R = \text{MLD}/\text{MLD}_{\text{ref}}$ , for different schemes (same as Figure 9c) in different surface forcing regimes. (a) Langmuir (L), Langmuir + Convection (LC), and Convection (C) regimes, with destabilizing monthly mean surface buoyancy flux. (b) Langmuir (L\*), Langmuir + Convection (LC\*), and Convection (C\*) regimes, with stabilizing monthly mean surface buoyancy flux. See the text for the physical meaning of the regimes with a star and their differences from those without a star. Gray bars show the percentage occurrence of each regime, estimated from the JRA55-do forced GOTM5 simulations.

881 excellent reference cases, but these more complex LES studies require more detail than this  
882 comparison allows.

#### 883 4.4.1 Entrainment

884 This section compares the 11 schemes to LES in simulating the entrainment by  
885 mixed Langmuir turbulence and convection in *Li and Fox-Kemper* [2017]. Under constant  
886 destabilizing surface forcing conditions, the mixed layer keeps deepening as a result of  
887 the entrainment process, eroding the constant stratification below and converting turbu-  
888 lent kinetic energy into potential energy. In a quasi-equilibrium state, the rate of mixed

# Arctic Today Article



**Figure 14.** Same as Figure 9d, but for different surface forcing regimes as described in the caption of Figure 13.

layer deepening reaches a constant, and the potential energy increases approximately linearly with time as a result of the constant energy input from the surface. Thus, the rate of change in potential energy describes the intensity of entrainment and can be used to evaluate different schemes against LES.

For both LES and GOTM5 simulations, the potential energy referenced to the bottom of the simulation domain was first computed from the simulated buoyancy profiles at each output time step. The rate of change in potential energy was then diagnosed from the slope of a linear fit to the potential energy time series over the last inertial period. The rate of change in potential energy in a set of LES of *Li and Fox-Kemper* [2017] under different wind, wave and surface cooling conditions are shown in Figure 15a, plotted against the parameter  $h/L_L$ . As in section 4.1 and 4.3.2, the parameter  $h/L_L$  describes the relative importance of convection and Langmuir turbulence, with greater values indicating greater influences from convection. Note that both Langmuir turbulence and convection enhances entrainment, as shown by the black and gray arrows. However, convection starts to dominate the entrainment when surface cooling is strong.

Figure 15b shows the ratios of the rate of change in potential energy in GOTM5 simulations for each scheme to that of the LES. The closer this ratio is to one, the better the agreement with the LES. Note that smaller denominators (panel (a), e.g., when  $h/L_L \sim 0.1$ ) in these ratios do not necessarily lead to greater deviations from one. We conclude from Figure 15b that different schemes tend to agree with LES in the convection regime (except KPPLT-R16), but diverge from LES in different ways in the Langmuir turbulence regime. In particular, all non-Langmuir schemes significantly underpredict entrainment in the Langmuir turbulence regime, with ePBL being the closest to LES. Of all the Langmuir schemes tested here, ePBL-LT agrees with LES the most. This agreement

913 is expected as ePBL-LT was tuned against the same set of LES to match both the rate of  
 914 change in potential energy and the rate of change in SST [Reichl and Li, 2019].

915 While the scaling of entrainment buoyancy flux in KPPLT-LF17 was also derived  
 916 from the same set of LES, KPPLT-LF17 suffers from the underlying strong sensitivities  
 917 of KPP to numerical details as discussed in *Van Roekel et al.* [2018] and Appendix A.1.  
 918 Therefore, the results of KPPLT-LF17 (as well as KPPLT-VR12) shown in this figure de-  
 919 pend strongly on how well the stratification at the base of the boundary layer is resolved  
 920 and thereby the choice of numerical details in KPP. With the present KPP setup following  
 921 the recommendation of *Van Roekel et al.* [2018], KPPLT-LF17 gives too strong entrain-  
 922 ment as compared to LES. Sensitivity tests show that, with an alternative KPP setup (see  
 923 the discussion in Appendix A.1), KPPLT-LF17 may give weaker entrainment in Langmuir  
 924 turbulence regime [e.g., seen in *Reichl and Li*, 2019]. However, it is important to note  
 925 that the scaling relationships used in both KPPLT-LF17 and KPPLT-VR12 were derived  
 926 without any assumptions of such numerical details in KPP. The entrainment is generally  
 927 too strong in KPPLT-R16 in both convection and Langmuir turbulence regimes, while too  
 928 weak in KPPLT-VR12, SMCLT and OSMOSIS in the Langmuir turbulence regime.

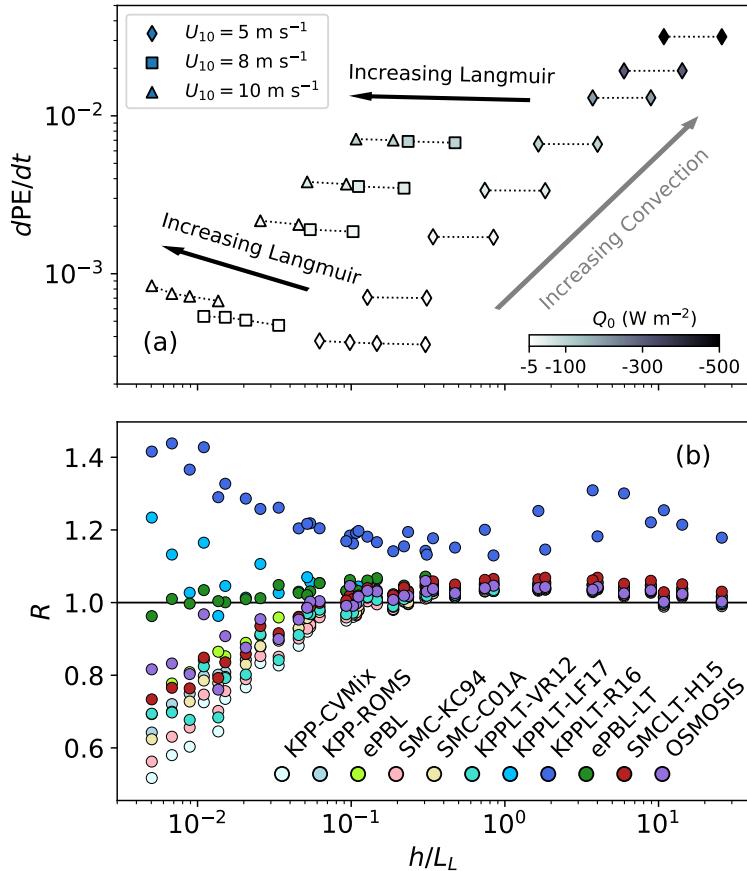
#### 929 **4.4.2 Idealized Hurricane**

930 This section compares the 11 schemes to LES in an idealized hurricane case of *Re-*  
 931 *ichl et al.* [2016]. The most significant feature in this case is rapid mixed layer deepen-  
 932 ing under strong transient hurricane wind (up to  $65 \text{ m s}^{-1}$ ) and associated strong waves  
 933 ( $\text{La}_t \rightarrow 0.3$  during peak wind). The time series of the water-side friction velocity  $u_*$ , wind  
 934 direction and turbulent Langmuir number  $\text{La}_t$  are shown in Figure 16a. The time evolu-  
 935 tion of the temperature profile is shown in Figure 16b. Note that the surface buoyancy  
 936 flux was set to nearly neutral ( $5 \text{ W m}^{-2}$  surface cooling) throughout the run. So unlike the  
 937 previous section, the convection does not play a significant role in driving the entrainment  
 938 here.

939 Figures 16c,d show the time series of simulated SST and change in potential en-  
 940 ergy in LES (gray dashed lines) and GOTM5 simulations with different schemes (shown  
 941 by the differences from the LES, solid lines in color), in which significant spread is seen  
 942 for both. Except KPP-ROMS, all schemes under-predict the deepening of the mixed layer  
 943 during the hurricane, resulting in warmer SST and smaller change of potential energy  
 944 than LES after the hurricane. But the Langmuir schemes outperform the non-Langmuir  
 945 turbulence ones. As expected, KPPLT-R16 performs the best in correctly simulating the  
 946 SST, as it is tuned against a set of idealized hurricane LES of *Reichl et al.* [2016] includ-  
 947 ing this one, though slightly under-predicts the change in potential energy. KPPLT-LF17  
 948 also correctly predicts the SST after the hurricane, but is slightly too warm in SST during  
 949 the hurricane, and also under-predicts the change in potential energy. ePBL-LT performs  
 950 better in predicting the change in potential energy after the hurricane than KPPLT-R16  
 951 and KPPLT-LF17, but performs less well during the mixed layer deepening phase, and  
 952 also over-predicts the SST. SMCLT-H15 outperforms its non-Langmuir turbulence coun-  
 953 terpart, SMC-KC94, but it is farther away from the LES solution than the non-Langmuir  
 954 LES case. The OSMOSIS scheme performs poorly in this case, giving overly warm SST  
 955 and too little change in potential energy, probably due to that the effects of resolved shear  
 956 are not currently represented in this implementation of the OSMOSIS scheme. Note that  
 957 these biases of the simulated SST and change in potential energy are almost as big as the  
 958 overall change during the simulation and larger than the differences between LES with and  
 959 without the wave forcing driving Langmuir turbulence (dashed line in black).

960 Overall, the effects of wave forcing are apparent, affecting both the schemes and the  
 961 LES. Some schemes perform badly in this test (KPP-ROMS, OSMOSIS, KPPLT-VR12,  
 962 and SMCLT-H15). KPPLT-LF17, ePBL-LT, and KPPLT-R16 schemes are able to capture

# Accepted Article



**Figure 15.** The rate of change in potential energy in the entrainment cases of Li and Fox-Kemper [2017] in LES and schemes. (a) The LES rate of change in potential energy ( $\text{J m}^{-2} \text{s}^{-1}$ ), versus the parameter  $h/L_L$ . Different LES cases with the same surface cooling (color) and wind forcing (symbols) are grouped together and connected by dotted lines. In each group, surface wave forcing increases moving from the right to the left. Arrows in black and gray show the directions of increasing Langmuir turbulence and increasing convection, respectively, both enhance the entrainment. (b) The ratio of the rate of change in potential energy for each scheme to that in LES (i.e., the vertical axis shown in panel (a)) versus the parameter  $h/L_L$  from each comparison.

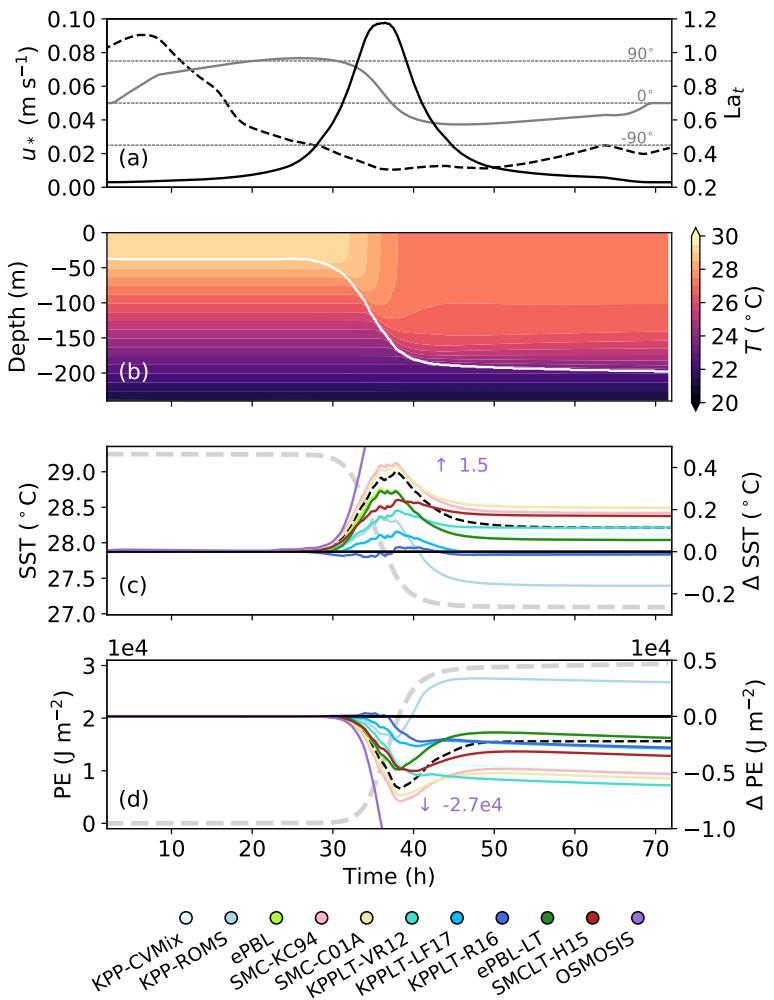
the additional Langmuir deepening, while the remaining non-Langmuir schemes (i.e., except KPP-ROMS) agree with the non-Langmuir LES.

## 5 Discussion

### 5.1 Overall Comments on the Langmuir Turbulence Schemes

The enhancement factor to the turbulent velocity scale in KPPLT-VR12 is probably the most straightforward way to account for some of Langmuir turbulence effects in KPP. This enhancement factor represents the bulk effect of Langmuir turbulence on enhancing the vertical velocity variance in the observations [D'Asaro, 2001; Tseng and D'Asaro, 2004; D'Asaro *et al.*, 2014] and LES [McWilliams *et al.*, 1997; Grant and Belcher, 2009]. Refining the enhancement factor formula to account for additional properties of Lang-

Accepted Article



**Figure 16.** Comparison of different schemes to LES in an idealized hurricane case of Reichl *et al.* [2016]. (a) Time series of friction velocity (solid line in black, vertical axis on the left), turbulent Langmuir number (dashed line in black, vertical axis on the right) and wind direction (solid line in gray, angle counter-clockwise from the direction of the hurricane path). (b) Hovmöller diagram of the simulated temperature in LES. White line marks the boundary layer depth defined by the depth at which the stratification reaches its maximum. (c) Time series of simulated sea surface temperature (SST) in LES (thick dashed line in gray, vertical axis on the left) and the differences of simulated SST in GOTM with different schemes from the LES (thin solid lines in color, vertical axis on the right). Dashed line in black shows the difference simulated by the same LES but without wave forcing of Langmuir turbulence. (d) Same as (c) but for vertically integrated potential energy, referenced to the bottom of the simulation domain.

973  
974  
975  
976  
977  
978  
979

muir turbulence, e.g., penetration depth relative to the boundary layer depth [Harcourt and D'Asaro, 2008; Kukulka and Harcourt, 2017] and direction relative to the wind [Van Roekel *et al.*, 2012; McWilliams *et al.*, 2014], improves the bias reduction attributable to Langmuir turbulence in a climate model [e.g., Li *et al.*, 2016]. The results here support this claim, as KPPLT-VR12 performs similarly or better than KPP-CVMix in comparison to LES. But further refinements of the enhancement factor, i.e., adding additional degrees of freedom, result in only small changes in the simulated monthly mean MLD [e.g., Li *et al.*, 2017].

980 Thus other effects of Langmuir turbulence, such as entrainment, need to be considered on  
 981 top of the enhanced turbulent diffusivity.

982 KPPLT-LF17 takes into account the fact that the entrainment process is scaled by a  
 983 different velocity scale than the square root of the vertical velocity variance [Li and Fox-  
 984 Kemper, 2017]. This is achieved by further modifying the parameterization of entrainment  
 985 processes in KPP. As a result, KPPLT-LF17 gives stronger mixed layer deepening than  
 986 KPPLT-VR12. This additional deepening shows significantly better performance versus  
 987 KPPLT-VR12 in the LES comparison case 3 here, and much stronger Langmuir effects un-  
 988 der realistic forcing (cases 1 and 2). Consistently, KPPLT-LF17 improves the simulated  
 989 MLD in a climate model by enhancing the Langmuir turbulence induced mixed layer en-  
 990 traignment in the extratropical regions where waves are relatively strong but slightly reduc-  
 991 ing this effect in the tropical regions as compared to KPPLT-VR12 [Li and Fox-Kemper,  
 992 2017]. It is yet to be seen how the discrepancies among other Langmuir schemes shown  
 993 here translate to biases in simulated MLD and other important quantities in a climate  
 994 model.

