Comparing ocean surface boundary vertical mixing schemes including langmuir turbulence

JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2019MS001810

Comparing Ocean Surface Boundary Vertical Mixing Schemes Including Langmuir Turbulence

Qing Li[1,2], Brandon G. Reichl[3,4], Baylor Fox-Kemper[1], Alistair J. Adcroft[3,4], Stephen E. Belcher[5], Gokhan Danabasoglu[6], Alan L. M. Grant[7], Stephen M. Griffies[3,4], Robert Hallberg[3,4], Tetsu Hara[8], Ramsey R. Harcourt[9], Tobias Kukulka[10], William G. Large[6], James C. McWilliams[11], Brodie Pearson[1,12], Peter P. Sullivan[6], Luke Van Roekel[13], Peng Wang[11], and Zhihua Zheng[9]

[1]Department of Earth, Environmental and Planetary Sciences, Brown University, Providence, RI, USA, [2]Now at Fluid Dynamics and Solid Mechanics, Los Alamos National Laboratory, Los Alamos, NM, USA, [3]NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA, [4]Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, USA, [5]Met Office Hadley Centre, Exeter, UK, [6]National Center for Atmospheric Research, Boulder, CO, USA, [7]Department of Meteorology, University of Reading, Reading, UK, [8]Graduate School of Oceanography, University of Rhode Island, Narragansett, RI, USA, [9]Applied Physics Laboratory, University of Washington, Seattle, WA, USA, [10]School of Marine Science and Policy, University of Delaware, Newark, DE, USA, [11]Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, CA, USA, [12]College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR, USA, [13]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

Correspondence to:
Q. Li,
qingli@lanl.gov

Citation:
Li, Q., Reichl, B. G., Fox-Kemper, B., Adcroft, A. J., Belcher, S. E., Danabasoglu, G., et al. (2019). Comparing ocean surface boundary vertical mixing schemes including Langmuir turbulence. *Journal of Advances in Modeling Earth Systems, 11*, 3545–3592. https://doi.org/10.1029/2019MS001810

Received 11 JUL 2019
Accepted 21 OCT 2019
Accepted article online 30 OCT 2019
Published online 15 NOV 2019

Abstract Six recent Langmuir turbulence parameterization schemes and five traditional schemes are implemented in a common single-column modeling framework and consistently compared. These schemes are tested in scenarios versus matched large eddy simulations, across the globe with realistic forcing (JRA55-do, WAVEWATCH-III simulated waves) and initial conditions (Argo), and under realistic conditions as observed at ocean moorings. Traditional non-Langmuir schemes systematically underpredict large eddy simulation vertical mixing under weak convective forcing, while Langmuir schemes vary in accuracy. Under global, realistic forcing Langmuir schemes produce 6% (−1% to 14% for 90% confidence) or 5.2 m (−0.2 m to 17.4 m for 90% confidence) deeper monthly mean mixed layer depths than their non-Langmuir counterparts, with the greatest differences in extratropical regions, especially the Southern Ocean in austral summer. Discrepancies among Langmuir schemes are large (15% in mixed layer depth standard deviation over the mean): largest under wave-driven turbulence with stabilizing buoyancy forcing, next largest under strongly wave-driven conditions with weak buoyancy forcing, and agreeing during strong convective forcing. Non-Langmuir schemes disagree with each other to a lesser extent, with a similar ordering. Langmuir discrepancies obscure a cross-scheme estimate of the Langmuir effect magnitude under realistic forcing, highlighting limited understanding and numerical deficiencies. Maps of the regions and seasons where the greatest discrepancies occur are provided to guide further studies and observations.

1. Introduction

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

The OSBL is a turbulent region acting as a buffer between the rapid variations of the atmosphere and the slowly varying ocean interior. The competition between turbulent mixing in the OSBL and restratification mechanisms (buoyant and dynamical) determines the depth of this layer, which modulates the exchange of heat, momentum, and trace gasses (such as CO_2) between the atmosphere and ocean, thus affecting the workings of the broader Earth climate system. For example, seasonal variability of the OSBL affects mode water formation and thereby water properties in the ocean interior (e.g., Sallée et al., 2013a, 2013b;

Stommel, 1979) and biological activities in the surface ocean and thereby global biogeochemical cycles (e.g., Moore et al., 2013; Rodgers et al., 2014). On a shorter time scale, diurnal variability of the OSBL is important for the air-sea heat flux by affecting the diurnal cycling of sea surface temperature (SST; Large & Caron, 2015; Price et al., 1986). Due to their small horizontal spatial scale (10–100 m), short temporal scale (10^3–10^5 s) and nonhydrostatic nature, turbulent boundary layer mixing processes are not resolved in OGCMs. Therefore, such models require parameterizations (a.k.a. closures or subgrid schemes) that are both physically accurate and numerically efficient and stable.

Direct observational assessments of the quality of upper ocean mixing scaling relationships are rare (e.g., D'Asaro et al., 2014; Sutherland et al., 2014). The present paper takes a process-based, multiple scheme approach: contrasting parameterizations in an identical numerical and forcing framework, specifically constructed to avoid influences from biases of the driving model, and considering a range of realistic conditions. This approach is particularly good at identifying disagreement among parameterizations, but assessing parameterization skill is challenged by the lack of an appropriate "truth" for evaluation. Here large eddy simulations (LESs) 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 & Leibovich, 1976; Teixeira & Belcher, 2010). While breaking waves are responsible for the elevated near-surface turbulent dissipation in field measurements (Agrawal et al., 1992; Sullivan et al., 2007; Terray et al., 1996), which can be subsequently amplified due to the distortion of turbulence by Stokes drift (Grant & Belcher, 2009; Teixeira, 2012), Langmuir turbulence is largely responsible for a greater vertical velocity contribution to the turbulent kinetic energy (TKE) within the wavy OSBL than is found in a wall-bounded layer (D'Asaro, 2001; D'Asaro et al., 2014; Sutherland et al., 2014; Tseng & D'Asaro, 2004). 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 & Fox-Kemper, 2017), probably due to its coherent structures including deeply penetrating jets (Polton & Belcher, 2007).

Among many wave-related physical processes that affect the global climate, Langmuir turbulence is prominent (Belcher et al., 2012; Cavaleri et al., 2012; D'Asaro, 2014; D'Asaro et al., 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 (MLD) bias in the Southern Ocean (Sallée et al., 2013b). Nonbreaking wave turbulence (e.g., Babanin & Haus, 2009; Qiao et al., 2004) is another theory for wave-turbulence interactions akin to Langmuir turbulence in that stronger nonbreaking waves cause more mixing and which has stimulated parameterization studies (e.g., Chen et al., 2018; Fan & Griffies, 2014; Qiao et al., 2016). One model (FIO) among the CMIP5 ensemble did include nonbreaking wave turbulence. However, nonbreaking 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 nonlocal 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.

Journal of Advances in Modeling Earth Systems

10.1029/2019MS001810

Table 1
A Summary of Langmuir Turbulence Schemes Compared in this Study

Name	LT Parameters	K_λ	Down $\partial_z \boldsymbol{u}^S$ Momentum Flux	Entrainment	Non-LT Counterpart	References
KPPLT-VR12	$\mathrm{La}_{\mathrm{SLP}}$	$\mathcal{E}$ on q (A4)	Implicit	$\mathcal{E}$ on q (A4)	KPP-CVMix	McWilliams and Sullivan (2000); Van Roekel et al. (2012); Li et al. (2016)
KPPLT-LF17	$\mathrm{La}_{\mathrm{SLP}}$, $\mathrm{La}_{\mathrm{SL}}$	$\mathcal{E}$ on q (A4)	Implicit	LT-dependent $\overline{w'b'}_e$ in U_{tL}^2 (A5)	KPP-CVMix	Li and Fox-Kemper (2017)
KPPLT-R16	$\mathrm{La}_{\mathrm{SLP}}$, $\boldsymbol{u}^S(z)$	$\mathcal{E}_K$ on K_λ (A6)	$-K_\lambda^L \partial_z \boldsymbol{u}^L$	$\mathcal{E}_{U_t^2}$ on U_t^2 (A6)	KPP-CVMix + retuned Ri_c (A7)	Reichl et al. (2016)
ePBL-LT	$\mathrm{La}_{\mathrm{SL}}$	q via (A11)	Implicit	Extra energy due to LT: $m_{*\mathrm{LT}} u_*^3$ (A9)	ePBL	Reichl and Li (2019)
OSMOSIS	La_t, δ^S	K_λ via (A16)	-	Prognostic Eqn. of h_b	-	Appendix A3
SMCLT-H15	$u^S(z)$	SP in q & l Eqns., SP in ARSM for C_λ	$-K_u \partial_z \boldsymbol{u} - K_u^S \partial_z \boldsymbol{u}^S$ (A18)	–	SMC-KC94	Kantha and Clayson (2004); Harcourt (2013),Harcourt (2015)

Note. The second to fifth columns correspond to elements from sections 2.1, 2.2, 2.5, and 2.7, showing parameters for Langmuir turbulence (LT), modifications to the turbulent diffusivity K_λ in equation (2) via velocity scale q, length scale l or coefficient C_λ, down-Stokes drift shear momentum flux, and enhanced entrainment for each scheme. Down $\partial_z \boldsymbol{u}^S$ momentum flux is *implicit* for schemes that do not explicitly write out the Coriolis-Stokes force in the momentum equations, effectively assuming the simulated velocity being Lagrangian (see more discussion in Appendix A1). The sixth 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.

Table 2
A Summary of Schemes Without Langmuir Turbulence Compared in this Study

Name	K_λ	Γ_λ	References
KPP-CVMix	Empirical q & C_λ, $l = h_b$ determined by discretized Ri_b (A2)	Nonzero Γ_θ	Large et al. (1994); Danabasoglu et al. (2006); Griffies et al. (2015)
KPP-ROMS	Empirical q & C_λ, $l = h_b$ determined by integrated Ri_b (A3)	Nonzero Γ_θ	Large et al. (1994); McWilliams et al. (2009)
ePBL	Empirical q, l & C_λ constraining mechanical energy	–	Reichl and Hallberg (2018)
SMC-C01A	Dynamic q & l Eqns. (k-ϵ), C_λ from ARSM (weak-equilibrium)	–	Rodi (1987); Canuto et al. (2001); Warner et al. (2005)
SMC-KC94	Dynamic q & l Eqns. (k-kl), C_λ from ARSM (quasi-equilibrium)	–	Mellor and Yamada (1982); Kantha and Clayson (1994)

Note. The second and third columns show the different choices of the turbulent diffusivity K_λ and nonlocal flux Γ_λ in equation (2) in each scheme. ARSM = Algebraic Reynolds stress model.