995 KPPLT-R16 also considers the influence of Langmuir turbulence on the entrainment  
 996 separate from its enhancement effect on the turbulent diffusivity but in a different formu-  
 997 lation than KPPLT-LF17. In particular, it explicitly uses the Lagrangian shear instead of  
 998 the Eulerian shear to parameterize the entrainment due to Stokes drift modulated resolved  
 999 shear and uses a different enhancement factor for the entrainment due to unresolved shear  
 1000 associated with Langmuir turbulence. It further retunes the critical Richardson number  
 1001 to remove any implicit effects of Langmuir turbulence previously included in the tuning  
 1002 parameters of KPP [Reichl *et al.*, 2016]. However, KPPLT-R16 was tuned against LES of  
 1003 idealized hurricanes with neutral surface stability conditions and it lacks explicit consid-  
 1004 erations of scenarios where surface buoyancy flux is destabilizing. As a result, it greatly  
 1005 overestimates the deepening of mixed layer in most test cases shown here except the hur-  
 1006 rricane case. Improvements to KPPLT-R16 can be made by explicitly considering the ef-  
 1007 fect of surface stability condition on its enhancement factor to the unresolved shear as in  
 1008 KPPLT-LF17.

1009 In general, all the above three versions of KPPLT are expected to improve consid-  
 1010 ering the other effects of Langmuir turbulence introduced in section 2. It should also be  
 1011 noted though that they can be affected by numerical details in a manner similar to the un-  
 1012 derlying KPP scheme [e.g., Van Roekel *et al.*, 2018].

1013 The strength of ePBL-LT is that it explicitly sets the rate of mechanical energy con-  
 1014 version to potential energy by mixing, which helps alleviate sensitivity to vertical resolu-  
 1015 tion and time step. In comparison to LES here, ePBL-LT performs the best in the entrain-  
 1016 ment cases and reasonably well in the hurricane case. In other test cases, it gives similar  
 1017 results as KPPLT-LF17, though the difference between ePBL-LT and ePBL is less in mag-  
 1018 nitude than the differences between KPPLT-LF17 and KPP-CVMix, probably due to the  
 1019 different tuning strategies. ePBL-LT and ePBL were tuned to LES with and without Lang-  
 1020 muir turbulence, respectively; whereas KPP-CVMix was tuned to observations that prob-  
 1021 ably included Langmuir to some degree, and KPPLT-LF17 was derived by applying the  
 1022 Langmuir effects seen in LES to KPP-CVMix (see Appendices A.1 and A.2). ePBL-LT  
 1023 also generally performs better in numerical robustness than schemes based on KPP (see  
 1024 Appendix B: ).

1025 SMCLT-H15 agrees well with ePBL-LT and KPPLT-LF17 under realistic forcing,  
 1026 except its Langmuir effects over its non-Langmuir counterpart are somewhat less seasonal  
 1027 and extratropical (Figure 10). Numerically, it is much more sensitive to time step and ver-  
 1028 tical resolution, as is its non-Langmuir counterpart. Thus, in present implementation it is  
 1029 better suited for regional modeling rather than for the coarse vertical resolution and time  
 1030 stepping typical of climate modeling.

1031 OSMOSIS is a new scheme without clear precedent among those compared here.  
 1032 It shows excellent potential, especially in handling restratifying buoyancy forcing and in  
 1033 tracking aspects of the OSMOSIS observations, and it is only moderately sensitive to res-  
 1034 olution. However, it performs very differently from the other schemes especially at low-  
 1035 and mid-latitudes where it is shallower than the non-Langmuir schemes typically, and it  
 1036 performs badly in comparison to the LES in the hurricane case, which may be related to  
 1037 the lack of representation of the effects of resolved shear. None of these results necessarily  
 1038 reject the formulation of the OSMOSIS scheme, only that it requires more testing under  
 1039 diverse regimes and tuning.

1040 KPP-CVMix and KPP-ROMS are in wide use in climate and regional modeling.  
 1041 The results here demonstrate that the numerical distinctions between these schemes, in  
 1042 particular the different algorithms to diagnose the boundary layer depth, are consequential,  
 1043 and KPP-ROMS has difficulty in comparison to the LES cases chosen here, greater sensi-  
 1044 tivity to vertical and time step resolution, and exceptionally strong sensitivity to shear at  
 1045 the mixed layer base. Collectively, these results suggest that the KPP-CVMix formulation  
 1046 is more trustworthy, especially with the recommendations of *Van Roekelet et al.* [2018]. Fur-  
 1047 thermore, it is clear that non-Langmuir schemes are systematically inconsistent with LES  
 1048 including wave forcing, thus KPP-ROMS needs a Langmuir version.

## 1049 5.2 Limitations of the Comparisons

1050 We are using a single column model GOTM5 for the comparison among different  
 1051 Langmuir schemes for its simplicity and cleanliness, minimizing the influences of calling  
 1052 model biases, interactions with lateral processes and feedbacks of coupled system, which  
 1053 can complicate the interpretation of the comparison results. However, for the same rea-  
 1054 sons, this approach also precludes us from doing a direct evaluation against observations  
 1055 and therefore answering the question of which scheme performs the best, except in com-  
 1056 parison to LES which are severely deficient in capturing the forcing regimes needed for  
 1057 climate model applications. For example, the simulated temperature at OCS-Papa with  
 1058 all schemes exhibit increasing warming biases throughout the year when compared with  
 1059 observations. This behavior is due to the lack of horizontal and vertical advection, and  
 1060 lateral mixing in the single column model, which transport the heat gained by the net  
 1061 positive incoming surface heat flux away. To compare with observations at OCS-Papa,  
 1062 this net gain of heat needs to be balanced [see, e.g., *Large*, 1996]. Furthermore, we have  
 1063 made no attempt to evaluate the quality of the GOTM5 numerics here, and it may be that  
 1064 some schemes would perform differently under different numerical schemes (for example,  
 1065 *Large et al.* [1994] implements an additional iterative step to make KPP more robust than  
 1066 is not included in GOTM5). Therefore, when implemented in an OGCM, the discrepan-  
 1067 cies among schemes seen here are likely to be quantitatively different.

1068 Consistent surface forcing data are used to force simulations with different Langmuir  
 1069 schemes. However, due to their different design, different Langmuir schemes may use the  
 1070 same surface forcing data differently. For example, the full profile of Stokes drift is used  
 1071 in KPPLT-R16 and SMCLT-H15, whereas only its surface value and decay depth are used  
 1072 in OSMOSIS, in which an exponential profile is assumed, even though the full profile  
 1073 tends to decay faster than an exponential with depth [Webb and Fox-Kemper, 2011, 2015].  
 1074 The other three Langmuir schemes use the surface layer averaged Langmuir number, in-  
 1075 corporating some integrated information of the profile shape of Stokes drift in addition to  
 1076 its surface value. The directional information of Stokes drift is only used in KPPLT-VR12  
 1077 and partially in KPPLT-LF17. These inconsistencies may contribute to some of the dis-  
 1078 crepancies among schemes seen here, and a thoughtful refactoring of any of the schemes  
 1079 here to take advantage of good ideas in other schemes is likely to be beneficial. In addi-  
 1080 tion, the limited quality of the surface forcing data may also complicate the interpretation  
 1081 of the differences in the comparison results. Comparison between JRA55-do and CORE-  
 1082 II forcing is underway to quantify these uncertainties. Agreement among schemes does

1083 indicate that the schemes perform similarly under this forcing, but may also result from  
 1084 the lack of important features in the forcing data to reveal the different physics, such as  
 1085 sea state effects on stresses [e.g., Janssen, 1989; Shimura *et al.*, 2017] and submesoscale-  
 1086 OSBL interactions [e.g., Hamlington *et al.*, 2014; Suzuki *et al.*, 2016; Callies and Ferrari,  
 1087 2018] which are known to be important but are outside of the forcing considered here.  
 1088 This comparison has nonetheless demonstrated why we need global atmospheric forcing  
 1089 datasets such as CORE-II and JRA55-do to cover typical surface forcing conditions over  
 1090 the global ocean. Continuing improvements in these datasets reduce these uncertainties,  
 1091 which is why we prefer the higher space and time resolution of JRA55-do over its prede-  
 1092 cessor CORE-II. That is, JRA55-do should have a better representation of high-frequency  
 1093 and extreme events.

### 1094 5.3 Future Directions

1095 One critical question yet to be addressed is how well do these Langmuir schemes  
 1096 perform? While section 4.4 presents some preliminary comparisons against LES of some  
 1097 idealized cases, a full evaluation of these schemes requires additional LES cases, espe-  
 1098 cially those with more realistic configurations and transient surface forcing where the  
 1099 model ensemble spread was wide in this comparison. For example, comparisons with LES  
 1100 of mixed layer deepening [Kukulka *et al.*, 2009], diurnal restratification [Kukulka *et al.*,  
 1101 2013], stable surface forcing [Pearson *et al.*, 2015] and LES of the Southern Ocean Flux  
 1102 Station [Schulz *et al.*, 2012; Large *et al.*, 2019], will be worthwhile for improved evalua-  
 1103 tion and refined distinctions among schemes under transient surface forcing. Direct mea-  
 1104 surements of turbulent mixing, such as the turbulent fluxes, under various surface forcing  
 1105 conditions are potentially extremely valuable. Examining cost function effects of different  
 1106 schemes in ocean reanalyses fit to observations such as the Southern Ocean State Estimate  
 1107 [Mazloff *et al.*, 2010] may help indicate the degree to which lateral processes and vertical  
 1108 advection play a role.

1109 An immediate question following the evaluation step is how to improve the par-  
 1110 ameterizations of Langmuir turbulence? The greater uncertainties in Figure 13b suggest that  
 1111 more LES and observational studies of Langmuir turbulence in regimes with various sta-  
 1112 bilizing surface buoyancy flux, such as diurnal heating, are needed following Min and Noh  
 1113 [2004], Pearson *et al.* [2015], and Walker *et al.* [2016]. As shown in the regime diagram  
 1114 of Figure 1b, the parameter space covered by these equilibrium LES studies is rather lim-  
 1115 ited. In addition, better understanding of the extra dimensions that are not represented  
 1116 in the regime diagram of Figure 1 may also be important in explaining the discrep-  
 1117 ances among these schemes. It is perhaps easiest to cross-pollinate these ideas among these  
 1118 schemes and study the consequences in comparison to studies quantifying their effects in  
 1119 LES, such as: penetration depth of Stokes drift [Harcourt and D’Asaro, 2008; Kukulka and  
 1120 Harcourt, 2017], wind and wave misalignment [Van Roekel *et al.*, 2012], Earth’s rotation  
 1121 [e.g., Liu *et al.*, 2018], and interactions between Langmuir turbulence with other processes  
 1122 such as submesoscale eddies and fronts [e.g., Hamlington *et al.*, 2014; Suzuki *et al.*, 2016],  
 1123 inertial currents [e.g., Wang *et al.*, 2018] and combined effects of sensible, latent, and pen-  
 1124 etrative heat fluxes [e.g., Min and Noh, 2004; Kukulka *et al.*, 2013; Pearson *et al.*, 2015].

1125 Another interesting question is related to the sensitivity of each scheme to the sur-  
 1126 face forcing. Preliminary comparisons between GOTM5 simulations forced by JRA55-do  
 1127 and CORE-II show that the relative discrepancies among schemes seen here are insensi-  
 1128 tive to the use of JRA55-do or CORE-II datasets. However, these two datasets do result in  
 1129 significant differences in the simulated mean state (e.g., the MLD), which is comparable  
 1130 to the discrepancies among schemes. A closer comparison between GOTM5 simulations  
 1131 forced by the two datasets is ongoing and will be helpful in this regard. Free from compli-  
 1132 cated interactions among processes as in an OGCM, such single column simulations may  
 1133 also assist the analyses and interpretation of global OGCM simulations forced by CORE-II  
 1134 and JRA55-do.

Finally, the suite of six Langmuir schemes compared here each emphasizes one or more of the elements of Langmuir turbulence parameterization (section 2). But this is by no means an exhaustive list, there are other Langmuir turbulence parameterization schemes emphasizing different elements, in particular those on the modified vertical distribution (shape) of turbulent diffusivity [Sinha *et al.*, 2015; Yang *et al.*, 2015], on the non-local fluxes [e.g., Smyth *et al.*, 2002], and misalignment between momentum flux and shear [McWilliams *et al.*, 2012]. It will be interesting to compare these schemes with the six examined here, with the comparison framework of realistic and LES forcing used here as a guide. In addition, five of the six Langmuir schemes compared here (except SMCLT-H15) are built on first-moment turbulence closure models, largely reflecting the fact that these first-moment closure schemes are widely used among most global ocean models. However, some representations of higher-moments may be needed to fully capture the unique structure of Langmuir turbulence shown in LES [e.g., McWilliams *et al.*, 1997] and its effect on the vertical mixing in the OSBL.

## 6 Summary and Conclusions

A suite of six different Langmuir turbulence parameterization schemes has been implemented in a common single column modeling framework using GOTM5. This setup allows, for the first time, a direct comparison among these Langmuir schemes, as well as a consistent comparison to five traditional schemes without Langmuir turbulence. We documented results using three test scenarios to emphasize the performance of these schemes from multiple perspectives. The three scenarios explored here are: (i) idealized conditions commonly used in LES studies of Langmuir turbulence; (ii) realistic conditions based on field measurements at ocean stations; (iii) climatological forcings based on global atmospheric states targeted for use in forcing global ocean-ice models.

The extended GOTM5 code and test suite developed here (see Appendix D: for more details) represent an initial effort towards a common modeling framework and test suite for testing ocean boundary layer physics. This framework offers the means to systematically compare different ocean boundary vertical mixing schemes with and without Langmuir turbulence, and to do so in a consistent way under different scenarios as evaluated against different sets of data. Our efforts with this framework in the present paper have led to the following conclusions.

- As expected, Langmuir schemes generally predict a deeper mixed layer depth (pointwise in monthly-means, 6%, -1% to 14% for 90% confidence range or 5.2 m, -0.2 m to 17.4 m for 90% confidence; 5-25% deeper in zonal mean in extratropical latitudes) than the non-Langmuir schemes, especially when convection is weak such as the austral summer Southern Ocean. Significant differences exist both among Langmuir schemes (with a mean MLD standard deviation of 15%) and among non-Langmuir schemes (6%), resulting in overlaps between the two groups thus complicating a scheme-independent estimation of the amount of Langmuir deepening. The substantial differences among the Langmuir schemes highlight the necessity for better constraints on the effects of Langmuir turbulence from observations and/or LES studies.
- Including Langmuir turbulence changes the behavior of vertical mixing schemes. This behavior was demonstrated by comparing individual Langmuir schemes to their non-Langmuir counterparts. In general, the most significant Langmuir turbulence induced deepening of the mixed layer (9% on average, 2% to 17% for 90% confidence range, in Langmuir turbulence regime) is found in extratropical regions, especially the Southern Ocean in austral summer.
- Discrepancies among schemes are larger in the Langmuir turbulence regime than in the convection regime, and larger in stabilizing surface conditions than in destabilizing surface conditions.

- 1186 • Correctly predicting the entrainment rate in the presence of Langmuir turbulence  
1187 remains challenging for most of the Langmuir schemes, even under idealized steady  
1188 forcing. It is even more difficult under transient surface forcing conditions.
- 1189 • Many important regimes in the parameter space of surface forcing that are typi-  
1190 cal in the global ocean have not been well explored in LES. For example, previ-  
1191 ous LES studies of Langmuir turbulence have mostly focused on neutral surface  
1192 forcing with only a few exceptions, whereas the real world appears to mostly lie in  
1193 the combined Langmuir turbulence and convection, and Langmuir turbulence with  
1194 stabilizing surface forcing regimes. More LES and observational studies in these  
1195 regimes and the regions and seasons indicated in Fig. 12 are required to better con-  
1196 strain the effects of Langmuir turbulence and boundary layer turbulence in general.
- 1197 • The formulation and numerical implementation method of a scheme, e.g., KPP-  
1198 CVMix versus KPP-ROMS, can also lead to very different results or different nu-  
1199 matical robustness even though the underlying physical foundation is the same.

1200 Although offering more questions than answers, the results of this intercomparison pro-  
1201 vide guidance for further research into the parameterization of ocean boundary layer mix-  
1202 ing. Future efforts will benefit from a focus on forcing regimes where the present suite of  
1203 mixing schemes diverge, with an emphasis on strategies presented here to diagnose skills  
1204 and differences. The disagreement of the schemes in even simple, idealized regimes sug-  
1205 gests that the idealized process study approach is far from exhausted. Furthermore, the  
1206 complexities of the realistic simulations suggest a broad parameter space that remains rel-  
1207 atively unexplored. Continued development of robust and accurate parameterizations re-  
1208 mains a critical endeavor for advancing regional and global model simulations.

## 1209 A: Description and Implementation of Different Schemes

### 1210 A.1 KPP and KPPLT

1211 Here we only briefly review the key ingredients of KPP that are relevant to the mod-  
1212 ifications to include the effects of Langmuir turbulence in the three KPPLT variants (KPPLT-  
1213 VR12, KPPLT-LF17 and KPPLT-R16) in our comparison suite. Other components of  
1214 KPP, e.g., the non-local term, are treated the same among all the KPP variants in this  
1215 study (KPP-CVMix, KPP-ROMS, and the above three versions of KPPLT), and there-  
1216 fore excluded in the discussion here for brevity. The full description of KPP can be found  
1217 in *Large et al. [1994]* and *Van Roekel et al. [2018]*, with the specific adaptation in KPP-  
1218 CVMix detailed in *Griffies et al. [2015]*.

1219 KPP parameterizes the eddy diffusivity  $K_\lambda$  in equation (2) from the boundary layer  
1220 depth  $h_b$ , a turbulent velocity scale  $w_\lambda(\sigma)$  depending on the friction velocity  $u_*$  and the  
1221 Monin-Obukhov similarity functions, and a dimensionless shape function  $G_\lambda(\sigma)$ :

$$1222 K_\lambda(\sigma) = h_b w_\lambda(\sigma) G_\lambda(\sigma), \quad (\text{A.1})$$

1223 with  $\sigma = z/h_b$  a dimensionless vertical coordinate.

1224 KPP-CVMix follows the bulk Richardson number calculation of *Large et al. [1994]*  
1225 and *Griffies et al. [2015]* to diagnose the boundary layer depth  $h_b$ . In this approach, the  
1226 boundary layer depth is diagnosed by finding the shallowest depth where the bulk Richard-  
son number,

$$1227 \text{Ri}_b(z) = \frac{z \left[ b_r - \bar{b}(z) \right]}{[\mathbf{u}_r - \bar{\mathbf{u}}(z)]^2 + U_t^2(z)}, \quad (\text{A.2})$$

1228 reaches a critical value,  $\text{Ri}_c = 0.3$ . The reference velocity  $\mathbf{u}_r$  and buoyancy  $b_r$  are found  
1229 by averaging over the surface layer ( $z > -0.1h_b$ ) to reduce the resolution dependency. The  
term  $U_t^2(z)$  in the denominator aims to account for the effects of unresolved shear.

1230 KPP-ROMS uses a different approach, whereby  $h_b$  is diagnosed as the first non-zero  
 1231 depth at which  $\text{Cr}(z) = 0$ , where

$$\text{Cr}(z) = \int_z^0 J(z') \left[ (\partial_z \bar{\mathbf{u}})^2 - \frac{N^2(z)}{\text{Ri}_c} - C_{\text{ek}} f^2 \right] dz' + \frac{U_t^2(z)}{|z|}, \quad (\text{A.3})$$

1232 with  $N^2 = \partial_z \bar{b}$  the square of the Brunt-Väisälä frequency,  $J(z) = |z|/(|z| + 0.1 h_b)$  a weight-  
 1233 ing function, and  $f$  the Coriolis parameter. Note that the last term in the integral with  $C_{\text{ek}}$   
 1234 a constant represents the stabilizing effect of rotation, which is absent in the bulk Richard-  
 1235 son number formula in equation (A.2). See Appendix B of *McWilliams et al.* [2009] for  
 1236 more discussion on this formula.

1237 In accordance with the observation that the vertical velocity variance within the  
 1238 OSBL is enhanced in the presence of Langmuir turbulence, the most straightforward mod-  
 1239 ification of KPP is applying an enhancement factor  $\mathcal{E}$  to the turbulent velocity scale  $w_\lambda$ ,  
 1240 or  $w_{\lambda L} = \mathcal{E} w_\lambda$ . Following the idea of *McWilliams and Sullivan* [2000], various formulas  
 1241 of  $\mathcal{E}$  as a function of Langmuir number have been proposed [e.g., *McWilliams and Sulli-*  
 1242 *vian*, 2000; *Smyth et al.*, 2002; *Takaya et al.*, 2010; *Li et al.*, 2016]. For simplicity, only one  
 1243 such KPPLT model is presented here (KPPLT-VR12). The relative differences among the  
 1244 different formulas of  $\mathcal{E}$  are illustrated in Figure A.1. In KPPLT-VR12, the enhancement  
 1245 factor is written as a function of the surface layer averaged and projected Langmuir num-  
 1246 ber  $\text{La}_{\text{SLP}}$  defined in (6) based on the LES work of *Van Roekel et al.* [2012],

$$\mathcal{E} = |\cos \theta_{\text{wl}}| [1 + (3.1 \text{La}_{\text{SLP}})^{-2} + (5.4 \text{La}_{\text{SLP}})^{-4}]^{1/2}, \quad (\text{A.4})$$

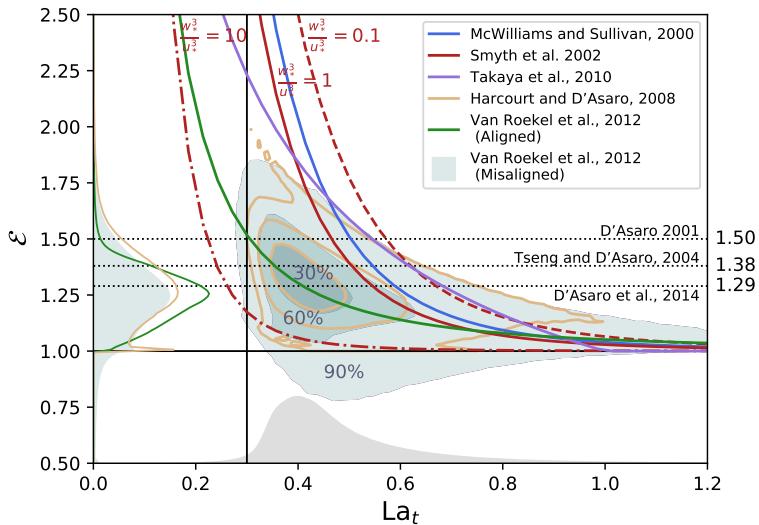
1247 where  $\theta_{\text{wl}}$  is the angle between wind and Langmuir cells. It is expected from Figure A.1  
 1248 that enhancement factors based on *McWilliams and Sullivan* [2000], *Smyth et al.* [2002],  
 1249 or *Takaya et al.* [2010] will lead to much stronger enhanced vertical mixing than equa-  
 1250 tion (A.4), as shown in *Li et al.* [2016] and *Ali et al.* [2019]. We note, however, that the  
 1251 VR12 case in *Ali et al.* [2019] is different from KPPLT-VR12 detailed here by the use of  
 1252  $\text{La}_t$  and the additional Stokes drift term in the bulk Richardson number in KPP [see more  
 1253 discussions on this term in *Li et al.*, 2016].

1254 In addition to the direct effect of enhanced  $w_\lambda$  that increases the eddy diffusivity  
 1255 according to equation (A.1), the entrainment at the base of the OSBL is also enhanced ac-  
 1256 cording to equation (A.2), where the unresolved shear term  $U_t^2$  is a function of  $w_\lambda$ . There-  
 1257 fore, this approach simply assumes that the entrainment is affected by Langmuir turbu-  
 1258 lence in the same way as the vertical turbulent diffusion, which turns out to be insufficient  
 1259 [*Li and Fox-Kemper*, 2017]. Improvements are possible by separately considering the ef-  
 1260 ffects of Langmuir turbulence on those processes.

1261 *Li and Fox-Kemper* [2017] show that the entrainment buoyancy flux is affected dif-  
 1262 ferently by the presence of Langmuir turbulence and follows a different scaling law than  
 1263 the vertical velocity variance within the OSBL. KPPLT-LF17 therefore further incorpo-  
 1264 rates this new scaling law of entrainment buoyancy flux into KPP in addition to the en-  
 1265 hancement factor as in KPPLT-VR12 by modifying the unresolved shear term in equa-  
 1266 tion (A.2) to,

$$U_{tL}^2(z) = \frac{C_v N(z) w_\lambda(z) |z|}{\text{Ri}_c} \left[ \frac{0.15 w_*^3 + 0.17 u_*^3 (1 + 0.49 \text{La}_{\text{SLP}}^{-2})}{w_\lambda(z)^3} \right]^{1/2}, \quad (\text{A.5})$$

1267 where  $C_v$  is a dimensionless coefficient and  $w_* = (-B_0 h)^{1/3}$  is the convective veloc-  
 1268 ity scale. Note that, unlike in KPPLT-VR12, in KPPLT-LF17 the enhancement factor in  
 1269 equation (A.4) is only applied to the eddy diffusivity in equation (A.1) and thereby  $w_\lambda$  in  
 1270 equation (A.5) is not enhanced. It is also important to note that all coefficients in equa-  
 1271 tions (A.4) and (A.5) are derived from LES and no tuning is required when applied in  
 1272 KPP.



**Figure A.1.** Comparison among the formulas of enhancement factor  $\mathcal{E}$  from *McWilliams and Sullivan* [2000], *Smyth et al.* [2002], *Takaya et al.* [2010], *Harcourt and D'Asaro* [2008], aligned and misaligned versions of *Van Roekel et al.* [2012]. Three cases with  $w_*^3/u_*^3 = [0.1, 1, 10]$ , with  $w_* = (-B_0 h)^{1/3}$  the convective velocity scale, are shown for *Smyth et al.* [2002]. Since the conversion from  $\text{La}_{\text{SL}}$ , or  $\text{La}_{\text{SLP}}$ , to  $\text{La}_t$  requires additional degrees of freedom, such as the full Stokes drift profile shape and  $h_b$  which are not included in this diagram, the isolines of the joint probability distributions function (PDF) of  $\mathcal{E}$  and  $\text{La}_t$  that enclose 30%, 60% and 90% of all instances centered at the highest joint PDF are shown by the contours and shadings for *Harcourt and D'Asaro* [2008] and misaligned version of *Van Roekel et al.* [2012], respectively. The PDFs of  $\mathcal{E}$  for selected cases are shown on the left side of the diagram. The PDF of  $\text{La}_t$  is shown by the gray shading at the bottom. All PDFs are estimated from the 3-hourly output of the GOTM simulations forced by JRA55-do over 12 months. Enhancement factors inferred from direct measurements of *D'Asaro* [2001], *Tseng and D'Asaro* [2004] and *D'Asaro* [2014] are marked by the dotted lines and labeled on the right for reference. Horizontal and vertical reference lines in black mark  $\mathcal{E} = 1$  and  $\text{La}_t = 0.3$ , respectively.

KPPLT-R16 uses different enhancement factors for eddy diffusivity and unresolved shear, respectively,

$$\begin{aligned}\mathcal{E}_K &= 1 + (\mathcal{E}' - 1)G_\lambda(\sigma)/\max[G_\lambda(\sigma)], \\ \mathcal{E}' &= \min(2.25, 1 + \text{La}_{\text{SLP}}^{-1}), \\ \mathcal{E}_{U_t^2} &= 1 + 2.3 \text{La}_{\text{SLP}}^{-1/2},\end{aligned}\quad (\text{A.6})$$

where the enhancement of the eddy diffusivity is concentrated near its peak value by the weighting function  $G_\lambda(\sigma)/\max[G_\lambda(\sigma)]$ . Note that here a slightly different definition of  $\text{La}_{\text{SLP}}$  than Equation (6) is used [see Eq. (25) of *Reichl et al.*, 2016, and the discussion therein]. In addition, the empirical parameters in KPP are retuned to first remove any implicit effects of Langmuir turbulence before explicitly adding those effects by applying an enhancement factor. In particular, *Reichl et al.* [2016] find that a smaller critical Richardson number,  $Ri'_c = 0.235$ , gives optimal agreement between KPP and LES under tropical cyclone conditions without Langmuir turbulence. This value may require retuning depending on factors such as the vertical resolution, which is consistent in the experiments presented here. They further use the Lagrangian shear in the definition of bulk Richardson number instead of the Eulerian shear to represent the effects of down-Stokes drift shear

1286 mixing on the entrainment. The criterion to find  $h_b$  in KPPLT-R16 is therefore the shal-  
1287 lowest depth where

$$\text{Ri}'_b(z) = \frac{z \left[ b_r - \bar{b}(z) \right]}{\left[ u_r^L - \bar{u}^L(z) \right]^2 + U_t^2(z) \mathcal{E}_{U_t^2}} \leq \text{Ri}'_c = 0.235. \quad (\text{A.7})$$

1288 Here  $\mathbf{u}^L = \mathbf{u} + \mathbf{u}^S$  is the Lagrangian current.

1289 Note that the Coriolis-Stokes force is explicitly written out in the horizontal momen-  
1290 tum equation in KPPLT-R16, but is implicit in KPPLT-VR12 and KPPLT-LF17 where the  
1291 simulated  $\mathbf{u}$  is treated as the Lagrangian flow. In this sense, the usage of  $\mathbf{u}$  in the bulk  
1292 Richardson number in KPPLT-VR12 and KPPLT-LF17 is generally consistent with the us-  
1293 age of  $\mathbf{u}^L$  in KPPLT-R16.

1294 It should also be noted that parameters in KPPLT-R16 were tuned against LES of  
1295 idealized hurricanes in *Reichl et al.* [2016] with nearly neutral surface conditions, in con-  
1296 trast to KPPLT-LF17 which were tuned against LES with weak to moderate wind and var-  
1297 ious destabilizing surface buoyancy fluxes. This may explain some of the different behav-  
1298 iors between these two, especially under strong destabilizing surface forcing conditions.

1299 Consistent with *Van Roekel et al.* [2018], we found that the results of KPP-CVMix  
1300 (thereby all the KPPLT schemes here) are quite sensitive to the choice of where  $N$  (de-  
1301 fined at cell interfaces) is evaluated in the unresolved shear term  $U_t^2$  in the bulk Richard-  
1302 son number  $\text{Ri}_b$  (both defined at cell centers) in equation (A.2). Default setup in CVMix  
1303 uses the values at the interfaces below the cell centers following *Danabasoglu et al.* [2006],  
1304 which does not behave very well with the relatively high vertical resolution ( $\Delta z = 1$  m)  
1305 used in our runs here. Using this setup generally results in too small  $N$  (and thereby  $U_t^2$ )  
1306 and strongly affect the apparent effects of Langmuir turbulence enhanced entrainment in  
1307 KPPLT-LF17 and KPPLT-R16. Ultimately this strong sensitivity to the numerical details  
1308 is related to how well the stratification at the base of the boundary layer is resolved. In  
1309 theory, the maximum stratification defines the boundary layer base in the limit of pure  
1310 convection, from which  $U_t^2$  is derived. We therefore follow the recommendation of *Van Roekel*  
1311 *et al.* [2018] by using the maximum value of  $N$  above and below the cell center.

1312 KPPLT-VR12, KPPLT-LF17 and KPPLT-R16 were implemented in the CVMix pack-  
1313 age instead of being directly implemented in GOTM5. Thus, simulations with these schemes,  
1314 as well as KPP-CVMix, were realized by calling the CVMix library. This approach takes  
1315 advantage of the significant effort of implementing some of the KPPLT schemes in CVMix  
1316 [e.g., *Li et al.*, 2017], but also means that the software, algorithm, and numerics of the  
1317 CVMix implementation—just as they would be used in a climate model—are being eval-  
1318 uated alongside the physical differences between schemes. The formulation of KPP-CVMix  
1319 and parameters selected are based on the default configuration in CVMix [*Danabasoglu*  
1320 *et al.*, 2006; *Griffies et al.*, 2015]. With the exception of the method to diagnose the value  
1321 of  $N^2$  noted in the preceding paragraph, the recommendations in *Van Roekel et al.* [2018]  
1322 are not included, which are expected to make a smaller difference on the results than KP-  
1323 PLT schemes.

1324 KPP-ROMS uses a different critical condition for determining the boundary layer  
1325 depth, replacing the original critical Richardson number condition (A.2) by the condi-  
1326 tion (A.3). As yet it includes no Langmuir turbulence parameterization effects. It is de-  
1327 scribed in *McWilliams et al.* [2009] and *Lemarié et al.* [2012]. It also uses the particular  
1328 space-time algorithms and discretizations employed in the UCLA version of the Regional  
1329 Oceanic Modeling System [ROMS *Shchepetkin and McWilliams*, 2005, 2009]. KPP-ROMS  
1330 was implemented in GOTM5 directly as a subroutine following the KPP subroutine of  
1331 UCLA ROMS. The time-stepping and conservation laws are all controlled by GOTM5.