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).

A typical approach for examining OSBL schemes compares different schemes in a common OGCM (e.g., Ali et al., 2019; Fan & Griffies, 2014; Large et al., 1997; Li et al., 2016; Li & Fox-Kemper, 2017), 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, for example, 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, 2016; Sallée et al., 2013b). 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 & 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 Reichl & Hallberg, 2018; Van Roekel et al., 2018). CVMix (Griffies et al., 2015) is also used as an external library, where many KPP-based Langmuir schemes have been implemented (e.g., Li & Fox-Kemper, 2017; Reichl et al., 2016). 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 disagreements highlight gaps in our understanding of the OSBL physics where either more LES studies or more ocean observations are needed. This information is therefore valuable for both ocean modelers and observationalists.

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

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

The remainder of this paper is organized as follows. In the next section, the elements of Langmuir turbulence parameterization schemes that conceptually distinguish them from non-Langmuir schemes are reviewed for reference in later sections. In section 3 the data and experimental setup of three test cases (ocean station moorings, global, and versus LES) for the comparison are detailed. The results of the three test cases are presented in section 4. Generally, the Langmuir schemes mix more strongly and deeply than their non-Langmuir partners, but they are also in less agreement than non-Langmuir schemes. Both numerical and physical constructions are involved in the discrepancies. A comparison of the regimes of forcing realized in the realistic and idealized test cases in section 4.1 highlights when the schemes most disagree. Discussions of the comparison results, highlighting remaining issues, key regimes for evaluation, and future research directions are presented in section 5. A brief summary and major conclusions of the comparison are presented in section 6. This paper includes four appendices, including a summary of the key components of each Langmuir scheme and the implementation methods in Appendix A, sensitivity tests for vertical resolutions and time steps in Appendix B, notes on the regimes of forcing (section 4.1) in Appendix C, and a description of source code and data accessibility in Appendix D.

2. Elements of Langmuir Turbulence Parameterization Schemes

The aim of an OSBL vertical mixing scheme is to accurately and efficiently approximate the vertical turbulent flux of an arbitrary field λ in terms of a parametric dependence on known (model resolved) properties and parameters. Here λ can be either a tracer or a component of the momentum. A general form of the parameterized vertical turbulent flux can be written as

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

where w is the vertical component of the velocity $\boldsymbol{u} = (u, v, w)$; $\overline{()}$ is an ensemble mean, often approximated by a horizontal and time mean (e.g., over a grid cell and time step as commonly implemented in OGCMs); and $()'$ is the fluctuations from this mean. The known variables in brackets can include mean fields $\bar{\lambda}$, the vertical shear of the mean horizontal velocity $\overline{\boldsymbol{S}} = (\partial_z \bar{u}, \partial_z \bar{v}, 0)$, stratification or the Brunt-Väsälä frequency N, surface boundary conditions (e.g., the surface buoyancy flux B_0), and nondimensional parameters (e.g., the turbulent Langmuir number La_t to be introduced in equation (4)). Some *second-moment closure* schemes evolve a system of covariances or second moments, including 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

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

The eddy diffusivity ($K_\lambda > 0$) is commonly written as the product of a velocity scale q, a length scale l, and a nondimensional coefficient C_λ (Prandtl, 1925); all can be functions of depth z,

$$K_\lambda = C_\lambda q l. \tag{3}$$

The last term in equation (2), Γ_λ, represents any flux not proportional to the local gradient of $\bar{\lambda}$ and is thus known as the *nonlocal 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 (Kenyon, 1969; Myrhaug et al., 2014; Stokes, 1847; van den Bremer & Breivik, 2018; Webb & Fox-Kemper, 2011, 2015) is the key forcing component of the theory of Langmuir circulation and Langmuir turbulence (Craik & Leibovich, 1976; McWilliams et al., 1997). The full Stokes drift profile is required for some schemes (Harcourt, 2013, 2015; Kantha & Clayson, 2004; Noh et al., 2016; Reichl et al., 2016; Sinha et al., 2015). 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 & 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., Breivik et al., 2016; Li et al., 2017; Webb & Fox-Kemper, 2015; Sullivan et al., 2012).

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

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

(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 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 δ^S assuming an exponential profile. This definition in (4) is attractive for its simplicity and its utility in scaling the energy budget (Grant & Belcher, 2009; McWilliams et al., 1997). Although many of the schemes here use variants other than La_t, this parameter is used for all regime diagnoses in this paper (sections 4.1 and 4.3.2).

Notably, La_t neglects whether the Stokes drift is a result of long waves or short waves, that is, the influence of the Stokes drift decay scale δ^S. The degree to which the Stokes drift extends into the OSBL is important (Kukulka & 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 $\mathrm{La}_{\mathrm{SL}}$ from the surface layer (a fraction of the mixed layer, upper 20% in their definition) averaged Stokes drift $\langle u^S\rangle_{\mathrm{SL}}$ and a reference Stokes drift u^S_{ref} near the base of the mixed layer,

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

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

The projected Langmuir number takes account for the effects of misalignment between wind and waves (Hanley et al., 2010; Reichl et al., 2016; Van Roekel et al., 2012). One variant based on $\mathrm{La}_{\mathrm{SL}}$ is written as

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

where θ_{ww} is the misalignment between wind and waves and θ_{wl} between wind and Langmuir cells. In GCMs, coupling with a wave model is required to estimate these angles (e.g., Li et al., 2016). Different variants of $\mathrm{La}_{\mathrm{SLP}}$ are used in KPPLT-VR12, KPPLT-R16, and the scaling of Langmuir-enhanced diffusivity in KPPLT-LF17.

2.2. Enhanced Turbulent Diffusivity K_λ

There is a general consensus that the diffusivity K_λ appearing in equation (3) should be elevated in the presence of Langmuir turbulence to represent the enhanced vertical mixing found in field measurements. The question is how? We can do so by enhancing K_λ directly (e.g., KPPLT-R16) or enhancing the velocity scale q (e.g., McWilliams & Sullivan, 2000; KPPLT-VR12 and KPPLT-LF17) in first-moment closure schemes using an Langmuir number-dependent enhancement factor. Figure A1 compares different formulas of the enhancement factor applied to q adopted in some of these schemes. Note that directly enhancing q affects

other components of the scheme that depend on q, which is one of the differences between KPPLT-VR12 and KPPLT-R16. In second-moment closure schemes, this is usually done by adding the contribution of Stokes drift to the TKE and length scale equations and C_λ in equation (3) (e.g., Axell, 2002; Kantha & Clayson, 2004; SMCLT-H15) or enhancing the length scale l in some other variants (e.g., Noh et al., 2016).

2.3. Modified Vertical Profile of K_λ

In addition to its magnitude, the functional form of K_λ versus z is also modified by the presence of Langmuir turbulence (e.g., Sinha et al., 2015; Yang et al., 2015). This modification is a direct result of the distinctive structures of Langmuir turbulence (Teixeira & Belcher, 2010), with enhanced vertical TKE versus its horizontal components. This effect can be accounted for by modifying the coefficient C_λ in equation (3), corresponding to the shape function G_λ in KPP-based schemes (e.g., Sinha et al., 2015; Yang et al., 2015) and the stability function S_λ in algebraic second-moment closure models such as SMCLT-H15. However, a consensus has not been reached on such modifications. The Stokes drift decay depth seems to affect this profile shape significantly (Kukulka & Harcourt, 2017).

2.4. Modified Nonlocal Flux Γ_λ

The nonlocal flux Γ_λ in equation (2) is usually attributed to coherent structures in the OSBL such as convective plumes (Large et al., 1994; Troen & Mahrt, 1986). Given the coherent windrows associated with Langmuir circulation commonly seen in ocean surface observations (Thorpe, 2004) and the deeply penetrating jets seen especially in LES of Langmuir turbulence (Polton & Belcher, 2007), it is expected that such structures of Langmuir turbulence will contribute to Γ_λ. The nonlocal fluxes of scalar fields and momentum are discussed extensively in the context of KPP (Griffies et al., 2015; McWilliams & Sullivan, 2000; Sinha et al., 2015; Smyth et al., 2002; Van Roekel et al., 2018), but an agreement on the effects of Langmuir turbulence has not been reached, perhaps because nonlocal fluxes are typically diagnosed as the residual of local fluxes, causing errors in local fluxes to appear in nonlocal flux diagnoses. Therefore, none of the KPP schemes compared here alter this term.

2.5. Down-Stokes Drift Shear Momentum Flux

A down-Stokes drift shear momentum flux is "nonlocal" to the Eulerian mean flow and can be treated as a nonlocal momentum flux in KPP-based schemes (e.g., McWilliams & Sullivan, 2000; Sinha et al., 2015). The necessity of this term is also justified by the "anti-Stokes" effect (Haney et al., 2015; McWilliams & Fox-Kemper, 2013; Pearson, 2018; Suzuki & Fox-Kemper, 2016; Uchiyama et al., 2010), which significantly modifies the Eulerian mean velocity profile. The idea of down-Lagrangian shear mixing is extensively examined in McWilliams et al. (2012) and adopted in Langmuir turbulence parameterization schemes such as the KPPLT-R16, where the same turbulent viscosity is used for both the down-Eulerian shear and the down-Stokes drift shear components. Other schemes, such as SMCLT-H15, require a generally different turbulent viscosity for the down-Stokes drift shear component, due to the distinct dynamical effects of the Stokes drift and the Eulerian shear.