1332 **A.2 ePBL and ePBL-LT**

1333 The ePBL parameterization is the OSBL turbulent mixing parameterization as de-  
 1334 scribed in *Reichl and Hallberg* [2018]. It is constructed as a framework for climate simu-  
 1335 lations by introducing turbulent mixing with a relatively weak dependence on model verti-  
 1336 cal resolution and time step. The technical details that provide such capability exploit the  
 1337 fact that the change in column potential energy can be exactly computed from the linear  
 1338 integral of the change in specific volume through a vertical coordinate transformation to  
 1339 pressure. The buoyancy diffusion problem, assuming a known turbulent diffusion coeffi-  
 1340 cient, is also linear to the close approximation that the equation of state is nearly linear  
 1341 over small changes in temperature and salinity. This linearity makes it possible to track  
 1342 the integrated change of potential energy of the entire water column as a result of chang-  
 1343 ing the diffusion coefficient anywhere within the water column. The algorithm's implicit  
 1344 diffusion solver (e.g., the tridiagonal solver) can thus be implicitly coupled to the energetic  
 1345 considerations within the ePBL turbulence parameterization.

1346 Similar to equation A.1, ePBL defines the turbulent diffusivity profile  $K_\lambda(z)$  from  
 1347 profiles of an empirical velocity  $w(z)$  and a length scale  $L(z)$ ,

$$1348 K_\lambda(z) = C_\lambda w(z)L(z), \quad (\text{A.8})$$

1349 where  $C_\lambda$  is a dimensionless coefficient. There is no restriction in ePBL on the specific  
 1350 form used to estimate  $w(z)$  and  $L(z)$ . The primary emphasis of ePBL is in setting the in-  
 1351 tegral of the vertical turbulence buoyancy flux that describes potential energy change asso-  
 1352 ciated with turbulent mixing,

$$1353 M_e = - \int_{-h_b}^0 K_b(z) \max(0, \partial_z \bar{b}) dz = - \int_{-h_b}^0 C_b w(z) L(z) \max(0, \partial_z \bar{b}) dz, \quad (\text{A.9})$$

1354 where  $\bar{b}$  is the mean buoyancy (the calling model's resolved buoyancy), and the non-local  
 1355 flux is zero. This constraint is imposed by pre-defining  $M_e$  via various parameterizations,  
 1356 and then seeking solutions for the boundary layer depth  $h_b$  and the corresponding  $K_b(z)$   
 1357 profile that satisfy equation (A.9). In practice this strategy is applied for temperature and  
 1358 salinity. Here we describe the specific forms of  $w(z)$ ,  $L(z)$ , and  $M_e$  adopted for this study,  
 1359 which are based on the version of ePBL described in *Reichl and Hallberg* [2018]; *Reichl*  
 and *Li* [2019].

1359 The value of  $L(z)$  in ePBL is given as,

$$1360 L(z) = (z_0 + |z|) \max \left[ \frac{l_b}{h_b}, \left( \frac{h_b - |z|}{h_b} \right)^\gamma \right], \quad (\text{A.10})$$

1361 with  $\gamma = 2$  providing a similar shape to KPP [*Large et al.*, 1994] and  $l_b$  being a bottom  
 1362 length scale which is dependent on bottom roughness or interior stratification and prevents  
 1363  $L$  from becoming zero at the base of the OSBL. The value of  $w(z)$  in ePBL is given by

$$1364 w(z) = C_{w_*} \left( \int_z^0 \overline{w' b'} dz \right)^{1/3} + (c_\mu^0)^{1/3} u_* \left[ 1 - a \cdot \min \left( 1, \frac{|z|}{h_b} \right) \right], \quad (\text{A.11})$$

1365 where  $C_{w_*}$  and  $c_\mu^0$  are empirical coefficients and  $a$  is a fixed vertical decay scale.

1366  $M_e$  is parameterized from the surface buoyancy flux  $B_0$ , the surface friction veloc-  
 1367 ity  $u_*$ , the boundary layer depth  $h_b$ , the Coriolis parameter  $f$ , and for ePBL-LT, the  
 1368 surface layer averaged Langmuir number  $\text{La}_{SL}$ . The formula of  $M_e$  was originally optimized  
 1369 to represent integral properties of the turbulent mixing inferred from  $k$ - $\epsilon$  (with *Schu-*  
 1370 *mann and Gerz* [1995] stability functions) simulations of buoyancy and shear driven tur-  
 1371 bulent boundary layers [see *Reichl and Hallberg*, 2018]. When applied to a global ocean  
 1372 model this version of ePBL yields shallow-mixing biases in mechanically-driven tur-  
 1373 bulence regimes, a result consistent with missing the Langmuir turbulence contribution such

as found using KPP by *Li et al.* [2016]. Similar to the KPPLT approach, an additional Langmuir turbulence term as a function of  $La_{SL}$  was thus introduced to represent the effect of surface waves and ultimately reduce this bias [*Reichl and Li*, 2019]. The effect of this additional term is to provide additional energy to facilitate mixing against stable stratification and deepen the boundary layer in a manner consistent with the LES results of *Li and Fox-Kemper* [2017]. The final relationship is expressed as

$$M_e = (m_* + m_{*LT})u_*^3 + n_* \int_{-h_b}^0 \max(0, \overline{w'b'}) dz, \quad (\text{A.12})$$

where the coefficient for convective mixing  $n_*$  is taken as 0.065 and the coefficient for mechanical mixing  $m_*$  is given depending on the forcing conditions. Here we adopt forms for  $m_*$  of

$$m_* = m_{*N}\Psi + m_{*S}, \quad (\text{A.13})$$

which separates the stabilizing ( $S$ ) and non-stabilizing ( $N$ ) regimes as:

$$m_{*N} = c_{N1} \left\{ 1 - [1 + c_{N2} \exp(-c_{N3} h_b |f|/u_*)]^{-1} \right\}, \quad (\text{A.14a})$$

$$m_{*S} = c_{S1} \left[ \max(0, B_0)^2 h_b / u_*^5 |f| \right]^{c_{S2}}, \quad (\text{A.14b})$$

$$\Psi = 1 - c_\Psi \frac{\max(0, -B_0)}{\max(0, -B_0) + 2m_{*N} u_*^3 / h_b}, \quad (\text{A.14c})$$

where the empirical coefficients of *Reichl and Hallberg* [2018] are adopted in our experiments ( $c_{N1} = 0.275$ ,  $c_{N2} = 8$ ,  $c_{N3} = 5$ ,  $c_{S1} = 0.2$ ,  $c_{S2} = 0.4$ , and  $c_\Psi = 0.67$ ). A maximum  $m_*$  is set to 10, preventing the stabilizing relationship from running away as  $f \rightarrow 0$ . The coefficient for Langmuir mixing  $m_{*LT}$  used here is given by,

$$m_{*LT} = 0.105 \left[ La_{SL} \left( 1 + 0.8 \frac{|B_0|}{f u_*^2} \right) \right]^{-1}. \quad (\text{A.15})$$

In addition to this  $K_\lambda(z)$  profile, the *Jackson et al.* [2008] shear-driven mixing parameterization is employed, as discussed in detail by *Reichl and Hallberg* [2018]. When values of  $K_\lambda(z)$  in both ePBL and Jackson are present, the larger of the two is used to govern the vertical mixing.

A version of the ePBL source code taken from the MOM6 repository was adapted to a form capable of calling directly by GOTM5 here. The emphasis of ePBL on implicit numerics is critically maintained in our adaption of the algorithm from MOM6 to GOTM5. To achieve this implicitness, the ePBL algorithm embeds an implicit solver that is iterated for a given model time step over both the temperature and salinity fields as well as the model  $K$  profile and non-dimensional forcing relations (that determine the amount of mixing allowed to occur). For ePBL to be fully interchangeable with other  $K$ -profile type mixing parameterizations with minimal code modification, the ePBL algorithm is designed to return the equivalent  $K$  profile at its conclusion rather than directly modifying the model state itself. This  $K$  profile can then be easily interchanged with the various other parameterizations and applied with the GOTM5 tridiagonal routine to apply the appropriate turbulent mixing. Since GOTM5 has a Crank-Nicholson type time-stepping scheme this requires GOTM5 with ePBL to be run in backward Euler mode by setting the Crank-Nicholson coefficient to be one.

### A.3 OSMOSIS

The OSMOSIS parameterization scheme was implemented in GOTM by modifying an existing KPP model which parameterizes turbulent fluxes using equation (2) [*Large et al.*, 1994]. The full description is underway and will be published in a separate paper. This appendix summarizes the main ideas of the OSMOSIS scheme.

The OSMOSIS scheme makes several modifications to the KPP scheme to account for the effects of Langmuir turbulence and to improve upon the scheme. The first modification is that the OSMOSIS scheme separates the boundary layer into two regions, a mixed layer and a pycnocline, which have finite depth and differing structure. This is in contrast to KPP where only the boundary layer is defined and the interface between this layer and the exterior is free to develop as the boundary layer rises and falls. This means that the OSMOSIS scheme has several non-dimensional vertical coordinates in addition to  $\sigma = z/h_b$  (equation (A.1)), which can affect shape functions within the pycnocline and mixed layer. It also means that the gradients of properties vary through the depth of the boundary layer.

The second OSMOSIS modification is that both the diffusivity  $K_\lambda$  and non-local fluxes  $\Gamma_\lambda$  of scalars ( $\lambda = T, S$  etc.) and of momentum ( $\lambda = u, v$ ) include Stokes drift effects. Specifically, the diffusivity is

$$K_\lambda = \mathcal{F}_{\lambda,\phi}(u_*, w_{*L}, w_*, \sigma_i, h_b, h_m, \phi), \quad (\text{A.16})$$

where  $u_* = (u_*^2 u_0^S)^{1/3} = (-B_0 L_L)^{1/3}$  and  $w_*$  are the respective velocity scales of shear-driven, Langmuir and convective turbulence,  $\sigma_i$  denotes multiple non-dimensional co-ordinates based on the depths of the boundary layer  $h_b$  and mixed layer  $h_m$  and the thickness of the pycnocline ( $h_b - h_m$ ), and  $\phi = h_b/L_*$  is a stability parameter with  $L_*$  a stability length scale. In equation (A.16) the functional form of  $\mathcal{F}_{\lambda,\phi}$  varies for different diffused variables ( $\lambda$ ) and between stable and unstable conditions ( $\phi$ ).

The non-local fluxes in the OSMOSIS scheme can be decomposed as,

$$\Gamma_\lambda = \Gamma_{\text{st}} + \Gamma_{\text{buoy}} + \Gamma_{\text{trnsp}} + \Gamma_{\text{pyc}} + \Gamma_{\text{ent}}, \quad (\text{A.17})$$

where the terms on the right respectively denote fluxes caused by Stokes drift, buoyancy, non-local transport, pycnocline structure and entrainment at the base of the boundary layer. Like the diffusivity, the functional form of the fluxes depend on the variable being transported and the stability of the boundary layer. Notably, the mechanism driving  $\Gamma_{\text{buoy}}$  switches from convection under unstable conditions, to boundary layer shoaling under stable conditions ( $\Gamma_{\text{buoy}} = 0$  in a shoaling boundary layer). The fluxes in equation (A.17) are functions of velocity scales, surface fluxes, and non-dimensional coordinates ( $\sigma_i$ ) including  $z/\delta^S$  where  $\delta^S$  is the Stokes decay depth.

The third unique aspect of the OSMOSIS scheme is its different treatment of behavior under stable and unstable conditions. This partly manifests as the dependence of  $K_\lambda$  and  $\Gamma_\lambda$  on the stability ( $\phi$ ), similar to KPP. However, the OSMOSIS scheme solves prognostic equations for the boundary layer depth whether the boundary layer is deepening or shoaling, while many other schemes use a diagnostic equation for boundary layer depth under shoaling conditions. Part of the reason for this is that the OSMOSIS scheme was developed using an array of LES, including several simulations of Langmuir turbulence stabilized by a variety of radiative and sensible heat fluxes (red dots in Figure 1b; note that the parameter space of radiative heating profiles is not visible in the figure).

#### A.4 SMCLT

The Harcourt [2015] second moment closure (SMC) of Langmuir turbulence (SMCLT-H15) was implemented in GOTM5 by modifying the existing implementation of the *Kantha and Clayson* [1994] quasi-equilibrium version of the ‘level 2.5’  $q^2 \cdot q^2 l$  SMC of *Mellor and Yamada* [1974]. These SMC models combine two non-equilibrium prognostic equations for the turbulent kinetic energy (TKE =  $q^2/2$ ) and its product  $q^2 l$  with the dissipation length scale  $l$ , with a linearized algebraic Reynolds stress model (ARSM) that assumes local equilibrium balance for each Reynolds stress and flux tensor component. The quasi-equilibrium assumption of *Galperin et al.* [1988] simplifies the stability functions whereby the ARSM relates  $q^2$  and  $l$  to vertical turbulent fluxes, and distinguishes

1456 this ‘level 2 1/4’ model from MY2.5 and several other weak equilibrium formulations also  
 1457 implemented in GOTM.

1458 SMCLT-H15 differs from earlier SMCs of Langmuir turbulence [e.g., *Kantha and*  
 1459 *Clayson, 2004*] which account for CL vortex production [*Craik and Leibovich, 1976*] in  
 1460 the equations for  $q^2$  and  $q^2l$  by also including vortex production terms in the ARSM. For  
 1461 implementation of an SMC in an upper ocean model, the solution to the ARSM for the  
 1462 vertical stress and flux components is encapsulated by stability functions  $S_x$  that are  
 1463 rational polynomials in nondimensional stratification  $G_H = -l^2q^{-2}N^2$  and shear  $G_M =$   
 1464  $l^2q^{-2}|\partial_z \bar{\mathbf{u}}|^2$  of the general form  $S_x = \text{Num}\{S_x\}/\text{Den}\{S_x\}$ . With the simplifying quasi-  
 1465 equilibrium assumption, dependence on  $G_M$  in the polynomials  $\text{Num}\{S_x\}, \text{Den}\{S_x\}$  is  
 1466 eliminated as in *Kantha and Clayson [1994]*. But with the inclusion of CL vortex force  
 1467 production terms in the ARSM, the closure expressions for stress and flux become:

$$\overline{\mathbf{u}'w'} = -S_M q l \partial_z \bar{\mathbf{u}} - S_M^S q l \partial_z \mathbf{u}^S, \quad (\text{A.18})$$

$$\overline{w'\theta'} = -S_H q l \partial_z \bar{\theta}, \quad (\text{A.19})$$

1468 with corresponding expressions for other scalar components. Here the new eddy coeffi-  
 1469 cient  $K_M^S = S_M^S q l$  directs momentum flux down the Stokes gradient  $\partial_z \mathbf{u}^S$  and the nu-  
 1470 merators and denominators of the stability functions are now polynomials in new nondi-  
 1471 mensional forcing  $G_V = l^2q^{-2}\partial_z \bar{\mathbf{u}} \cdot \partial_z \mathbf{u}^S$  and  $G_S = l^2q^{-2}|\partial_z \mathbf{u}^S|^2$  as well as on  $G_H$ . An  
 1472 important modification in *Harcourt [2015]* near the surface is that all appearances of the  
 1473 Stokes shear in  $G_V$ ,  $G_S$ , and  $\overline{\mathbf{u}'w'}$  are modified by  $\partial_z \mathbf{u}^S \rightarrow (1 - f_z^S)\partial_z \mathbf{u}^S$ , where  $f_z^S$  is a  
 1474 surface proximity function that decays with depth from  $f_z^S = 1$  at the surface. Expressions  
 1475 for  $f_z^S$ ,  $S_H$ ,  $S_M$ , and  $S_M^S$  are those provided in *Harcourt [2015]*.

1476 The implementation of SMCLT-H15 is translated into the native notation of GOTM  
 1477 (where e.g.,  $G_M$  and  $G_H$  are supplanted by  $\alpha_M$  and  $-\alpha_N$ , respectively), but effectively  
 1478 includes the following components: (i) The TKE and  $q^2l$  equations were modified after  
 1479 equations (5)-(6) in *Harcourt [2015]* to include CL vortex production, with coefficient  
 1480  $E_6 = 6.0$  applied to the new  $q^2l$  source term; (ii) Routines computing the nondimensional  
 1481 forcing functions were modified to compute  $G_V$  and  $G_S$ , modified by  $(1 - f_z^S)$  and  $(1 -$   
 1482  $f_z^S)^2$ , respectively; (iii) Routines computing the evolution of horizontal momentum  $\partial_t \bar{\mathbf{u}}$   
 1483 were modified to include a body force  $\partial_z [K_M^S(1 - f_z^S)\partial_z \mathbf{u}^S]$  due to momentum flux down  
 1484 the Stokes gradient; and (iv) A new subroutine to compute the stability functions  $S_H$ ,  $S_M$ ,  
 1485 and  $S_M^S$ .

1486 To improve SMCLT-H15 performance under variable and convectively unstable forc-  
 1487 ing, the crude limiters (e.g.  $G_H \leq G_H^{\max}$ ) have been replaced by formulations more spe-  
 1488 cific to the functional dependence of the stability functions on the turbulence timescale  
 1489  $l/q$ . Under both unstable and stable conditions, the limiters are applied consistently across  
 1490 nondimensional forcings  $G_H$ ,  $G_V$ ,  $G_S$ , and  $G_M$  by limiting the turbulence timescale  $l/q$   
 1491 to values for which the ARSM is able to provide realizable values for Reynolds tensor  
 1492 components and associated stability functions that are both physically possible and nu-  
 1493 merically stable. To do this  $l^2/q^2$  is rescaled to the equilibrium curve as  $r_{\lim}l^2/q^2$  when  
 1494 the distance from the origin to  $[G_H, G_V]$  exceeds the distance in the same direction to the  
 1495 equilibrium curve for  $G_S = G_M = 0$  in the  $G_H$  -  $G_V$  plane. Here  $r_{\lim}$  is a coefficient be-  
 1496 between 0 and 1. The equilibrium curve is approximated by shifting the curve defined by the  
 1497 zero of the  $S_H$  denominator  $\text{Den}\{S_H\} = f(G_H, G_V)$  in equation (33b) of *Harcourt [2015]*  
 1498 by  $\Delta G_H = 0.003$  and  $\Delta G_V = 0.006$ , and solving for  $r_{\lim}$  as the smallest positive root of  
 1499  $f(r_{\lim}G_H + \Delta G_H, r_{\lim}G_V + \Delta G_V) = 0$ . An additional limiter is applied for unstable  $G_H > 0$   
 1500 to control oscillations that arise when the closure is able to support the same scalar flux  
 1501 with two different scalar gradients. To avoid this, the requirement  $\partial(S_H/G_H)/\partial G_H > 1$  of  
 1502 [*Burchard and Deleersnijder, 2001*] is approximated by rescaling the nondimensional forcing  
 1503 by  $r_{\lim}$  whenever the the solution to  $f(2r_{\lim}G_H, r_{\lim}G_V) = 0$  gives  $0 < r_{\lim} < 1$ . This  
 1504 limiter unfortunately constrains the nondimensional forcing away from the equilibrium  
 1505 state under strongly unstable convective conditions, effectively reducing the  $O(1)$  stability

functions by up to half. However, concerns over the questionable physical basis for this should not distract from the larger problem with SMC's under these conditions, namely the lack of non-local buoyancy fluxes. It appears likely that adding non-local fluxes to quasi-equilibrium SMCs could serendipitously alleviate the need for this second limiter.

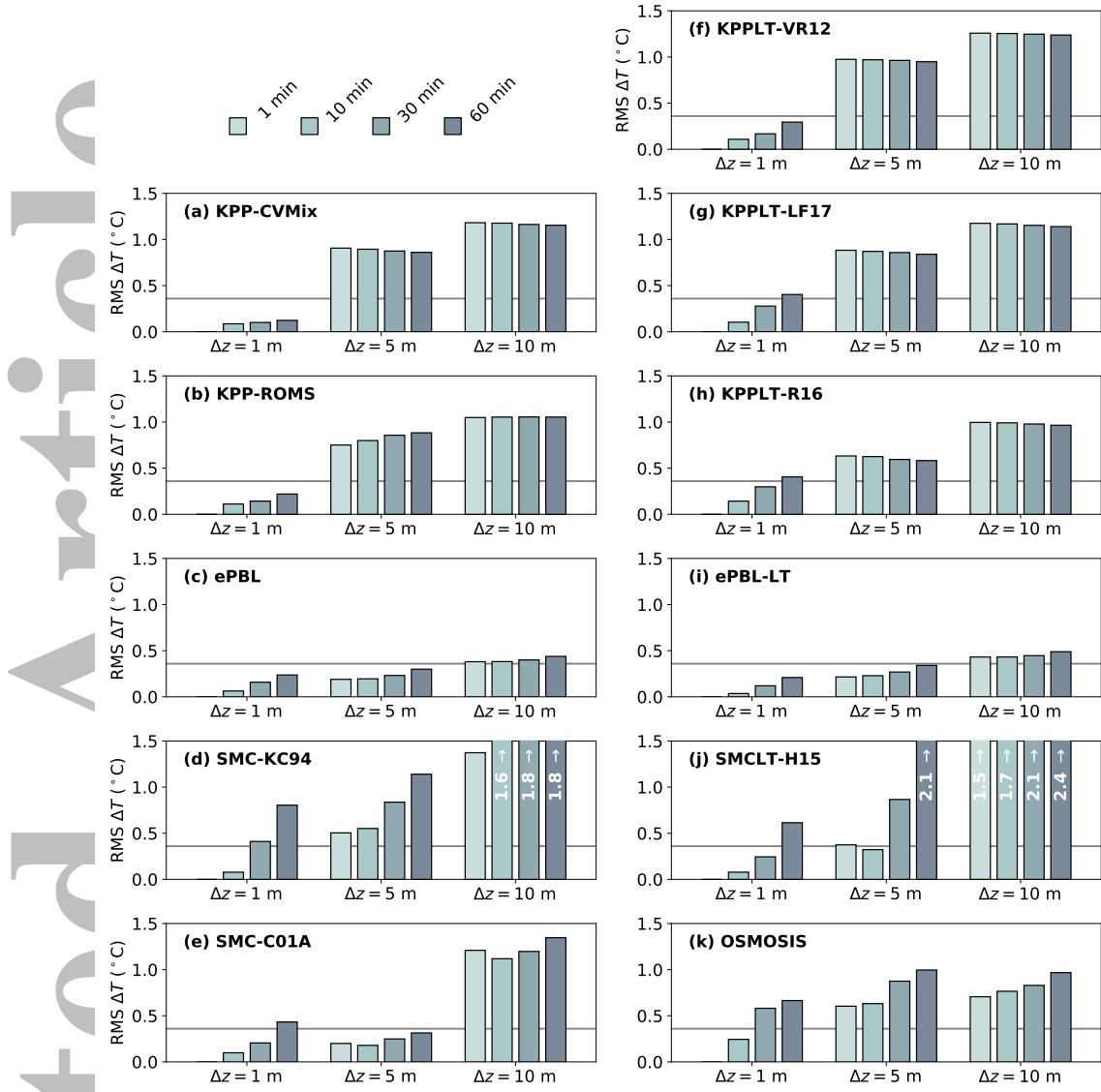
On the unstable side, there is no equilibrium state for a given set of dimensional forcing, and turbulence is only maintained by transport divergence, a feature omitted from the ARSM. Under these typically stratified conditions, the timescale is limited by rescaling  $l^2/q^2$  to  $0.28/N^2$  whenever  $G_H < -0.28$ . This limit is applied only to rescaling the nondimensional forcing that determines stability functions, and it is combined with increasing the coefficient of buoyancy production in the  $q^2 l$  equation to  $E_3 = 5.0$ . This differs from the existing length limiter option in GOTM5, which directly restricts  $l$  in the dynamic prediction of  $q^2 l$  and  $ql$  multiplying stability functions to  $l < q\sqrt{0.28/N^2}$ . The existing limiter can produce reasonable results when combined with the default  $E_3 = 1.8$ , but spreading of the thermocline at the mixed layer base under strong shear is different, possibly less realistic.

Initial difficulties with implementing SMCLT-H15 in GOTM5 were traced to a problem in the existing implementation of the  $q^2 l$  equation for both weak and quasi-equilibrium  $q^2 - q^2 l$  SMC models in GOTM5: Using a predicted value of  $q^2$  to determine the dissipation term for the concurrent advance of the  $q^2 l$  equation generated artificial instabilities requiring very short time steps for stability.

## B: Sensitivity to Vertical Resolution and Time Step

To test the sensitivity of each scheme to the vertical resolution and time step, the simulations at ocean stations were conducted with three different vertical grid spacings of  $\Delta z = [1, 5, 10]$  m, and four different time steps of  $\Delta t = [1, 10, 30, 60]$  min, in total 12 configurations. These configurations span the range of vertical grid spacing and time step commonly used in global and regional OGCMs. The simulated time evolution of the temperature profile with each configuration was compared with that with the finest grid spacing and smallest time step ( $\Delta z = 1$  m,  $\Delta t = 1$  min). The RMS differences for simulations with different schemes in the OCS-Papa case are shown in Figure B.1. The same analyses were performed for OSMOSIS-Winter and OSMOSIS-Spring cases, but are not shown here as the results are qualitatively similar. Note that the RMS differences are computed separately for each scheme against its finest resolution run, which differ substantially among schemes (standard deviation shown by the gray line).

The first conclusion from Figure B.1 is that most schemes appear to be less sensitive to changes in time step than to changes in vertical resolution. This is expected as the mixed layer depth at OCS-Papa ranges from around 20 m to 60 m (Figure 4) and coarsening the vertical resolution from 1 m to 10 m significantly reduces the available grid points within the mixed layer. In particular, the performances of SMC-KC94 and SMCLT-H15 degrade significantly with a vertical grid spacing of  $\Delta z = 10$  m or a time step of  $\Delta t = 60$  min, especially when mixed layer is shallow (not shown). SMC-C01A is much less sensitive to changes in both time step and vertical resolution than SMC-KC94 and SMCLT-H15. The performances of all variants of KPP show some robustness to long time steps even at  $\Delta t = 60$  min, especially when the vertical resolution is coarse. But they also degrade when coarsening the vertical resolution. KPP-ROMS appears to be more sensitive to vertical resolution and time step than KPP-CVMix, highlighting the influences of the detailed formulation of a scheme. The sensitivity of OSMOSIS is in between the KPP variants and the SMC variants. The ePBL and ePBL-LT schemes show the least sensitivity to vertical resolution and time step of all schemes considered here, which is one of the key motivations for the ePBL framework [Reichl and Hallberg, 2018].



**Figure B.1.** Sensitivity to vertical resolution and time step for different schemes in the OCS-Papa case.

Each panel shows the root-mean-square (RMS) differences of the simulated temperature ( $^{\circ}\text{C}$ ) using different vertical resolutions and time steps as compared to the finest resolution simulation ( $\Delta t = 1 \text{ min}$ ,  $\Delta z = 1 \text{ m}$ ). Results with three different vertical grid spacings ( $\Delta z = [1, 5, 10] \text{ m}$ , x-axis) and four different time steps ( $\Delta t = [1, 10, 30, 60] \text{ min}$ , bars) are shown. The gray line in each panel marks the standard deviation of the simulated temperature ( $0.47 ^{\circ}\text{C}$ ) across all schemes with the finest resolution.

### C: Notes on the Regime Diagram

The regime diagram in Figure 1 is intended to take advantage of the relative familiarity of such figures since *Belcher et al.* [2012]. Other papers have adapted this combination of log-log regime diagram with probability density superimposed as well [e.g. *Li and Fox-Kemper*, 2017]. However, in the years since *Belcher et al.* [2012] some deficiencies or potentially misleading aspects of this style of figure have been pointed out, and this appendix will attempt to aid the reader in interpreting the figure. In addition, some

1562 key improvements, such as mapping the location of LES studies and including the stable  
 1563 buoyancy forcing regime, are implemented here and these also deserve some mention.

1564 One of the most difficult aspects of creating the original figure in *Belcher et al.*  
 1565 [2012] while also keeping the  $h/L_L$  scaling approach was that  $h$  needed to be extracted at  
 1566 high temporal resolution from observations. The choice in *Belcher et al.* [2012] to present  
 1567 only the Southern Ocean data was motivated by this concern. In this work, this key disad-  
 1568 vantage can be easily avoided because all of the parameterization schemes and LES results  
 1569 can be mined for a suitable  $h$ . After some discussion and experimentation, the  $h$  used is  
 1570 the boundary layer depth diagnosed in KPP-CVMix instead of a density threshold-based  
 1571 mixed layer to estimate the parameter  $h/L_L$  as the active mixing layer (a.k.a. turbulent  
 1572 boundary layer) is much shallower than the mixed layer under stable surface forcing con-  
 1573 ditions [Pearson et al., 2015]. Using the boundary layer depth from other schemes using  
 1574 a diffusivity threshold [e.g., Noh and Lee, 2008] yields similar results. This approach is  
 1575 closer in spirit to the asymptotics in [Grant and Belcher, 2009] than using the mixed layer  
 1576 depth as well.

1577 A second potential point of confusion is how the probabilities are calculated from  
 1578 the area and what the notion of distance means in Figure 1. Figure C.1 shows a linear  
 1579 space version of Figure 1, with black points scattered through the domain to illustrate each  
 1580 of the 3-hourly JRA55-do data. A joint probability ( $p$ ) of attaining particular values of  
 1581  $La_t \in [a_1, a_2]$  and  $h/L_L \in [b_1, b_2]$  is the integral in linear space of the joint probability  
 1582 density function ( $\rho$ ):

$$p\left(a_1 \leq La_t \leq a_2, b_1 \leq \frac{h}{L_L} \leq b_2\right) = \int_{a_1}^{a_2} \int_{b_1}^{b_2} \rho\left(La_t, \frac{h}{L_L}\right) dLa_t d\left(\frac{h}{L_L}\right). \quad (\text{C.1})$$

1583 The probability of whether the logarithms fall within the range of  $\log_{10}(La_t) \in [c_1, c_2]$  and  
 1584  $\log_{10}(h/L_L) \in [d_1, d_2]$  is a different function:

$$\begin{aligned} & p\left(c_1 \leq \log_{10}(La_t) \leq c_2, d_1 \leq \log_{10}\left(\frac{h}{L_L}\right) \leq d_2\right) \\ &= \int_{c_1}^{c_2} \int_{d_1}^{d_2} \rho\left(\log_{10}(La_t), \log_{10}\left(\frac{h}{L_L}\right)\right) d\log_{10}(La_t) d\log_{10}\left(\frac{h}{L_L}\right). \end{aligned} \quad (\text{C.2})$$

1585 However, when taking appropriate bounds, i.e.,  $[c_1, c_2] = [\log_{10}(a_1), \log_{10}(a_2)]$  and  $[d_1, d_2] =$   
 1586  $[\log_{10}(b_1), \log_{10}(b_2)]$ , these two expressions are equivalent. It is clear from Figure C.1 that  
 1587 the interpretation of whether or not values at near-neutral buoyancy forcing are within the  
 1588 highly probable range depends sensitively on this choice. On the other hand, as the log-  
 1589 arithm is a monotonic function of positive real numbers, there is good agreement on the  
 1590 location of the joint probability density contours on the large magnitude ends of the distri-  
 1591 bution, the trouble is on the low magnitude of  $h/L_L$  ends. Relatedly, considering a point  
 1592 as “distant” from the high probability range depends sensitively on whether one uses log  
 1593 axes or not. The scatter points with  $h/L_L$  near 100 or  $10^2$  are clearly many, but are they  
 1594 close or far to those with a value at 1000 or  $10^3$ ?

1595 A third point of confusion is that negative (stable)  $h/L_L$  values are not geometri-  
 1596 cally or statistically connected to positive (unstable) values in the expected manner, where  
 1597 the negative values are located below the positive ones on a coordinate axis and the joint  
 1598 probability density functions are calculated including all of the data rather than the posi-  
 1599 tive and negative subsets separately. Instead, they are laid out side-by-side in two panels  
 1600 with the absolute value of  $h/L_L$  increasing upward in both cases. This choice is made for  
 1601 easier comparison between the stable and unstable cases, although it is unnecessary in the  
 1602 linear axis case (see Figure C.1).