2.6. Momentum Flux Distinct from Shear Orientation

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

2.7. Enhanced Entrainment Due to Unresolved Shear

Entrainment at the base of the OSBL is important for the vertical fluxes of heat and other tracers that have the source at the surface or below the OSBL (McWilliams et al., 2014). LES show enhanced entrainment in the presence of Langmuir turbulence (e.g., Grant & Belcher, 2009; Li & Fox-Kemper, 2017; McWilliams et al., 1997, 2014), partially due to plumes reminiscent of convective plumes (Li, 2018), associated with the penetrating jets of Langmuir turbulence (Polton & Belcher, 2007). To account for this effect, the unresolved shear term in KPP can be enhanced either by applying an enhancement factor (e.g., KPPLT-VR12 and KPPLT-R16)

Table 3
A Summary of Simulations at Ocean Stations

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

Note. The domain refers to the vertical extent of the water column considered.

or by explicitly allowing the entrainment buoyancy flux $\overline{w'b'}_e$ to depend on Langmuir number in the derivation of this term (e.g., McWilliams et al., 2014; KPPLT-LF17). The ePBL-LT scheme accounts for this effect by including the Langmuir turbulence contribution in the mechanical energy budget.

2.8. Enhanced Entrainment Due to Resolved Shear

Along with other mechanisms including wave breaking (Kudryavtsev et al., 2008) and a partition of the momentum flux between wave-induced and shear-induced components (Teixeira, 2018), enhanced vertical mixing of momentum by Langmuir turbulence also contributes to the weaker velocity shear near the surface than in a wall-bounded layer, as seen in both observations (e.g., Gargett et al., 2004; Schudlich & Price, 1998) and LES (e.g., Kukulka et al., 2010; McWilliams et al., 1997). As a result, the mean shear flow at the base of the OSBL is elevated and contributes significantly to the mixed layer deepening (Kukulka et al., 2010). In addition, the interaction between Langmuir turbulence and inertial currents also needs to be considered to correctly simulate the shear-induced mixing at the base of the OSBL, especially after a tropical cyclone passage (Wang et al., 2018). With sufficient vertical and temporal resolutions to resolve the shear near the base of the OSBL and inertial currents, this effect is included in most of the Langmuir schemes compared here, though it depends heavily on the parameterized momentum flux.

3. Data and Comparison Methods

The General Ocean Turbulence Model Version 5 (GOTM5; Umlauf et al., 2014; see updated version on gotm.net) is a one-dimensional water column model with flexible configurations of vertical mixing schemes. We make use of GOTM5 as a common driver for all vertical mixing schemes. All Langmuir schemes to be compared were implemented directly in GOTM5 by adapting their current implementations in popular OGCMs or single-column models, or in the external CVMix package (Griffies et al., 2015; see updated version on github.com/CVMix) and called from GOTM5. Implementations of these Langmuir schemes in GOTM5 were verified separately and in various degrees of detail by comparing with their original implementations under scenarios that were used for development. To provide the necessary wave information for all Langmuir schemes and, where applicable, Coriolis-Stokes force additions (see Table 1 and Appendix A1), we implemented in GOTM5 an interface that allows the input of Stokes drift via various methods. A list of all these schemes can be found in Tables 1 and 2. Brief descriptions of key components of each scheme and the implementation methods in GOTM5 can be found in Appendix A.

To reveal the differences among schemes in different scenarios, single-column simulations with GOTM5 were conducted in three test cases that serve distinct purposes. The initial conditions and surface forcing of these simulations range from idealized scenarios, as commonly used in LES studies, to more realistic scenarios, either as observed at Ocean Stations and by Argo floats or as covered in the global atmospheric surface forcing data set JRA55-do (Tsujino et al., 2018) designed to drive global ocean/sea-ice models. JRA55-do is derived from a blend of observations and reanalysis and represents an estimate of the global atmospheric state that we use to derive ocean surface forcing conditions.

All code and data are available online. See Appendix D for more details.

3.1. Test Case 1: Simulations at Ocean Stations

Observed temperature and salinity profiles and surface fluxes from two ocean stations: (1) the Ocean Surface Mixing, Ocean Submesoscale Interaction Study in the northeast Atlantic (OSMOSIS, 48.7° N, 16.2° W; Damerell et al., 2016) and (2) Ocean Climate Station Papa in the northern Pacific (OCS-Papa, 50.1° N, 144.9° W; www.pmel.noaa.gov/OCS/Papa) and waverider buoy data (J. Thomson, APL-UW) were used to initialize and force GOTM5 simulations in this test case. Three sets of simulations were conducted, as listed in

Table 3, focusing on the deepening of mixed layer in winter, shoaling of mixed layer in spring, and a full annual cycle. Each set consists of 11 runs with different schemes.

All simulations are run with $\Delta z = 1$ m vertical grid spacing and $\Delta t = 1$ min time step, and with 3-hourly instantaneous output for analysis. It is previously demonstrated that similar simulations are sensitive to numerical choices such as model time step and vertical resolution (e.g., Reichl & Hallberg, 2018; Van Roekel et al., 2018). We therefore choose this numerical configuration similar to their fine resolution cases, where the dependence of the simulation on model time step and vertical resolution is expected to converge. In addition, combinations of three different vertical grids, $\Delta z = [1, 5, 10]$ m, and four different time steps, $\Delta t = [1, 10, 30, 60]$ min, were tested to assess the sensitivity of each scheme to the vertical resolution and time steps that are more typical of contemporary climate modeling practice (see Appendix B for more details). For simplicity, a linear equation of state with thermal expansion coefficient, $\alpha_T = 1.66 \times 10^{-4}$ ° C^{-1}, and saline contraction coefficient, $\beta_S = 7.6 \times 10^{-4}$ kg/g, 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

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

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

with g the gravitational acceleration, and $T_{\text{ref}} = 10$ ° C and $S_{\text{ref}} = 35$ g/kg. 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, for example, 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., Skyllingstad 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, for example, 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 are 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° × 4° box over the global ocean south of 72° 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 24,000 sets of such 1-month single-column simulations (~2,000 [grid points] × 12 [months]). Each set of simulations contains 11 runs with different schemes, all initialized from the same Argo temperature and salinity profiles with zero velocity. Argo profile data from years 2004 to 2013 (WOD13; Argo, 2000; Boyer et al., 2013) were used to generate the initial 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 ±20 days relative 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 selected as the initial conditions for that 4° × 4° 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 reproducing 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.

The JRA55 (Japan Meteorological Agency/Japan, 2013) based surface data set 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, humidity, sea level pressure, precipitation, and short wave radiation. The values of these variables 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 SST and currents. Ocean surface wave information was obtained from a WAVEWATCH III simulation forced by the same JRA55-do surface data set for consistency. Instead of saving the full wave spectrum from WAVEWATCH III, the directional partitioned surface Stokes drift $\boldsymbol{u}_{0i}^{S}$ with three frequency bins ($n = 3$) were saved and input into GOTM5 to approximately reconstruct the full profile of Stokes drift from

$$\boldsymbol{u}^{S}(z) = \sum_{i=1}^{n} \boldsymbol{u}_{0i}^{S} \exp\left(2k_i z\right), \tag{9}$$

where $\boldsymbol{u}_{0i}^{S}$ is determined by integrating the full wave spectrum over each frequency bin with a mean wave number 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 were 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 & Yeager, 2009). The comparison of these forcing products in terms of their effects on mixing will be described in a separate paper.

The length of each simulation was set to 1 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.

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 without significant degradation in the quality of the results (see Appendix B). Three-hourly instantaneous output was 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.

3.3. Test Case 3: Idealized Simulations and Comparison with LES

LES with various constant wind, waves, and destabilizing surface buoyancy flux forcing (Li & 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 conditions and surface forcing in those studies are idealized, though the parameter space covered 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, and so forth. 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). 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 hr 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 m/s translation speed and maximum wind speed of 65 m/s (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 hr (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 A1).

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 A1 (see more discussion in Appendix A1). 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 La_t - h/L_L parameter space introduced by Belcher et al. (2012). The turbulent Langmuir number 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 h a length scale of the OSBL thickness and

$$L_L = -\frac{u_*^2 u_0^S}{B_0} \tag{10}$$

the Langmuir stability length. The surface buoyancy flux B_0 (dimensions of squared length per cubed time) is defined such that a positive value corresponds to stabilizing (e.g., ocean freshening or warming) conditions. Note that this is opposite to the definition used in Belcher et al. (2012) so that a positive value of L_L corresponds to convection here. However, since we also define h to be positive, the parameter h/L_L is consistent with the definition in Belcher et al. (2012).

Belcher et al. (2012) categorize the surface forcing regime under destabilizing surface buoyancy flux conditions ($B_0 < 0$) into shear turbulence, Langmuir turbulence, and convective turbulence regimes (see their Figure 3). They do so according to the relative 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,

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

with dimensionless coefficients ($A_S = 2[1 - \exp(-0.5\mathrm{La}_t)]$, $A_L = 0.22$ and $A_C = 0.3$) estimated from LES studies. Equation (11) is derived by scaling the three TKE production mechanisms and assuming the dissipation is proportional to the production for each mechanism (see the derivation and discussions in Belcher et al., 2012). This scaling serves as an example of many possible ways to divide the parameter space into different regimes in an asymptotic sense. Controversies exist on both the exact values of the dimensionless coefficients and the functional form of this scaling, and its applicability to accurately describe the sea surface conditions in field measurements (e.g., Large et al., 2019). For the purpose here, however, this regime diagram is helpful to contrast the parameter space already explored in LES of Langmuir turbulence with that in realistic global ocean data sets. More discussion on the scaling of TKE dissipation and a theoretical

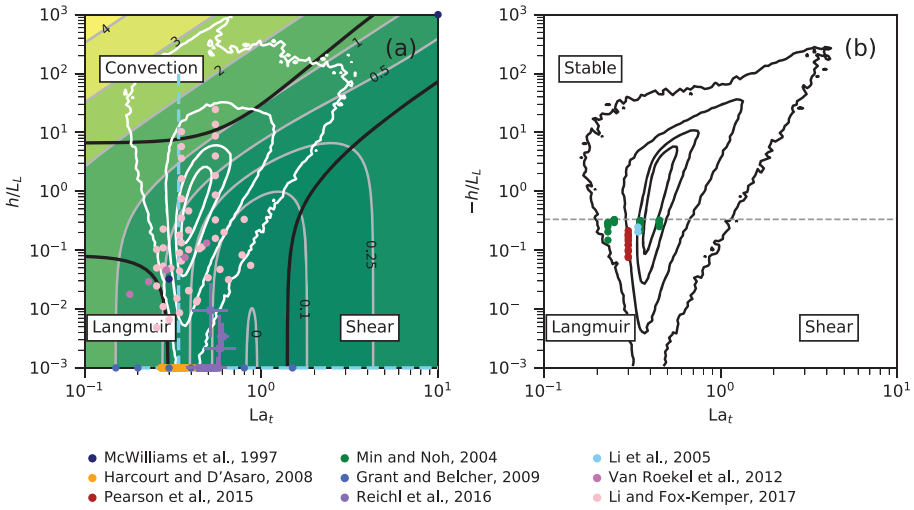

Figure 1. Regime diagrams in parameter spaces (a) La_t - h/L_L following Belcher et al. (2012) and (b) La_t - $-h/L_L$, which describe situations with destabilizing and stabilizing surface buoyancy fluxes, respectively. Both panels are plotted in log-log space. Filled contours in the background of panel a show the normalized turbulent dissipation rate $\log_{10}\left(\epsilon h/u_*^3\right)$ 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 $\log_{10}\left(La_t\right)$ and $\log_{10}\left(h/L_L\right)$ ($\log_{10}\left(-h/L_L\right)$ 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 ± 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).

model based on rapid distortion theory with some other assumptions, particularly focusing on the effect of Langmuir turbulence, can be found in Teixeira (2012).