1603 A fourth point of confusion is why there is any lower bound to  $h/L_L$  in the log-log  
 1604 version of the figure? Why not values near  $10^{-6}$ ? There are many physical reasons per-  
 1605 haps, but a key one is the rate of sampling and the diurnal cycle. Typical diurnal cycles

of heating and cooling are positive many hundreds of  $\text{W m}^{-2}$  during the day and negative by nearly the same amount at night, with an imbalance averaged over the whole day much lower than the peaks. The time when zero forcing occurs is relatively brief (even though the diurnal layer may not be persistent enough to fully mix the mixed layer). JRA55-do has 3-hourly sampling, and this limits the ability of the dataset and analysis method to sample these morning and evening transitions near zero. This problem is also appreciated in atmospheric studies [e.g. *Harvey et al.*, 2015].

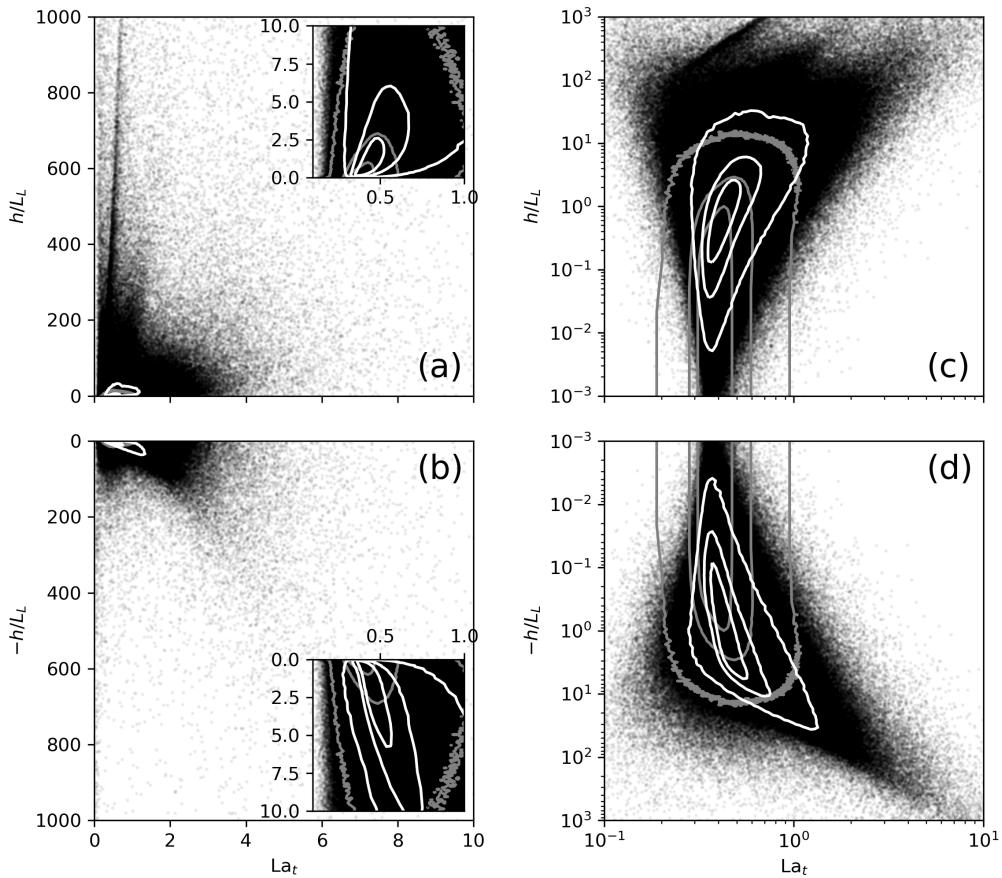
Finally, given these issues why stick with the log-log axes in Figure 1? Because we are building upon the asymptotic multiscale framework traditional in Langmuir studies [e.g. *Craik and Leibovich*, 1976; *McWilliams et al.*, 1997; *Grant and Belcher*, 2009; *Zhang et al.*, 2015]. In this framing it is not so important how small a small parameter is, but at what *order* it enters the dynamics. Thus, logarithms of small parameters are preferred in the probability space to indicate distance in asymptotic order as linear distances. When two data are far apart in Figure 1 they do not just differ in forcing magnitude, they might be expected to have different ordering of asymptotics and thus vastly different dynamics depending on different distinguished force balances.

## 1622 D: Source Code and Data

The source code used here is hosted on GitHub ([github.com/qingli411/gotm](https://github.com/qingli411/gotm)), which builds on existing code bases of GOTM5, but has been significantly extended to include six Langmuir schemes. An interface to read in wave data in various formats and compute the Stokes drift profile is implemented in the code, so that other Langmuir schemes can be easily incorporated and compared with the existing ones. Modifications to CVMix including the three variants of KPP with Langmuir turbulence described in Appendix A.1 have been merged into version v0.94b-beta of CVMix on GitHub ([github.com/CVMix/CVMix-src](https://github.com/CVMix/CVMix-src)). The test suite is also hosted on GitHub ([github.com/qingli411/gotmwork](https://github.com/qingli411/gotmwork)), which includes the initial and surface forcing data for different scenarios, tools to set up and run the tests, and scripts for data analysis and visualization. The archived version of the code as used in this paper and the forcing data are available at doi.org/10.26300/mknw-3842.

## 1634 Acknowledgments

We sincerely thank Hans Burchard and an anonymous reviewer for their thorough and very useful input on this manuscript, whose comments greatly helped to clarify our presentation. We thank Drs. Sonya Legg and Marion Alberty for useful comments on a draft of this manuscript. We thank Niki Zadeh at GFDL for providing the single-column version of the ePBL code as used in this study. This research was supported in part by the Kavli Institute for Theoretical Physics through the National Science Foundation (NSF) under Grant No. NSF PHY-1748958. This research was supported as part of the Energy Exascale Earth System Model (E3SM) project, funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research. Support from NSF 1258907 and 1350795 and ONR N00014-17-1-2393 to BF-K and QL is gratefully acknowledged. BGR and AJA acknowledge support from the Cooperative Institute for Climate Science and the Cooperative Institute for Modeling the Earth's System at Princeton University. BGR also acknowledges support from the Carbon Mitigation Initiative through Princeton's Environmental Institute at Princeton University. JCM and PW acknowledge support from ONR N00014-14-1-0626 and N00014-15-1-2645. TK acknowledges support from NSF Grant OCE-1634578. RRH and ZZ acknowledge support from NSF 1558459 and 1756115, and from ONR N00014-15-1-2308 and N00014-17-1-2859. BP acknowledges support from ONR N00014-17-1-2963. The National Center for Atmospheric Research is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977. All source code and data required to set up the simulations and reconstruct the figures in this paper are publicly available at doi.org/10.26300/mknw-3842.



**Figure C.1.** Joint distribution of  $\text{La}_t$  and  $h/L_L$  in (a, b) linear and (c, d) log-log spaces, estimated from 3-hourly output of JRA55-do forced GOTM5 simulations over 12 months. Black scatters show the raw data points with opacity 0.05. White contours show the joint PDF of  $\log_{10}(\text{La}_t)$  and  $\log_{10}(h/L_L)$  (computed separately for positive and negative  $h/L_L$ ) projected on linear and log-log spaces, with the latter being the contours shown in Figure 1. Gray contours show the joint PDF of  $\text{La}_t$  and  $h/L_L$  projected on linear and log-log spaces. For clarity only the isolines of joint PDF that enclose 30%, 60% and 90% of all instances centered at the highest PDF are shown. Insets in panels (a) and (b) highlight the regions of parameter space with the highest PDF.

## References

- Agrawal, Y. C., E. A. Terray, M. A. Donelan, P. A. Hwang, A. J. Williams, W. M. Drennan, K. K. Kahma, and S. A. Krtaigorodskii (1992), Enhanced dissipation of kinetic energy beneath surface waves, *Nature*, 359(6392), 219–220, doi:10.1038/359219a0.
- Ali, A., K. H. Christensen, Ø. Breivik, M. Malila, R. P. Raj, L. Bertino, E. P. Chassignet, and M. Bakhoday-Paskyabi (2019), A comparison of Langmuir turbulence parameterizations and key wave effects in a numerical model of the North Atlantic and Arctic Oceans, *Ocean Modelling*, 137, 76–97, doi:10.1016/j.ocemod.2019.02.005.
- Axell, L. B. (2002), Wind-driven internal waves and Langmuir circulations in a numerical ocean model of the southern Baltic Sea, *Journal of Geophysical Research*, 107(C11), 3204, doi:10.1029/2001JC000922.

- 1667 Ayotte, K. W., P. P. Sullivan, A. Andrén, S. C. Doney, A. A. M. Holtslag, W. G. Large,  
 1668 J. C. McWilliams, C.-H. Moeng, M. J. Otte, J. J. Tribbia, and J. C. Wyngaard (1996),  
 1669 An evaluation of neutral and convective planetary boundary-layer parameterizations rel-  
 1670 ative to large eddy simulations, *Boundary-Layer Meteorology*, 79(1-2), 131–175, doi:  
 1671 10.1007/BF00120078.
- 1672 Babanin, A. V., and B. K. Haus (2009), On the existence of water turbulence induced by  
 1673 nonbreaking surface waves, *Journal of Physical Oceanography*, 39(10), 2675–2679, doi:  
 1674 10.1175/2009JPO4202.1.
- 1675 Beare, R. J., M. K. Macvean, A. A. M. Holtslag, J. Cuxart, I. Esau, J.-C. Golaz, M. A.  
 1676 Jimenez, M. Khairoutdinov, B. Kosovic, D. Lewellen, T. S. Lund, J. K. Lundquist,  
 1677 A. McCabe, A. F. Moene, Y. Noh, S. Raasch, and P. Sullivan (2006), An intercompar-  
 1678 ision of large-eddy simulations of the stable boundary layer, *Boundary-Layer Meteorol-  
 1679 ogy*, 118(2), 247–272, doi:10.1007/s10546-004-2820-6.
- 1680 Belcher, S. E., A. L. M. Grant, K. E. Hanley, B. Fox-Kemper, L. Van Roekel, P. P. Sulli-  
 1681 van, W. G. Large, A. Brown, A. Hines, D. Calvert, A. Rutgersson, H. Pettersson, J.-R.  
 1682 Bidlot, P. A. E. M. Janssen, and J. A. Polton (2012), A global perspective on Langmuir  
 1683 turbulence in the ocean surface boundary layer, *Geophysical Research Letters*, 39(18),  
 1684 L18,605, doi:10.1029/2012GL052932.
- 1685 Boyer, T. P., J. I. Antonov, O. K. Baranova, H. E. Garcia, D. R. Johnson, A. V. Mis-  
 1686 honov, T. D. O'Brien, D. Seidov, Smolyar, I. (Igor), M. M. Zweng, C. R. Paver, R. A.  
 1687 Locarnini, J. R. Reagan, C. Coleman, and A. Grodsky (2013), World ocean database  
 1688 2013., U.S. Department of Commerce, National Oceanic and Atmospheric Adminstration,  
 1689 National Environmental Satellite, Data, and Information Service, National Oceanographic  
 1690 Data Center, Ocean Climate Laboratory., doi:10.7289/v5nz85mt.
- 1691 Breivik, O., J.-R. Bidlot, and P. A. E. M. Janssen (2016), A Stokes drift approximation  
 1692 based on the Phillips spectrum, *Ocean Modelling*, 100, 49–56, doi:10.1016/j.ocemod.  
 1693 2016.01.005.
- 1694 Burchard, H., and K. Bolding (2001), Comparative analysis of four second-moment tur-  
 1695 bulence closure models for the oceanic mixed layer, *Journal of Physical Oceanography*,  
 1696 31(8), 1943–1968, doi:10.1175/1520-0485(2001)031<1943:CAOFSM>2.0.CO;2.
- 1697 Burchard, H., and E. Deleersnijder (2001), Stability of algebraic non-equilibrium second-  
 1698 order closure models, *Ocean Modelling*, 3(1), 33–50, doi:10.1016/S1463-5003(00)  
 1699 00016-0.
- 1700 Burchard, H., P. D. Craig, J. R. Gemmrich, H. van Haren, P.-P. Mathieu, H. E. M. Meier,  
 1701 W. A. M. N. Smith, H. Prandke, T. P. Rippeth, E. D. Skillingstad, W. D. Smyth,  
 1702 D. J. S. Welsh, and H. W. Wijesekera (2008), Observational and numerical modeling  
 1703 methods for quantifying coastal ocean turbulence and mixing, *Progress in Oceanogra-  
 1704 phy*, 76(4), 399–442, doi:10.1016/j.pocean.2007.09.005.
- 1705 Callies, J., and R. Ferrari (2018), Baroclinic instability in the presence of convection,  
 1706 *Journal of Physical Oceanography*, 48(1), 45–60, doi:10.1175/JPO-D-17-0028.1.
- 1707 Canuto, V. M., A. Howard, Y. Cheng, and M. S. Dubovikov (2001), Ocean turbulence.  
 1708 Part I: One-point closure model—Momentum and heat vertical diffusivities, *Journal  
 1709 of Physical Oceanography*, 31(6), 1413–1426, doi:10.1175/1520-0485(2001)031<1413:  
 1710 OTPIOP>2.0.CO;2.
- 1711 Cavalieri, L., B. Fox-Kemper, and M. Hemer (2012), Wind waves in the coupled cli-  
 1712 mate system, *Bulletin of the American Meteorological Society*, 93(11), 1651–1661, doi:  
 1713 10.1175/BAMS-D-11-00170.1.
- 1714 Chen, S., F. Qiao, C. Huang, and Z. Song (2018), Effects of the non-breaking surface  
 1715 wave-induced vertical mixing on winter mixed layer depth in subtropical regions, *Jour-  
 1716 nal of Geophysical Research: Oceans*, 123(4), 2934–2944, doi:10.1002/2017JC013038.
- 1717 Craik, A. D. D., and S. Leibovich (1976), A rational model for Langmuir circulations,  
 1718 *Journal of Fluid Mechanics*, 73(03), 401–426, doi:10.1017/S0022112076001420.
- 1719 Damerell, G. M., K. J. Heywood, A. F. Thompson, U. Binetti, and J. Kaiser (2016),  
 1720 The vertical structure of upper ocean variability at the Porcupine Abyssal Plain dur-