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

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

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. This distribution indicates that when $B_0 < 0$ the typical scale of h/L_L over the global ocean places surface forcing conditions from the 3-hourly JRA55-do data between Langmuir turbulence and convection. The other 38% of forcing has stabilizing surface buoyancy fluxes ($B_0 > 0$), shown in the La_t - $-h/L_L$ parameter space in Figure 1b. Under both destabilizing and stabilizing surface conditions, wind-driven shear turbulence seldom dominates the dissipation

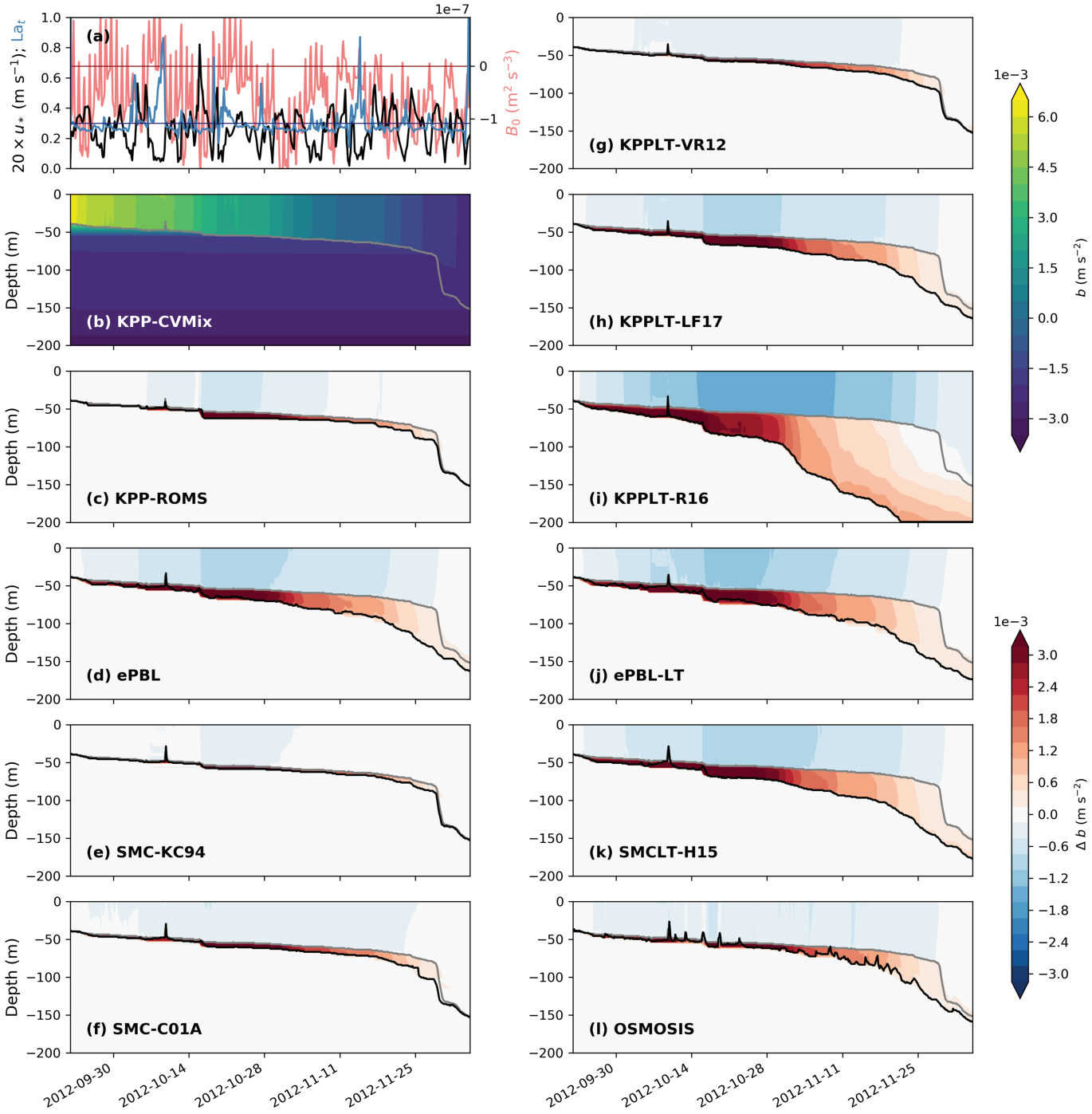

Figure 2. Comparison of simulated buoyancy with different schemes for the OSMOSIS-Winter case. Panel a shows time series of water-side surface friction velocity u_* (m/s) in black, turbulent Langmuir number La_t in blue ($La_t = 0.3$ in dark blue for reference), and surface buoyancy flux B_0 (m^2/s^3) in red ($B_0 = 0$ m^2/s^3 in dark red for reference). For clarity, $20 \times u_*$ is shown, sharing the same vertical axis with 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 (m/s^2) for KPP-CVMix, with the mixed layer depth (MLD) defined by the 0.03 kg/m^3 density criterion (de Boyer Montégut et al., 2004) marked in gray. For reference, a buoyancy anomaly of 3.0×10^{-3} m/s^2 is the equivalent of 1.8 ° C temperature or 0.40 g/kg 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.

by this measure, though its role on the vertical mixing in the OSBL is likely underrepresented in this regime diagram partially due to its strong coupling with Langmuir turbulence.

Scattered dot and square (with error bars) symbols and dashed lines in color mark the regimes explored in some representative LES studies of Langmuir turbulence in the literature, some of which were used in training the schemes here. In particular, orange dots mark the regimes of realistic wind and waves in pure wind sea explored in Harcourt and D'Asaro (2008), of which the LES data were used to tune parameters in SMCLT-H15. Blue dots mark the regimes explored in Grant and Belcher (2009) with various La_t, u_* and Stokes drift decay depth assuming an exponential profile, and red the regimes explored in Pearson et al. (2015) with various surface heating conditions and rotation effects. Both sets of LES were used to tune parameters in OSMOSIS. Neutral surface forcing conditions in both Harcourt and D'Asaro (2008) and Grant and Belcher (2009) correspond to $h/L_L = 0$, represented in this diagram by symbols on the bottom edge for convenience. Magenta dots mark the regimes of misaligned wind and waves explored in Van Roekel et al. (2012), of which the scaling for vertical velocity variance was used in KPPLT-VR12 to estimate the enhancement factor. Pink dots mark the regimes explored in LES of mixed Langmuir turbulence and convection by Li and Fox-Kemper (2017), which was used in both KPPLT-LF17 and ePBL-LT to constrain the entrainment buoyancy flux and integrated buoyancy flux over the entrainment zone, respectively. Purple squares with error bars (mean $\pm$ standard deviation) mark the regimes spanned in the transient conditions of idealized hurricanes explored in Reichl et al. (2016), which were used to tune parameters in KPPLT-R16.

For LES under constant forcing in stable regimes (colored dots in Figure 1b), h/L_L was estimated from h_i/L_L with h_i the initial MLD using equation 4 of Pearson et al. (2015). This equation suggests a maximum equilibrium value of $-h/L_L = 1/3$ in the limit of strong surface heating, which is shown by the gray dashed line. However, under transient forcing, $-h/L_L$ can be significantly larger as illustrated by the black contours.

It should be noted that the symbols and dashed lines here only represent the projections of the actual regimes explored in each study onto the La_t - h/L_L parameter space. There are other dimensions of the full parameter space explored in these studies that are not represented in this regime diagram, such as different penetration depths of Stokes drift in Harcourt and D'Asaro (2008), different wind-wave misalignment conditions in Van Roekel et al. (2012), and different combinations of surface and penetrative (radiative) heat fluxes in Pearson et al. (2015). Even so, this regime diagram reveals that many parts of the parameter space are still not well explored, especially those of Langmuir turbulence under transient stabilizing surface conditions (Figure 1b). Likewise, were the probability distributions for the station observations in Test Case 1 (section 3.1; Table 3) shown, it would be evident that the LESs differ greatly from the forcing regimes seen at the OCS-Papa and OSMOSIS sites. For this reason, it is insufficient to evaluate schemes here versus only LES.

4.2. Test Case 1: An Overview of the Comparison

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

OSMOSIS-Winter and OSMOSIS-Spring together represent a full annual cycle of the mixed layer at a midlatitude region in the North Atlantic, with strong variations in the MLD. The winter mixed layer deepening is rapid due to a strong destabilizing surface buoyancy flux, leading to a MLD of a few hundred meters starting from the rather shallow summer MLD of less than 50 m. Intermittent deepening events occur during the spring shoaling of the mixed layer. The water properties in the mixed layer are significantly influenced by horizontal advection of different water masses and submesoscale processes (Damerell et al., 2016; Thompson et al., 2016), which are ignored here in the single-column simulations. The turbulent Langmuir number is around 0.3 throughout the year (Figures 2a and 3a), suggesting significant influences of Langmuir turbulence on the vertical mixing.

OCS-Papa represents a full annual cycle of the mixed layer at midlatitude in the North Pacific. As a result of the much milder winter destabilizing surface buoyancy flux, the seasonal variations of MLD at this site is much weaker than that at the OSMOSIS site, with a summer MLD of about 20 m and a winter MLD of about 100 m. However, the wind is strong in winter, and surface gravity waves are intense with a turbulent

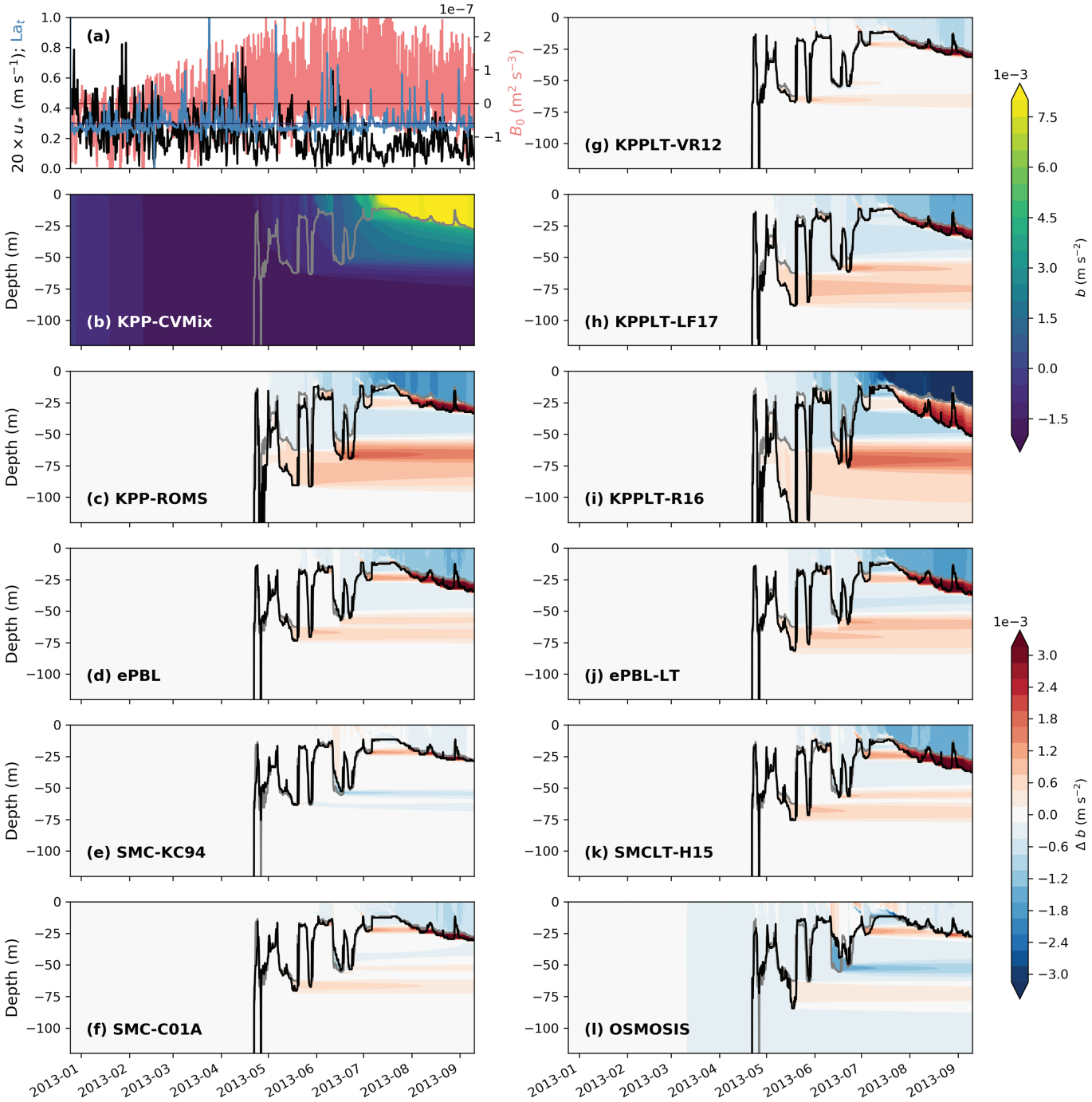

Figure 3. Same as Figure 2 but for the OSMOSIS-Spring case. Panel a shows water-side surface friction velocity u_* (m/s) in black, turbulent Langmuir number La$_t$ in blue, and surface buoyancy flux B_0 (m^2/s^3) in red. Panel b shows the simulated buoyancy b (m/s^2) for KPP-CVMix, with MLD defined by the 0.03 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 kg/m^3 MLD in black and the KPP-CVMix MLD in gray. For reference, a buoyancy anomaly of 3.0×10^{-3} m/s^2 is the equivalent of 1.8 ° C temperature or 0.40 g/kg salinity anomaly. Only the upper 120 m of the domain is shown for clarity.

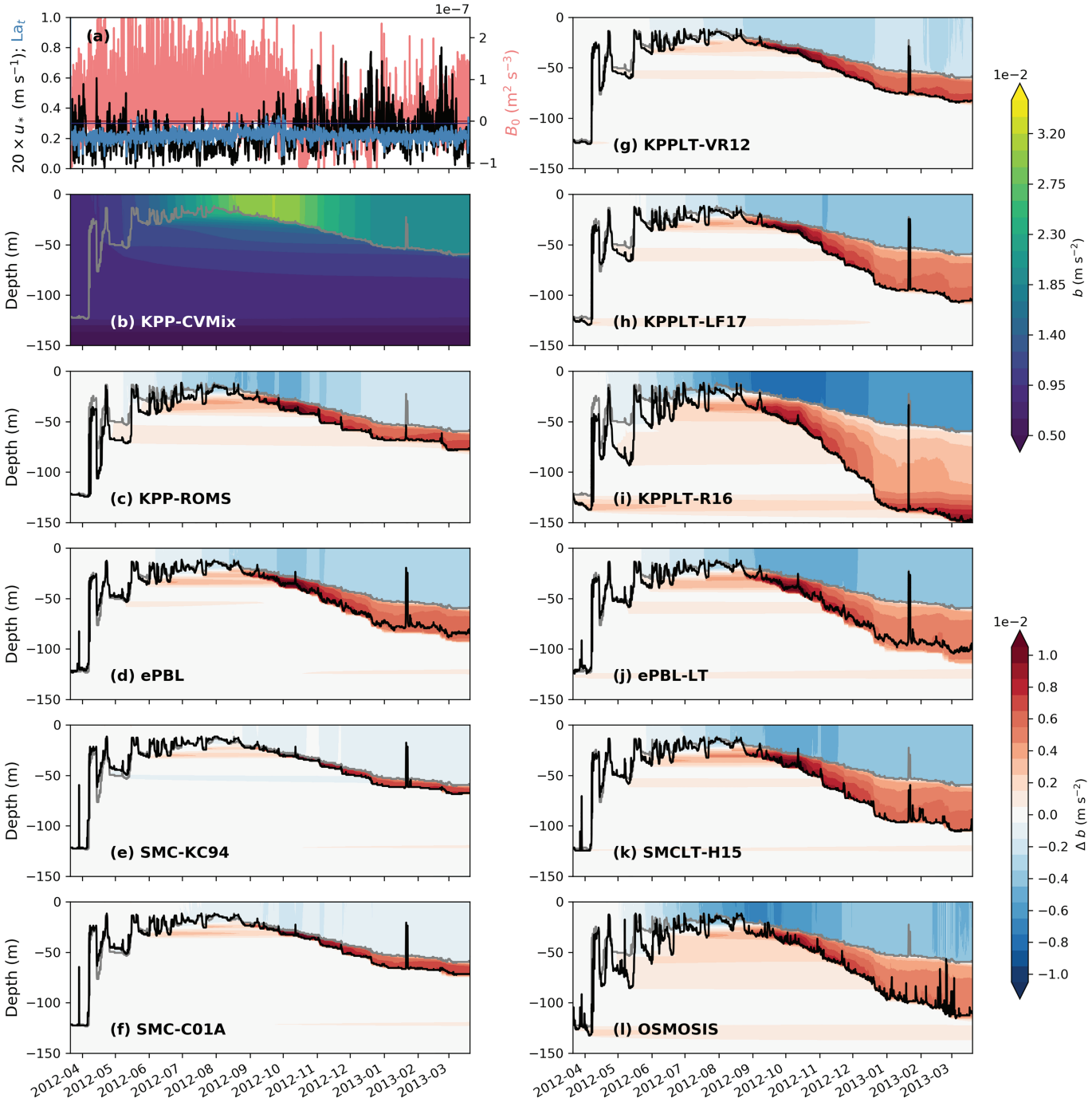

Figure 4. Same as Figure 2 but for the OCS-Papa case. Panel a shows water-side surface friction velocity u_* (m/s) in black, turbulent Langmuir number La$_t$ in blue, and surface buoyancy flux B_0 (m^2/s^3) in red. Panel b shows the simulated buoyancy b (m/s^2) for KPP-CVMix, with MLD defined by the 0.03 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 kg/m^3 MLD in black and the KPP-CVMix MLD in gray. For reference, a buoyancy anomaly of 3.0×10^{-3} m/s^2 is the equivalent of 1.8 ° C temperature or 0.40 g/kg salinity anomaly. Only the upper 150 m of the domain is shown for clarity.

Langmuir number of around 0.2 (Figure 4a). Therefore, considering the much shallower MLD than at the OSMOSIS site, we expect much greater influences of Langmuir turbulence on the vertical mixing at the OCS-Papa site.

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

A closer examination of these figures shows that most of the discrepancies appear to be in the mixed layer deepening phase, for example, evident in Figures 2 and 4 (September to January). Also note the discrepancies among schemes in sporadic mixed layer deepening events shown by the buoyancy anomaly extending in time in Figures 3 and 4. In particular, of the five non-Langmuir schemes, ePBL gives the most rapid deepening of mixed layer under destabilizing surface conditions, whereas KPP-CVMix gives the weakest, with the result of SMC-C01A closer to ePBL and SMC-KC94 closer to KPP-CVMix. Note, however, that the differences between the results of SMC-C01A and SMC-KC94 are relatively small in all three cases. KPP-ROMS seems to agree with KPP-CVMix in OSMOSIS-Winter case where MLD is relatively deep and destabilizing surface buoyancy flux is strong, but less so in the other two cases where MLD is relatively shallow, destabilizing surface buoyancy flux is relatively weak and wind is strong. Sensitivity tests show that the disagreements between the two versions of KPP decreases with increased damping of the horizontal velocity, mostly resulting from the changes in KPP-ROMS (not shown). This result suggests that KPP-ROMS is more susceptible to strong velocity shear at the base of the OSBL associated with inertial oscillations, which may also contribute to its larger sensitivity to vertical resolution than KPP-CVMix (Appendix B). Consistent with Van Roekel et al. (2018), we also notice quite strong sensitivity of KPP-CVMix to numerical details in determining the boundary layer depth from a bulk Richardson number criterion (see Appendix A1 for more details). All these results suggest that the algorithm and other numerical details in the formulation and implementation of a particular scheme are also important in addition to the underlying physics.