- ing 2012–2013, *Journal of Geophysical Research: Oceans*, 121(5), 3075–3089, doi: 10.1002/2015JC011423.
- Danabasoglu, G., W. G. Large, J. J. Tribbia, P. R. Gent, B. P. Briegleb, and J. C. McWilliams (2006), Diurnal coupling in the tropical oceans of CCSM3, *Journal of Climate*, 19(11), 2347–2365, doi:10.1175/JCLI3739.1.
- D'Asaro, E. A. (1985), The energy flux from the wind to near-inertial motions in the surface mixed layer, *Journal of Physical Oceanography*, 15(8), 1043–1059, doi:10.1175/1520-0485(1985)015<1043:TEFFTW>2.0.CO;2.
- D'Asaro, E. A. (2001), Turbulent vertical kinetic energy in the ocean mixed layer, *Journal of Physical Oceanography*, 31(12), 3530–3537, doi:10.1175/1520-0485(2002)031<3530:TVKEIT>2.0.CO;2.
- D'Asaro, E. A. (2014), Turbulence in the upper-ocean mixed layer., *Annual Review of Marine Science*, 6, 101–15, doi:10.1146/annurev-marine-010213-135138.
- D'Asaro, E. A., J. Thomson, A. Y. Shcherbina, R. R. Harcourt, M. F. Cronin, M. A. Hemer, and B. Fox-Kemper (2014), Quantifying upper ocean turbulence driven by surface waves, *Geophysical Research Letters*, 41(1), 102–107, doi:10.1002/2013GL058193.
- de Boyer Montégut, C., C. Madec, A. S. Fischer, A. Lazar, and D. Iudicone (2004), Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology, *Journal of Geophysical Research*, 109(C12), C12,003, doi:10.1029/2004JC002378.
- Fairall, C. W., E. F. Bradley, D. P. Rogers, J. B. Edson, and G. S. Young (1996), Bulk parameterization of air-sea fluxes for tropical ocean-global atmosphere coupled-ocean atmosphere response experiment, *Journal of Geophysical Research: Oceans*, 101(C2), 3747–3764, doi:10.1029/95JC03205.
- Fan, Y., and S. M. Griffies (2014), Impacts of parameterized Langmuir turbulence and nonbreaking wave mixing in global climate simulations, *Journal of Climate*, 27(12), 4752–4775, doi:10.1175/JCLI-D-13-00583.1.
- Galperin, B., L. H. Kantha, S. Hassid, and A. Rosati (1988), A quasi-equilibrium turbulent energy model for geophysical flows, *Journal of the Atmospheric Sciences*, 45(1), 55–62, doi:10.1175/1520-0469(1988)045<0055:AQETEM>2.0.CO;2.
- Gargett, A., J. Wells, A. E. Tejada-Martínez, and C. E. Grosch (2004), Langmuir supercells: A mechanism for sediment resuspension and transport in shallow seas, *Science*, 306(5703), 1925–1928, doi:10.1126/science.1100849.
- Grant, A. L. M., and S. E. Belcher (2009), Characteristics of Langmuir turbulence in the ocean mixed layer, *Journal of Physical Oceanography*, 39(8), 1871–1887, doi:10.1175/2009JPO4119.1.
- Griffies, S. M., A. Biastoch, C. W. Böning, F. O. Bryan, G. Danabasoglu, E. P. Chassignet, M. H. England, R. Gerdes, H. Haak, R. W. Hallberg, W. Hazeleger, J. Jungclaus, W. G. Large, G. Madec, A. Pirani, B. L. Samuels, M. Scheinert, A. S. Gupta, C. a. Severeijns, H. L. Simmons, A. M. Treguier, M. Winton, S. G. Yeager, and J. Yin (2009), Coordinated Ocean-ice Reference Experiments (COREs), *Ocean Modelling*, 26(1-2), 1–46, doi:10.1016/j.ocemod.2008.08.007.
- Griffies, S. M., M. Levy, A. J. Adcroft, G. Danabasoglu, R. W. Hallberg, D. Jacobsen, W. Large, and T. Ringler (2015), Theory and numerics of the Community Ocean Vertical Mixing (CVMix) project, *Tech. rep.*
- Griffies, S. M., G. Danabasoglu, P. J. Durack, A. J. Adcroft, V. Balaji, C. W. Böning, E. P. Chassignet, E. Curchitser, J. Deshayes, H. Drange, B. Fox-Kemper, P. J. Gleckler, J. M. Gregory, H. Haak, R. W. Hallberg, P. Heimbach, H. T. Hewitt, D. M. Holland, T. Ilyina, J. H. Jungclaus, Y. Komuro, J. P. Krasting, W. G. Large, S. J. Marshall, S. Masina, T. J. McDougall, A. J. G. Nurser, J. C. Orr, A. Pirani, F. Qiao, R. J. Stouffer, K. E. Taylor, A. M. Treguier, H. Tsujino, P. Uotila, M. Valdivieso, Q. Wang, M. Winton, and S. G. Yeager (2016), OMIP contribution to CMIP6: Experimental and diagnostic protocol for the physical component of the Ocean Model Intercomparison Project, *Geoscientific Model Development*, 9(9), 3231–3296, doi:<https://doi.org/10.5194/gmd-9-3231-2016>

- 1775 gmd-9-3231-2016.
- 1776 Hamlington, P. E., L. P. Van Roekel, B. Fox-Kemper, K. Julien, and G. P. Chini (2014),  
1777 Langmuir-submesoscale interactions: Descriptive analysis of multiscale frontal spin-  
1778 down simulations, *Journal of Physical Oceanography*, 44(9), 2249–2272, doi:10.1175/  
1779 JPO-D-13-0139.1.
- 1780 Haney, S., B. Fox-Kemper, K. Julien, and A. Webb (2015), Symmetric and geostrophic  
1781 instabilities in the wave-forced ocean mixed layer, *Journal of Physical Oceanography*,  
1782 45(12), 3033–3056, doi:10.1175/JPO-D-15-0044.1.
- 1783 Hanley, K. E., S. E. Belcher, and P. P. Sullivan (2010), A global climatology of wind-  
1784 wave interaction, *Journal of Physical Oceanography*, 40(6), 1263–1282, doi:10.1175/  
1785 2010JPO4377.1.
- 1786 Harcourt, R. R. (2013), A second-moment closure model of Langmuir turbulence, *Journal*  
1787 *of Physical Oceanography*, 43(4), 673–697, doi:10.1175/JPO-D-12-0105.1.
- 1788 Harcourt, R. R. (2015), An improved second-moment closure model of Langmuir turbu-  
1789 lence, *Journal of Physical Oceanography*, 45(4), 84–103, doi:10.1175/JPO-D-14-0046.1.
- 1790 Harcourt, R. R., and E. A. D'Asaro (2008), Large-eddy simulation of Langmuir turbulence  
1791 in pure wind seas, *Journal of Physical Oceanography*, 38(7), 1542–1562, doi:10.1175/  
1792 2007JPO3842.1.
- 1793 Harvey, N. J., R. J. Hogan, and H. F. Dacre (2015), Evaluation of boundary-layer type in  
1794 a weather forecast model utilizing long-term doppler lidar observations, *Quart. J. Roy.*  
1795 *Meteor. Soc.*, 141, 1345–1353.
- 1796 Jackson, L., R. Hallberg, and S. Legg (2008), A parameterization of shear-driven turbu-  
1797 lence for ocean climate models, *Journal of Physical Oceanography*, 38(5), 1033–1053,  
1798 doi:10.1175/2007JPO3779.1.
- 1799 Janssen, P. A. E. M. (1989), Wave-induced stress and the drag of air flow over sea waves,  
1800 *Journal of Physical Oceanography*, 19(6), 745–754, doi:10.1175/1520-0485(1989)  
1801 019<0745:WISATD>2.0.CO;2.
- 1802 Japan Meteorological Agency/Japan (2013), JRA-55: Japanese 55-year reanalysis, daily  
1803 3-hourly and 6-hourly data, doi:10.5065/D6HH6H41, Research Data Archive at the Na-  
1804 tional Center for Atmospheric Research, Computational and Information Systems Labo-  
1805 ratory, Boulder, Colo. (Updated monthly).
- 1806 Kantha, L. H., and C. A. Clayson (1994), An improved mixed layer model for geophysical  
1807 applications, *Journal of Geophysical Research*, 99(C12), 25,235–25,266, doi:10.1029/  
1808 94JC02257.
- 1809 Kantha, L. H., and C. A. Clayson (2004), On the effect of surface gravity waves on  
1810 mixing in the oceanic mixed layer, *Ocean Modelling*, 6(2), 101–124, doi:10.1016/  
1811 S1463-5003(02)00062-8.
- 1812 Kenyon, K. E. (1969), Stokes drift for random gravity waves, *Journal of Geophysical Re-*  
1813 *search*, 74(28), 6991, doi:10.1029/JC074i028p06991.
- 1814 Kudryavtsev, V., V. Shrira, V. Dulov, and V. Malinovsky (2008), On the vertical structure  
1815 of wind-driven sea currents, *Journal of Physical Oceanography*, 38(10), 2121–2144, doi:  
1816 10.1175/2008JPO3883.1.
- 1817 Kukulka, T., and R. R. Harcourt (2017), Influence of Stokes drift decay scale on Lang-  
1818 muir turbulence, *Journal of Physical Oceanography*, 47(7), 1637–1656, doi:10.1175/  
1819 JPO-D-16-0244.1.
- 1820 Kukulka, T., A. J. Plueddemann, J. H. Trowbridge, and P. P. Sullivan (2009), Significance  
1821 of Langmuir circulation in upper ocean mixing: Comparison of observations and simu-  
1822 lations, *Geophysical Research Letters*, 36(10), L10,603, doi:10.1029/2009GL037620.
- 1823 Kukulka, T., A. J. Plueddemann, J. H. Trowbridge, and P. P. Sullivan (2010), Rapid mixed  
1824 layer deepening by the combination of Langmuir and shear instabilities: A case study,  
1825 *Journal of Physical Oceanography*, 40(11), 2381–2400, doi:10.1175/2010JPO4403.1.
- 1826 Kukulka, T., A. J. Plueddemann, and P. P. Sullivan (2013), Inhibited upper ocean restratifi-  
1827 cation in nonequilibrium swell conditions, *Geophysical Research Letters*, 40(14), 3672–  
1828 3676, doi:10.1002/grl.50708.

- 1829 Langmuir, I. (1938), Surface motion of water induced by wind, *Science*, 87(2250), 119–  
1830 123.
- 1831 Large, W. G. (1996), An observational and numerical investigation of the climatological  
1832 heat and salt balances at OWS Papa, *Journal of Climate*, 9(8), 1856–1876, doi:10.1175/  
1833 1520-0442(1996)009<1856:AOANIO>2.0.CO;2.
- 1834 Large, W. G., and J. M. Caron (2015), Diurnal cycling of sea surface temperature, salin-  
1835 ity, and current in the CESM coupled climate model, *Journal of Geophysical Research: Oceans*, 120, 3711–3729, doi:10.1002/2014JC010691.Received.
- 1837 Large, W. G., and S. G. Yeager (2009), The global climatology of an interannually  
1838 varying air-sea flux data set, *Climate Dynamics*, 33(2-3), 341–364, doi:10.1007/  
1839 s00382-008-0441-3.
- 1840 Large, W. G., J. C. Mcwilliams, and S. C. Doney (1994), Oceanic vertical mixing: A re-  
1841 view and a model with a nonlocal boundary layer parameterization, *Reviews of Geo-  
1842 physics*, 32(4), 363–403.
- 1843 Large, W. G., G. Danabasoglu, S. C. Doney, and J. C. McWilliams (1997), Sensitivity  
1844 to surface forcing and boundary layer mixing in a global ocean model: Annual-mean  
1845 climatology, *Journal of Physical Oceanography*, 27(11), 2418–2447.
- 1846 Large, W. G., E. G. Patton, A. K. DuVivier, and P. P. Sullivan (2019), Similarity theory in  
1847 the surface layer of large-eddy simulations of the southern ocean with waves, *Journal of  
1848 Physical Oceanography*, doi:10.1175/JPO-D-18-0066.1, accepted.
- 1849 Lemarié, F., J. Kurian, A. F. Shchepetkin, M. Jeroen Molemaker, F. Colas, and J. C.  
1850 McWilliams (2012), Are there inescapable issues prohibiting the use of terrain-  
1851 following coordinates in climate models?, *Ocean Modelling*, 42, 57–79, doi:10.1016/j.  
1852 ocemod.2011.11.007.
- 1853 Li, Q. (2018), Langmuir turbulence and its effects on global climate, Ph.D. thesis, Brown  
1854 University, Providence, RI, doi:doi.org/10.26300/mj20-3x15.
- 1855 Li, Q., and B. Fox-Kemper (2017), Assessing the effects of Langmuir turbulence on the  
1856 entrainment buoyancy flux in the Ocean Surface Boundary Layer, *Journal of Physical  
1857 Oceanography*, 47(12), 2863–2886, doi:10.1175/JPO-D-17-0085.1.
- 1858 Li, Q., A. Webb, B. Fox-Kemper, A. Craig, G. Danabasoglu, W. G. Large, and M. Verten-  
1859 stein (2016), Langmuir mixing effects on global climate: WAVEWATCH III in CESM,  
1860 *Ocean Modelling*, 103, 145–160, doi:10.1016/j.ocemod.2015.07.020.
- 1861 Li, Q., B. Fox-Kemper, Ø. Breivik, and A. Webb (2017), Statistical models of global  
1862 Langmuir mixing, *Ocean Modelling*, 113, 95–114, doi:10.1016/j.ocemod.2017.03.016.
- 1863 Liu, J., J.-H. Liang, J. C. McWilliams, P. P. Sullivan, Y. Fan, and Q. Chen (2018), Effect  
1864 of planetary rotation on oceanic surface boundary layer turbulence, *Journal of Physical  
1865 Oceanography*, 48(9), 2057–2080, doi:10.1175/JPO-D-17-0150.1.
- 1866 Mazloff, M. R., P. Heimbach, and C. Wunsch (2010), An eddy-permitting Southern  
1867 Ocean state estimate, *Journal of Physical Oceanography*, 40(5), 880–899, doi:10.1175/  
1868 2009JPO4236.1.
- 1869 McWilliams, J. C., and B. Fox-Kemper (2013), Oceanic wave-balanced surface fronts and  
1870 filaments, *Journal of Fluid Mechanics*, 730, 464–490, doi:10.1017/jfm.2013.348.
- 1871 McWilliams, J. C., and P. P. Sullivan (2000), Vertical mixing by Langmuir circulations,  
1872 *Spill Science and Technology Bulletin*, 6(3), 225–237, doi:10.1016/S1353-2561(01)  
1873 00041-X.
- 1874 McWilliams, J. C., P. P. Sullivan, and C.-H. Moeng (1997), Langmuir turbulence in the  
1875 ocean, *Journal of Fluid Mechanics*, 334(1), 1–30.
- 1876 McWilliams, J. C., E. Huckle, and A. F. Shchepetkin (2009), Buoyancy effects in a  
1877 stratified Ekman layer, *Journal of Physical Oceanography*, 39(10), 2581–2599, doi:  
1878 10.1175/2009JPO4130.1.
- 1879 McWilliams, J. C., E. Huckle, J.-H. Liang, and P. P. Sullivan (2012), The wavy Ekman  
1880 layer: Langmuir circulations, breaking waves, and Reynolds stress, *Journal of Physical  
1881 Oceanography*, 42(11), 1793–1816, doi:10.1175/JPO-D-12-07.1.

- 1882 McWilliams, J. C., E. Huckle, J.-H. Liang, and P. P. Sullivan (2014), Langmuir tur-  
 1883 bulence in swell, *Journal of Physical Oceanography*, 44(3), 870–890, doi:10.1175/  
 1884 JPO-D-13-0122.1.
- 1885 Mellor, G. L., and T. Yamada (1974), A hierarchy of turbulence closure models for plane-  
 1886 tary boundary layers, *Journal of the Atmospheric Sciences*, 31, 1791–1806.
- 1887 Mellor, G. L., and T. Yamada (1982), Development of a turbulence closure model for geo-  
 1888 physical fluid problems, *Reviews of Geophysics and Space Physics*, 20(4), 851–875.
- 1889 Min, H. S., and Y. Noh (2004), Influence of the surface heating on Langmuir circulation,  
 1890 *Journal of Physical Oceanography*, 34(12), 2630–2641, doi:10.1175/JPOJPO-2654.1.
- 1891 Moore, J. K., K. Lindsay, S. C. Doney, M. C. Long, and K. Misumi (2013), Ma-  
 1892 rine ecosystem dynamics and biogeochemical cycling in the community earth sys-  
 1893 tem model [CESM1(BGC)]: Comparison of the 1990s with the 2090s under the  
 1894 RCP4.5 and RCP8.5 scenarios, *Journal of Climate*, 26(23), 9291–9312, doi:10.1175/  
 1895 JCLI-D-12-00566.1.
- 1896 Myrhaug, D., H. Wang, and L. E. Holmedal (2014), Stokes drift estimation for deep water  
 1897 waves based on short-term variation of wave conditions, *Coastal Engineering*, 88, 27–  
 1898 32, doi:10.1016/j.coastaleng.2014.01.014.
- 1899 Noh, Y., and W. S. Lee (2008), Mixed and mixing layer depths simulated by an OGCM,  
 1900 *Journal of Oceanography*, 64(2), 217–225, doi:10.1007/s10872-008-0017-1.
- 1901 Noh, Y., H. Ok, E. Lee, T. Toyoda, and N. Hirose (2016), Parameterization of Langmuir  
 1902 circulation in the ocean mixed layer model using LES and its application to the OGCM,  
 1903 *Journal of Physical Oceanography*, 46(1), 57–78, doi:10.1175/JPO-D-14-0137.1.
- 1904 Pearson, B. (2018), Turbulence-induced anti-Stokes flow and the resulting limitations of  
 1905 large-eddy simulation, *Journal of Physical Oceanography*, 48(1), 117–122, doi:10.1175/  
 1906 JPO-D-17-0208.1.
- 1907 Pearson, B., A. L. M. Grant, J. A. Polton, and S. E. Belcher (2015), Langmuir turbulence  
 1908 and surface heating in the ocean surface boundary layer, *Journal of Physical Oceanogra-  
 1909 phy*, 45, 2897–2911, doi:10.1175/JPO-D-15-0018.1.
- 1910 Polton, J. A., and S. E. Belcher (2007), Langmuir turbulence and deeply penetrating jets  
 1911 in an unstratified mixed layer, *Journal of Geophysical Research: Oceans*, 112(9), 1–11,  
 1912 doi:10.1029/2007JC004205.
- 1913 Prandtl, L. (1925), Bericht über die entstehung der turbulenz, *Z. Angew. Math. Mech.*, 5,  
 1914 136–139.
- 1915 Price, J. F., R. a. Weller, and R. Pinkel (1986), Diurnal cycling: Observations and models  
 1916 of the upper ocean response to diurnal heating, cooling, and wind mixing, *Journal of  
 1917 Geophysical Research*, 91(C7), 8411–8427, doi:10.1029/JC091iC07p08411.
- 1918 Qiao, F., Y. Yuan, Y. Yang, Q. Zheng, C. Xia, and J. Ma (2004), Wave-induced mixing  
 1919 in the upper ocean: Distribution and application to a global ocean circulation model,  
 1920 *Geophysical Research Letters*, 31(11), L11,303, doi:10.1029/2004GL019824.
- 1921 Qiao, F., Y. Yuan, J. Deng, D. Dai, and Z. Song (2016), Wave-turbulence interaction-  
 1922 induced vertical mixing and its effects in ocean and climate models, *Philosophical  
 1923 Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, 374(2065),  
 1924 20150,201, doi:10.1098/rsta.2015.0201.
- 1925 Reichl, B. G., and R. Hallberg (2018), A simplified energetics based planetary boundary  
 1926 layer (ePBL) approach for ocean climate simulations., *Ocean Modelling*, 132, 112–129,  
 1927 doi:10.1016/j.ocemod.2018.10.004.
- 1928 Reichl, B. G., and Q. Li (2019), A parameterization with a constrained potential energy  
 1929 conversion rate of vertical mixing due to Langmuir turbulence, *Journal of Physical  
 1930 Oceanography*, doi:10.1175/JPO-D-18-0258.1, in press.
- 1931 Reichl, B. G., D. Wang, T. Hara, I. Ginis, and T. Kukulka (2016), Langmuir turbulence  
 1932 parameterization in tropical cyclone conditions, *Journal of Physical Oceanography*,  
 1933 46(3), 863–886, doi:10.1175/JPO-D-15-0106.1.
- 1934 Rodgers, K. B., O. Aumont, S. E. Mikaloff Fletcher, Y. Plancherel, L. Bopp, C. de Boyer  
 1935 Montégut, D. Iudicone, R. F. Keeling, G. Madec, and R. Wanninkhof (2014), Strong