Of the six Langmuir schemes, KPPLT-R16 gives the most rapid deepening of mixed layer under destabilizing surface conditions. This behavior is expected as KPPLT-R16 was targeted to study strong wind and waves regimes under hurricane conditions without an explicit constraint on effects of destabilizing surface buoyancy flux (Reichl et al., 2016) and will likely overpredict the entrainment by convection (see also Li & Fox-Kemper, 2017). KPPLT-VR12 gives the weakest mixed layer deepening, especially when the destabilizing surface buoyancy flux is strong. The results of KPPLT-LF17, ePBL-LT, and SMCLT-H15 are similar in most respects. However, the net effects of Langmuir turbulence, as reflected by comparing each Langmuir scheme to their non-Langmuir counterpart, appear to agree better in the OCS-Papa and OSMOSIS-Spring cases but less so in the OSMOSIS-Winter case. The OSMOSIS scheme behaves similarly to other Langmuir schemes in mixed layer deepening events but differs significantly in mixed layer shoaling and restratification events. This is probably because it solves prognostic equations for boundary layer depth under both deepening and shoaling conditions (see more details in Appendix A3).

4.3. Test Case 2: A Global Perspective

In this section we present results of Test Case 2, which compares different schemes from a global perspective. Unlike the previous section, here we use the monthly mean MLD defined by the 0.03 kg/m^3 density criterion (de Boyer Montégut et al., 2004) as a diagnostic quantity to quantify the different behaviors of all the 11 schemes. Time series of MLD for each 1-month simulation is first diagnosed from the density field then averaged to get the monthly mean. Note that due to the monthly average and the identical initial conditions for all schemes at the beginning of every month, this measure tends to underestimate the differences among different schemes. Yet significant discrepancies among schemes are found by this measure. This measure is also directly relevant to the application in climate models. Other diagnostics, such as alternative definitions of MLD, the rate of change in potential energy and SST, were also examined, and they generally tell the same story, though the results are quantitatively different, reflecting the complex nature of the turbulent mixing problem that is not fully captured by the MLD. While we acknowledge the necessity of using multiple diagnostics to fully describe the different behaviors of these schemes, for brevity full documentation of all diagnostics is left out, though the reader is encouraged to access the simulation code and full suite of diagnostic results online (github.com/qingli411/gotm and github.com/qingli411/gotmwork).

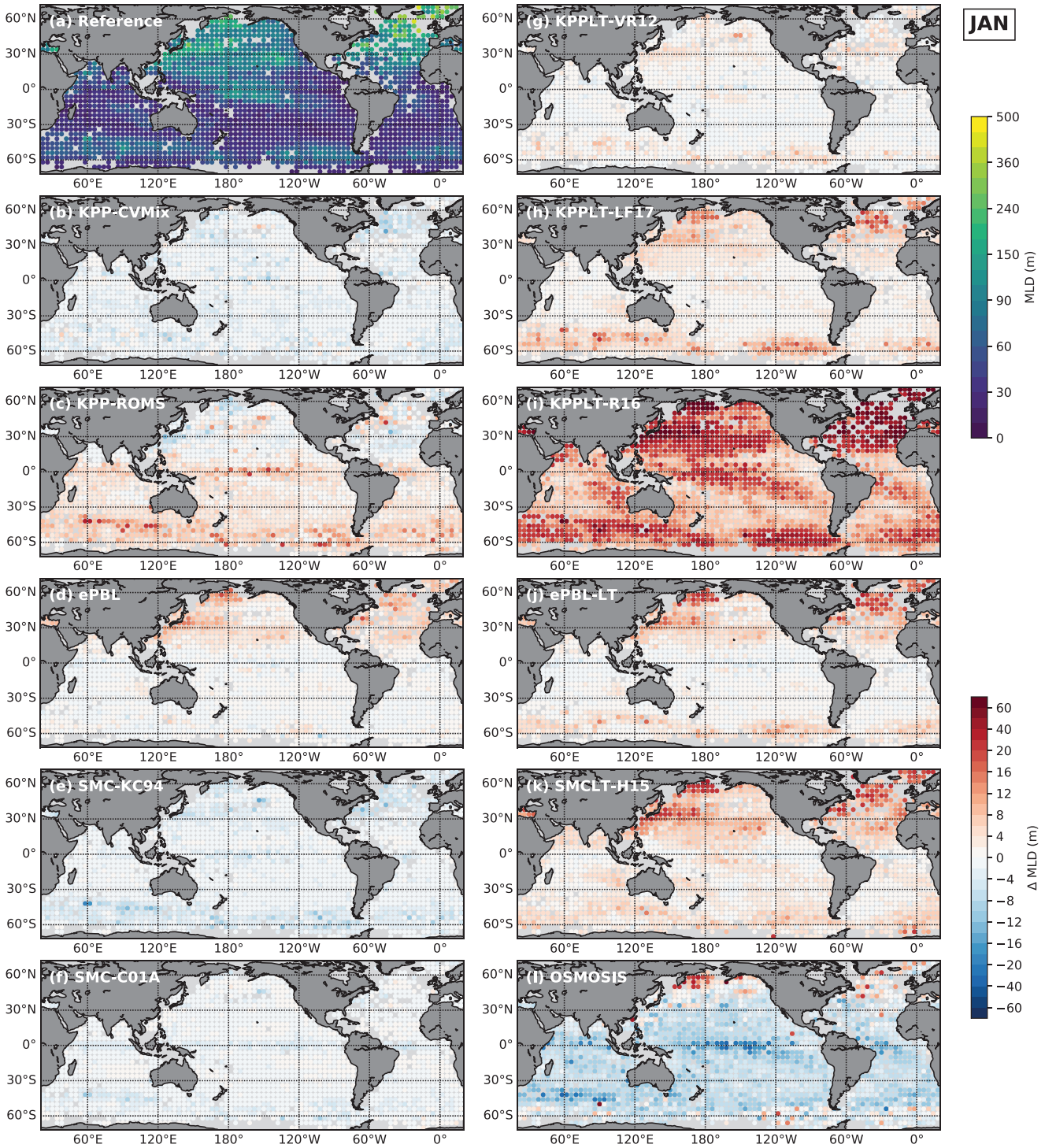

Figure 5. Comparison of January 2009 mean mixed layer depth (MLD) simulated by different schemes with atmospheric forcing from JRA55-do. Panel a shows the map of the reference MLD taken as the mean of the five non-Langmuir schemes (upper color bar in m). Panels b–l show the differences of MLD from that 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.

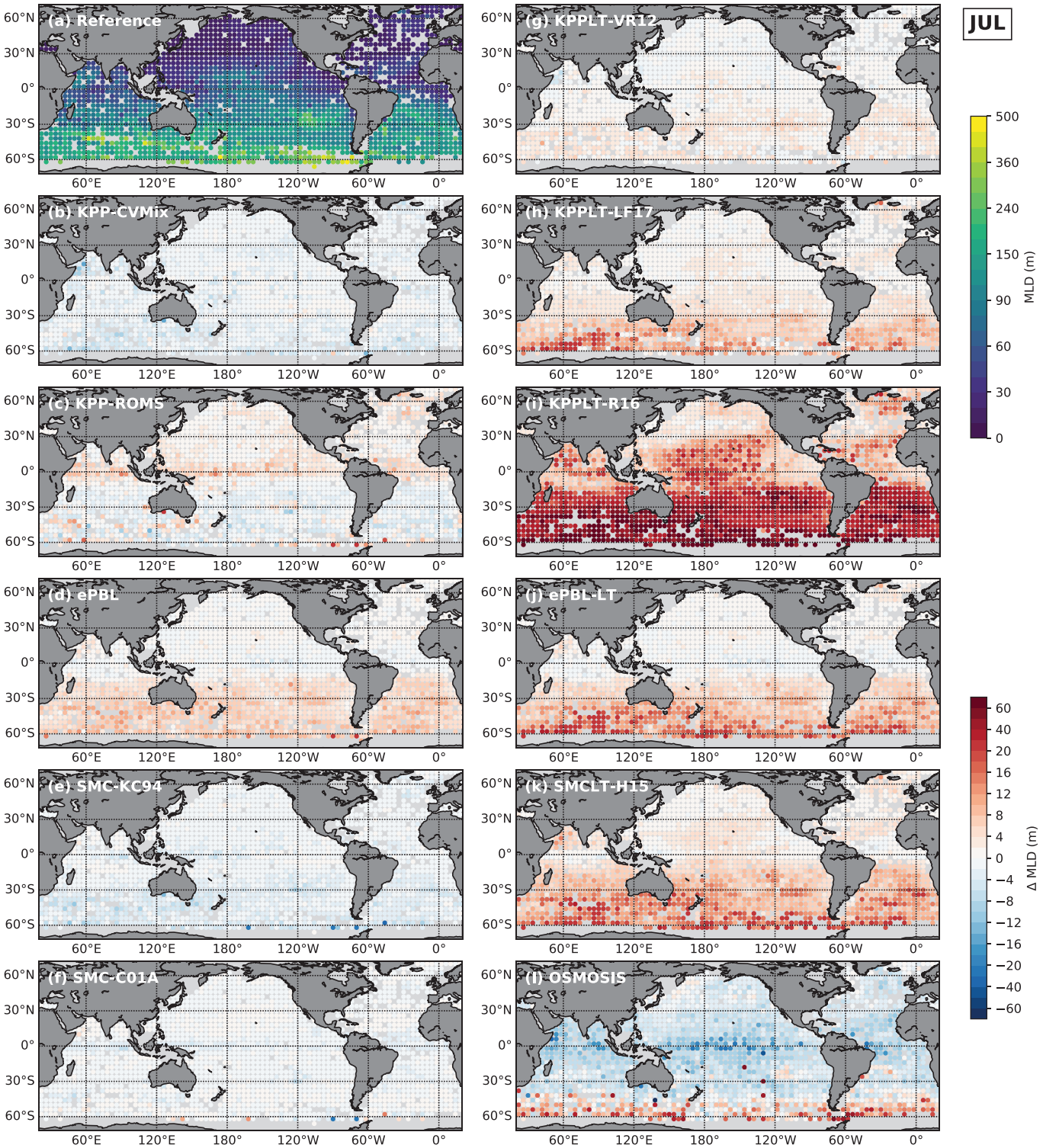

Figure 6. Same as Figure 5, but comparison of July 2008 mean mixed layer depth (MLD). (a) Mean of five non-Langmuir schemes, b–l show the differences from mean in a for each of the 11 schemes.

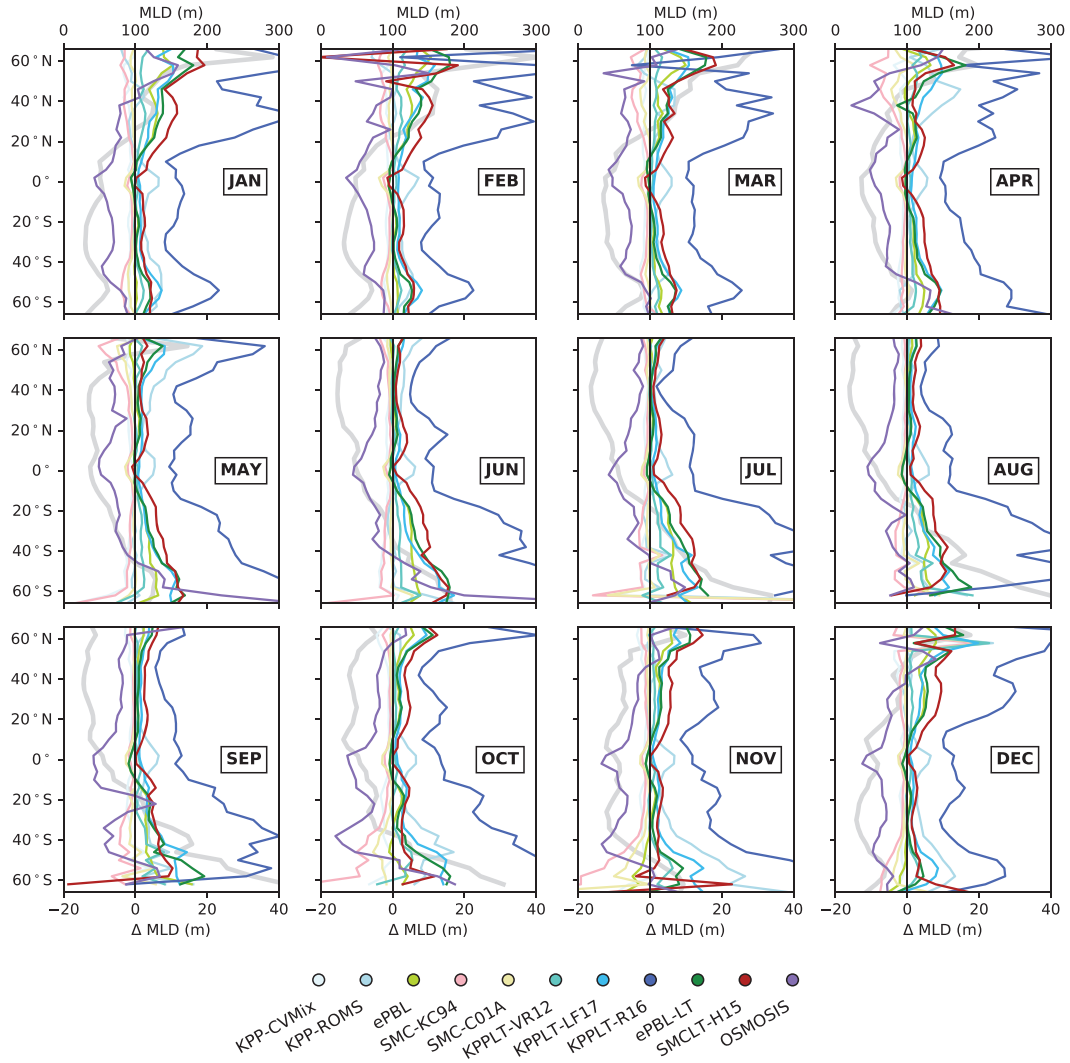

Figure 7. Comparison of zonal mean mixed layer depth (MLD; m) for each month. Thick line in gray shows the zonal mean MLD averaged over all the non-Langmuir cases (reference MLD; upper *x* axis). Thin lines in color show the difference of zonal mean MLD from the reference (Δ MLD; m) for each of the 11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

4.3.1. Comparisons by Region and Season

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

As in the ocean station comparison, significant discrepancies are seen among the Langmuir schemes, as well as among the non-Langmuir schemes. Of the five non-Langmuir schemes, SMC-C01A is closest to the reference MLD. ePBL gives systematically deeper MLD and SMC-KC94 systematically shallower, both with the differences from the mean approximately proportional to the mean MLD, that is, greater differences where the mean MLD is deep. Interestingly, the two non-Langmuir versions of KPP are at the opposite ends of the distribution, with KPP-CVMix having the shallower MLD and KPP-ROMS having deeper. The differences between these two appear smaller in winter hemispheres but substantial in summer hemispheres.

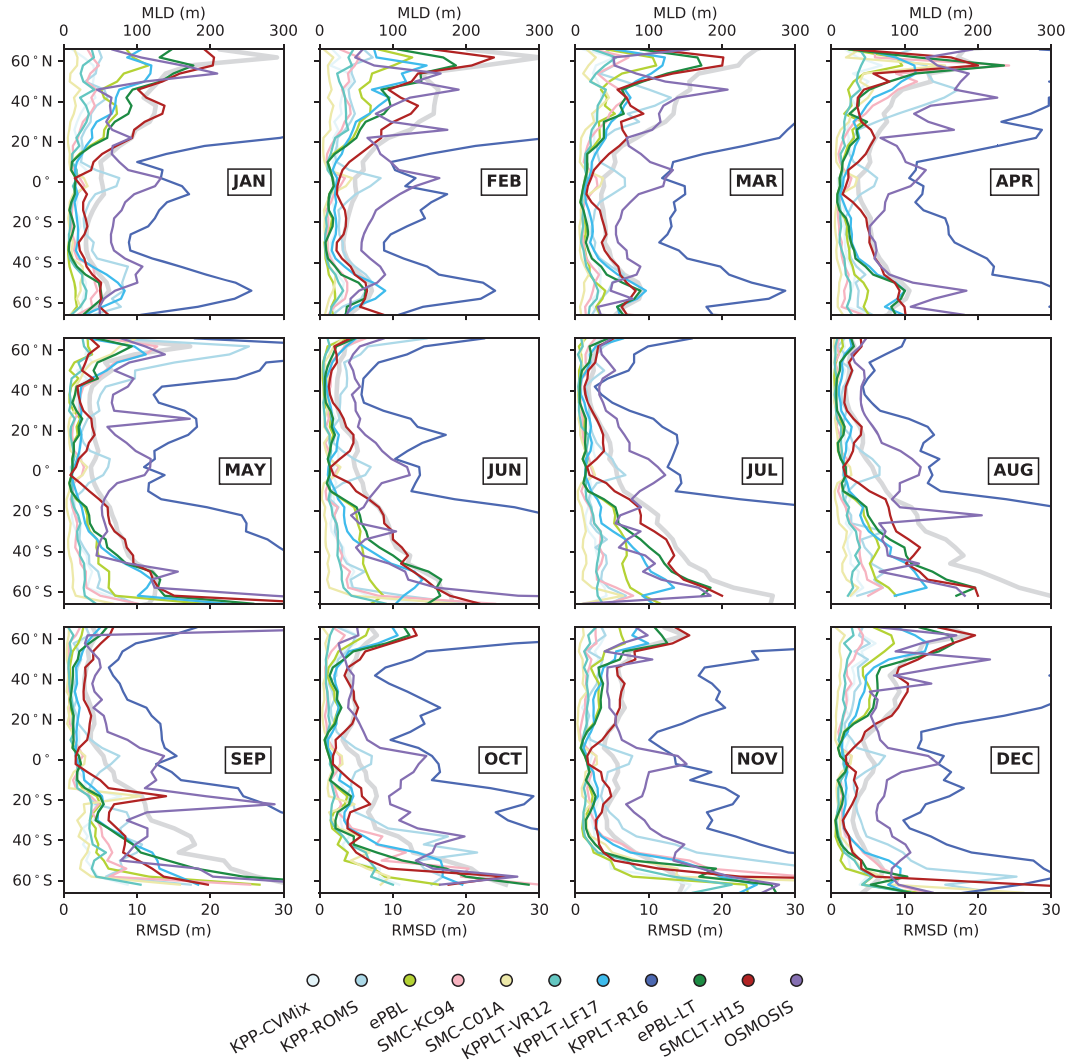

Figure 8. Same as Figure 7 but the thin lines show the root mean square differences (RMSD; m; lower *x* axis) from the reference MLD (gray thick line, upper *x* axis).

This result suggests that both the physical and numerical formulations of each scheme have a nontrivial impact on a global scale.