- 1936 sensitivity of Southern Ocean carbon uptake and nutrient cycling to wind stirring, *Bio-*  
 1937 *geosciences*, 11, 4077–4098, doi:10.5194/bg-11-4077-2014.
- 1938 Rodi, W. (1987), Examples of calculation methods for flow and mixing in stratified  
 1939 fluids, *Journal of Geophysical Research: Oceans*, 92(C5), 5305–5328, doi:10.1029/  
 1940 JC092iC05p05305.
- 1941 Sallée, J.-B., E. Shuckburgh, N. Bruneau, A. J. S. Meijers, T. J. Bracegirdle, Z. Wang, and  
 1942 T. Roy (2013a), Assessment of Southern Ocean water mass circulation and character-  
 1943 istics in CMIP5 models: Historical bias and forcing response, *Journal of Geophysical*  
 1944 *Research: Oceans*, 118(4), 1830–1844, doi:10.1002/jgrc.20135.
- 1945 Sallée, J.-B., E. Shuckburgh, N. Bruneau, A. J. S. Meijers, T. J. Bracegirdle, and Z. Wang  
 1946 (2013b), Assessment of Southern Ocean mixed-layer depths in CMIP5 models: Histori-  
 1947 cal bias and forcing response, *Journal of Geophysical Research: Oceans*, 118(4), 1845–  
 1948 1862, doi:10.1002/jgrc.20157.
- 1949 Schudlich, R. R., and J. F. Price (1998), Observations of seasonal variation in the  
 1950 Ekman layer, *Journal of Physical Oceanography*, 28(6), 1187–1204, doi:10.1175/  
 1951 1520-0485(1998)028<1187:OOSVIT>2.0.CO;2.
- 1952 Schulz, E. W., S. A. Josey, and R. Verein (2012), First air-sea flux mooring measurements  
 1953 in the Southern Ocean, *Geophysical Research Letters*, 39(16), L16,606, doi:10.1029/  
 1954 2012GL052290.
- 1955 Schumann, U., and T. Gerz (1995), Turbulent mixing in stably stratified shear flows, *Jour-  
 1956 nal of Applied Meteorology*, 34(1), 33–48, doi:10.1175/1520-0450-34.1.33.
- 1957 Shchepetkin, A. F., and J. C. McWilliams (2005), The regional oceanic modeling system  
 1958 (ROMS): A split-explicit, free-surface, topography-following-coordinate oceanic model,  
 1959 *Ocean Modelling*, 9(4), 347–404, doi:10.1016/j.ocemod.2004.08.002.
- 1960 Shchepetkin, A. F., and J. C. McWilliams (2009), Computational kernel algorithms  
 1961 for fine-scale, multiprocess, longtime oceanic simulations, in *Handbook of Numer-  
 1962 ical Analysis, Special Volume: Computational Methods for the Atmosphere and the  
 1963 Oceans*, vol. 14, edited by R. M. Temam and J. J. Tribbia, pp. 121–183, Elsevier, doi:  
 1964 10.1016/S1570-8659(08)01202-0.
- 1965 Shimura, T., N. Mori, T. Takemi, and R. Mizuta (2017), Long-term impacts of ocean  
 1966 wave-dependent roughness on global climate systems, *Journal of Geophysical Research: Oceans*, 122(3), 1995–2011, doi:10.1002/2016JC012621.
- 1968 Sinha, N., A. E. Tejada-Martínez, C. Akan, and C. E. Grosch (2015), Toward a K-profile  
 1969 parameterization of langmuir turbulence in shallow coastal shelves, *Journal of Physical*  
 1970 *Oceanography*, 45(12), 2869–2895, doi:10.1175/JPO-D-14-0158.1.
- 1971 Skillingstad, E. D., W. D. Smyth, and G. B. Crawford (2000), Resonant wind-driven mix-  
 1972 ing in the ocean boundary layer, *Journal of Physical Oceanography*, 30(8), 1866–1890,  
 1973 doi:10.1175/1520-0485(2000)030<1866:RWDMIT>2.0.CO;2.
- 1974 Smyth, W. D., E. D. Skillingstad, G. B. Crawford, and H. Wijesekera (2002), Nonlocal  
 1975 fluxes and Stokes drift effects in the K-profile parameterization, *Ocean Dynamics*, 52(3),  
 1976 104–115, doi:10.1007/s10236-002-0012-9.
- 1977 Stokes, G. G. (1847), On the theory of oscillatory waves, *Cambridge Philosophical Soci-  
 1978 ety*, 8, 197–237, doi:10.1017/CBO9780511702242.016.
- 1979 Stommel, H. (1979), Determination of water mass properties of water pumped down from  
 1980 the Ekman layer to the geostrophic flow below., *Proceedings of the National Academy of*  
 1981 *Sciences of the United States of America*, 76(7), 3051–3055, doi:10.1073/pnas.76.7.3051.
- 1982 Sullivan, P. P., J. C. McWilliams, and W. K. Melville (2007), Surface gravity wave effects  
 1983 in the oceanic boundary layer: Large-eddy simulation with vortex force and stochastic  
 1984 breakers, *Journal of Fluid Mechanics*, 593, 405–452, doi:10.1017/S002211200700897X.
- 1985 Sullivan, P. P., L. Romero, J. C. McWilliams, and W. K. Melville (2012), Transient evo-  
 1986 lution of Langmuir turbulence in ocean boundary layers driven by hurricane winds and  
 1987 waves, *Journal of Physical Oceanography*, 42, 1959–1980, doi:10.1175/JPO-D-12-025.  
 1988 1.

- 1989 Sullivan, P. P., J. C. Weil, E. G. Patton, H. J. J. Jonker, and D. V. Mironov (2016), Tur-  
 1990 bulent winds and temperature fronts in large-eddy simulations of the stable atmo-  
 1991 spheric boundary layer, *Journal of the Atmospheric Sciences*, 73(4), 1815–1840, doi:  
 1992 10.1175/JAS-D-15-0339.1.
- 1993 Sutherland, G., K. H. Christensen, and B. Ward (2014), Evaluating Langmuir turbulence  
 1994 parameterizations in the ocean surface boundary layer, *Journal of Geophysical Research: Oceans*, 119(3), 1899–1910, doi:10.1002/2013JC009537.
- 1995 Suzuki, N., and B. Fox-Kemper (2016), Understanding Stokes forces in the wave-averaged  
 1996 equations, *Journal of Geophysical Research: Oceans*, 121(5), 3579–3596, doi:10.1002/  
 1998 2015JC011566.
- 1999 Suzuki, N., B. Fox-Kemper, P. E. Hamlington, and L. P. Van Roekel (2016), Surface  
 2000 waves affect frontogenesis, *Journal of Geophysical Research: Oceans*, 121(5), 3597–  
 2001 3624, doi:10.1002/2015JC011563.
- 2002 Takaya, Y., J.-R. Bidlot, A. C. M. Beljaars, and P. A. E. M. Janssen (2010), Refinements  
 2003 to a prognostic scheme of skin sea surface temperature, *Journal of Geophysical Research: Oceans*, 115(C6), C06,009, doi:10.1029/2009JC005985.
- 2005 Teixeira, M. A. C. (2012), The influence of Langmuir turbulence on the scaling for  
 2006 the dissipation rate in the oceanic boundary layer, *Journal of Geophysical Research: Oceans*, 117(C5), C05,015, doi:10.1029/2011JC007235.
- 2008 Teixeira, M. A. C. (2018), A model for the wind-driven current in the wavy oceanic sur-  
 2009 face layer: Apparent friction velocity reduction and roughness length enhancement,  
 2010 *Journal of Physical Oceanography*, doi:10.1175/JPO-D-18-0086.1.
- 2011 Teixeira, M. A. C., and S. E. Belcher (2010), On the structure of Langmuir turbulence,  
 2012 *Ocean Modelling*, 31(3–4), 105–119, doi:10.1016/j.ocemod.2009.10.007.
- 2013 Terray, E., M. Donelan, Y. Agrawal, W. Drennan, K. Kahma, A. Williams, P. Hwang, and  
 2014 S. Kitaigorodskii (1996), Estimates of kinetic energy dissipation under breaking waves,  
 2015 *Journal of Physical Oceanography*, 26(5), 792–807, doi:10.1175/1520-0485(1996)  
 026<0792:EOKEDU>2.0.CO;2.
- 2017 Thompson, A. F., A. Lazar, C. Buckingham, A. C. Naveira Garabato, G. M. Damerell,  
 2018 and K. J. Heywood (2016), Open-ocean submesoscale motions: A full seasonal cycle of  
 2019 mixed layer instabilities from gliders, *Journal of Physical Oceanography*, 46(4), 1285–  
 2020 1307, doi:10.1175/JPO-D-15-0170.1.
- 2021 Thorpe, S. A. (2004), Langmuir circulation, *Annual Review of Fluid Mechanics*, 36(1), 55–  
 2022 79, doi:10.1146/annurev.fluid.36.052203.071431.
- 2023 Troen, I. B., and L. Mahrt (1986), A simple model of the atmospheric boundary layer;  
 2024 sensitivity to surface evaporation, *Boundary-Layer Meteorology*, 37(1–2), 129–148, doi:  
 2025 10.1007/BF00122760.
- 2026 Tseng, R.-S., and E. A. D'Asaro (2004), Measurements of turbulent vertical kinetic energy  
 2027 in the ocean mixed layer from Lagrangian floats, *Journal of Physical Oceanography*,  
 2028 34(9), 1984–1990, doi:10.1175/1520-0485(2004)034<1984:MOTVKE>2.0.CO;2.
- 2029 Tsujino, H., S. Urakawa, H. Nakano, R. Justin Small, W. M. Kim, S. G. Yeager, G. Dan-  
 2030 abasoglu, T. Suzuki, J. L. Bamber, M. Bentsen, C. Böning, A. Bozec, E. Chassignet,  
 2031 E. Curchitser, F. Boeira Dias, P. J. Durack, S. M. Griffies, Y. Harada, M. Illicak, S. A.  
 2032 Josey, C. Kobayashi, S. Kobayashi, Y. Komuro, W. G. Large, J. Le Sommer, S. J. Mars-  
 2033 land, S. Masina, M. Scheinert, H. Tomita, M. Valdivieso, and D. Yamazaki (2018),  
 2034 JRA-55 based surface dataset for driving ocean - sea-ice models (JRA55-do), *Ocean  
 2035 Modelling*, 130, 79–139, doi:10.1016/j.ocemod.2018.07.002.
- 2036 Uchiyama, Y., J. C. McWilliams, and A. F. Shchepetkin (2010), Wave-current interaction  
 2037 in an oceanic circulation model with a vortex-force formalism: Application to the surf  
 2038 zone, *Ocean Modelling*, 34(1), 16–35, doi:10.1016/j.ocemod.2010.04.002.
- 2039 Umlauf, L., and H. Burchard (2005), Second-order turbulence closure models for geophys-  
 2040 ical boundary layers. A review of recent work, *Continental Shelf Research*, 25, 795–827,  
 2041 doi:10.1016/j.csr.2004.08.004.

- 2042 Umlauf, L., H. Burchard, and K. Bolding (2014), GOTM sourcecode and test case docu-  
2043 mentation, *Tech. rep.*
- 2044 van den Bremer, T. S., and O. Breivik (2018), Stokes drift, *Phil. Trans. R. Soc. A*,  
2045 376(2111), 20170,104, doi:10.1098/rsta.2017.0104.
- 2046 Van Roekel, L., B. Fox-Kemper, P. P. Sullivan, P. E. Hamlington, and S. R. Haney (2012),  
2047 The form and orientation of Langmuir cells for misaligned winds and waves, *Journal of*  
2048 *Geophysical Research*, 117(C05001), C05,001, doi:10.1029/2011JC007516.
- 2049 Van Roekel, L., A. Adcroft, G. Danabasoglu, S. M. Griffies, B. Kauffman, W. Large,  
2050 M. Levy, B. G. Reichl, T. Ringler, and M. Schmidt (2018), The KPP boundary layer  
2051 scheme for the ocean: Revisiting its formulation and benchmarking one-dimensional  
2052 simulations relative to LES, *Journal of Advances in Modeling Earth Systems*, 10, 2647–  
2053 2685, doi:10.1029/2018MS001336.
- 2054 Walker, R., A. E. Tejada-Martínez, and C. E. Grosch (2016), Large-eddy simulation  
2055 of a coastal ocean under the combined effects of surface heat fluxes and full-depth  
2056 Langmuir circulation, *Journal of Physical Oceanography*, 46(8), 2411–2436, doi:  
2057 10.1175/JPO-D-15-0168.1.
- 2058 Wang, D., T. Kukulka, B. G. Reichl, T. Hara, I. Ginis, and P. P. Sullivan (2018), Interac-  
2059 tion of Langmuir turbulence and inertial currents in the ocean surface boundary layer  
2060 under tropical cyclones, *Journal of Physical Oceanography*, 48(9), 1921–1940, doi:  
2061 10.1175/JPO-D-17-0258.1.
- 2062 Warner, J. C., C. R. Sherwood, H. G. Arango, and R. P. Signell (2005), Performance  
2063 of four turbulence closure models implemented using a generic length scale method,  
2064 *Ocean Modelling*, 8(1), 81–113, doi:10.1016/j.ocemod.2003.12.003.
- 2065 Webb, A., and B. Fox-Kemper (2011), Wave spectral moments and Stokes drift estimation,  
2066 *Ocean Modelling*, 40(3-4), 273–288, doi:10.1016/j.ocemod.2011.08.007.
- 2067 Webb, A., and B. Fox-Kemper (2015), Impacts of wave spreading and multidirectional  
2068 waves on estimating Stokes drift, *Ocean Modelling*, 96, Part 1, 49–64, doi:10.1016/j.  
2069 ocemod.2014.12.007.
- 2070 Yang, D., B. Chen, M. Chamecki, and C. Meneveau (2015), Oil plumes and disper-  
2071 sion in Langmuir, upper-ocean turbulence: Large-eddy simulations and K-profile  
2072 parameterization, *Journal of Geophysical Research: Oceans*, 120, 4729–4759, doi:  
2073 10.1002/2014JC010542.
- 2074 Zhang, Z., G. P. Chini, K. Julien, and E. Knobloch (2015), Dynamic patterns in the re-  
2075 duced Craik-Leibovich equations, *Physics of Fluids (1994-present)*, 27(4), 046,605.