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

These discrepancies among schemes are consistent throughout the annual cycle, as shown by the zonal mean MLD for each month in Figure 7 and their root mean square (RMS) differences from the reference MLD in Figure 8. KPPLT-R16 gives the deepest MLD at all latitudes for all months. OSMOSIS gives the shallowest MLD except in the high latitudes, especially in winter. All other schemes seem to agree in simulated MLD in low latitudes for all months and in Northern Hemisphere middle to high latitudes in summer, while they disagree significantly in the Southern Ocean all year and Northern Hemisphere middle to high latitudes from October to May. Interestingly, KPP-ROMS gives up to 10-m-deeper MLD than other schemes

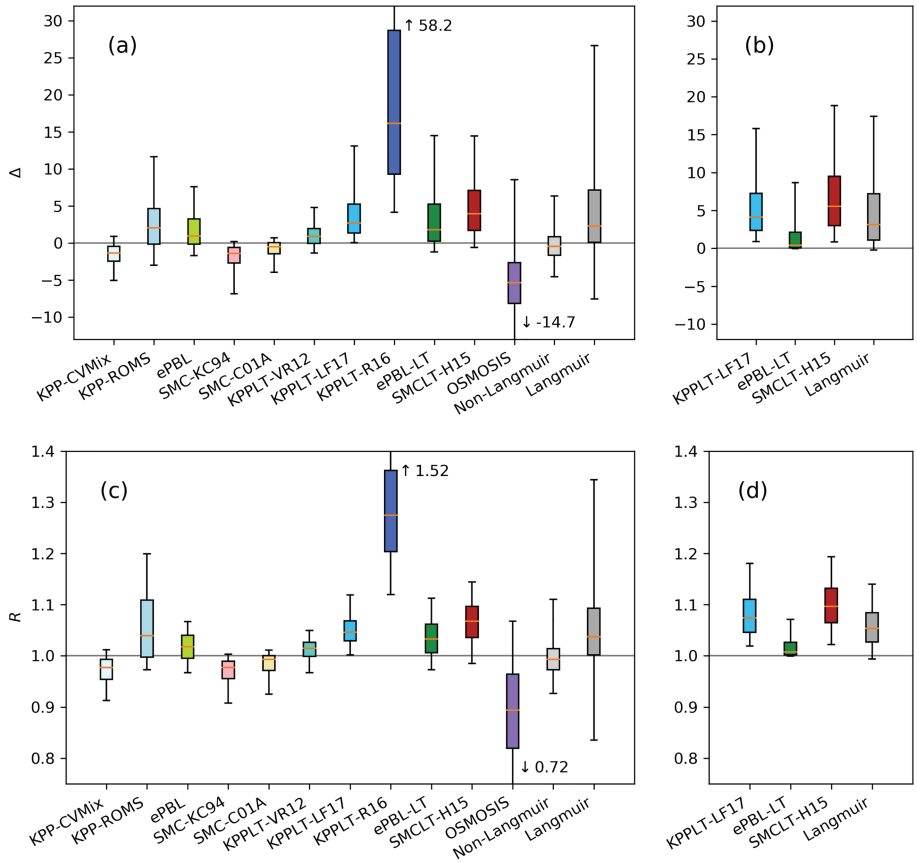

Figure 9. (a) Distribution of the difference of monthly mean mixed layer depth (MLD; m) from the reference MLD, $\Delta = \mathrm{MLD} - \mathrm{MLD_{ref}}$, for each scheme (colored), and all non-Langmuir turbulence and Langmuir schemes together, respectively (light and dark gray). The box marks the 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, $\Delta = \mathrm{MLD_{LT}} - \mathrm{MLD_{nLT}}$, 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, $R = \mathrm{MLD}/\mathrm{MLD_{ref}}$. (d) Same as c but for Langmuir schemes normalized by their non-Langmuir counterparts rather than the overall reference, $R = \mathrm{MLD_{LT}}/\mathrm{MLD_{nLT}}$, for KPPLT-LF17, ePBL-LT, SMCLT-H15, and the mean over all Langmuir schemes.

in this group near the equator (Figure 7). This behavior might be related to the strong vertical shear of near surface current in the equatorial Pacific picked up by the integral of the shear magnitude in KPP-ROMS (see equation (A3)).

A summary of the differences of simulated MLD among all the 11 schemes is presented in Figure 9. Figure 9a shows the distribution of the differences of monthly mean MLD in each scheme from the reference MLD, whereas Figure 9c shows the distribution of the ratios of these two. Again, significant differences are seen among both the Langmuir schemes and non-Langmuir schemes. Generally, we find greater spread in the Langmuir schemes than in the non-Langmuir schemes (see the distributions shown by the box and whisker in light and dark gray), with significant overlap. In addition, different categories of schemes show slightly different effects of Langmuir turbulence, as seen by comparing boxes with similar colors.

Figures 9b and 9d 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, and 9c), results of SMC-KC94 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

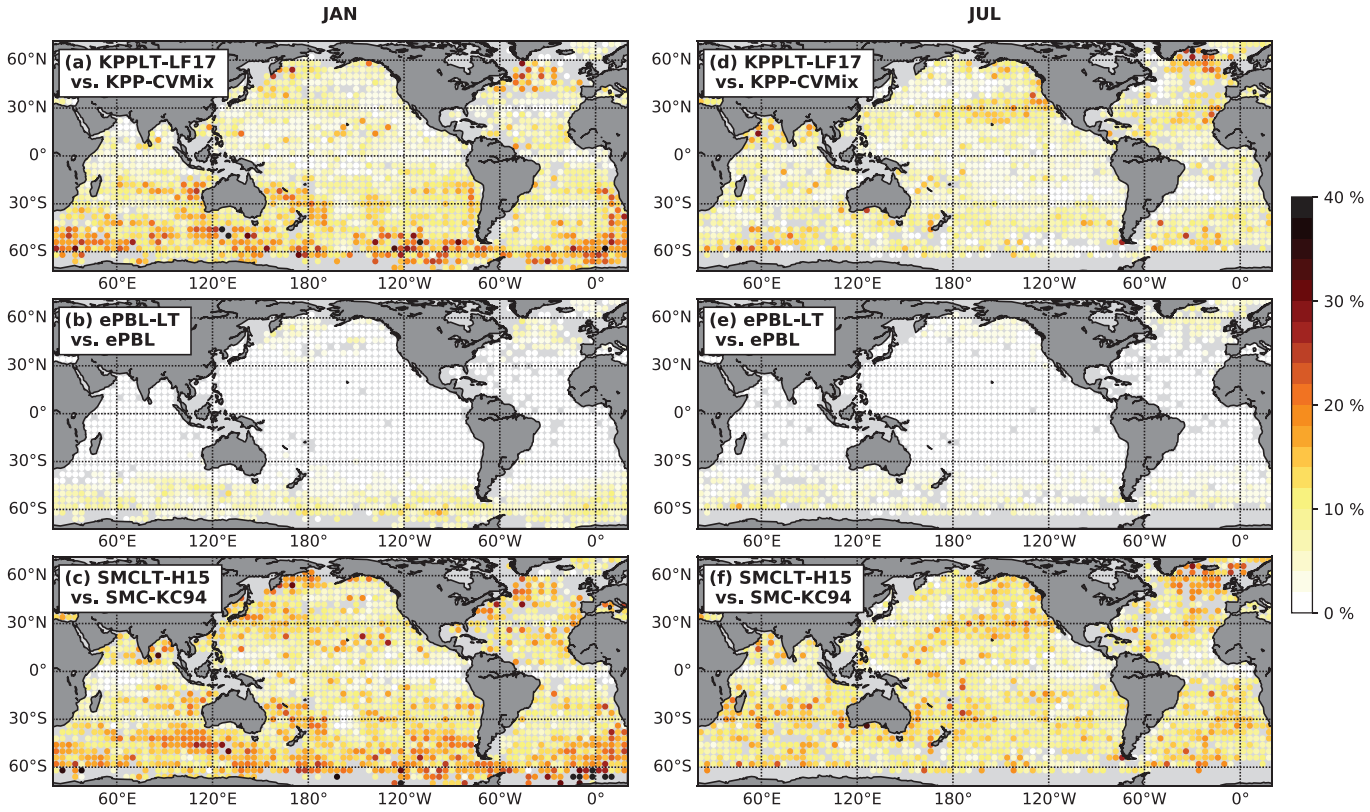

Figure 10. Maps of percentage change in simulated (a–c) January and (d–f) July mean mixed layer depth when including Langmuir turbulence, estimated by comparing (a and d) KPPLT-LF17 and KPP-CVMix; (b and e) ePBL-LT and ePBL; and (c and f) SMCLT-H15 and SMC-KC94.

in the simulated January and July mean MLD in KPPLT-LF17, ePBL-LT, and SMCLT-H15, as compared to KPP-CVMix, ePBL, and SMC-KC94, 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 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 dissipation scaling in equation (11). As shown by colors in Figure 11, the seven forcing regimes include three single forcing regimes for shear turbulence (S), Langmuir turbulence (L), and convective turbulence (C), where the contributions from the other two kinds of turbulence are both less than 25%; three combined forcing regimes (SL, SC, and LC), where both kinds of turbulence contribute more than 25% while the other contributes less than 25%; and one mixed forcing regime (SLC), where contributions of three kinds of turbulence are all greater than 25%. It is clear from the overlaid probability distribution that convection, Langmuir turbulence, and their combination are the dominant regimes in the 3-houly JRA55-do data.

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

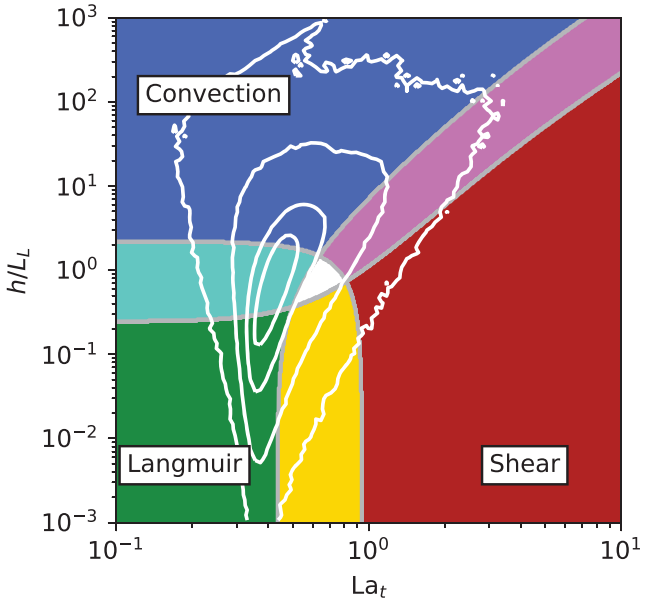

Figure 11. The same La_t - h/L_L parameter space as in Figure 1a divided into seven regimes according to the contribution of each of the three forcing mechanisms to the total turbulent dissipation (Red: S; Green: L; Blue: C; Magenta: SC; Yellow: SL; Cyan: LC; White: SLC). See the text for the criterion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

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

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

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

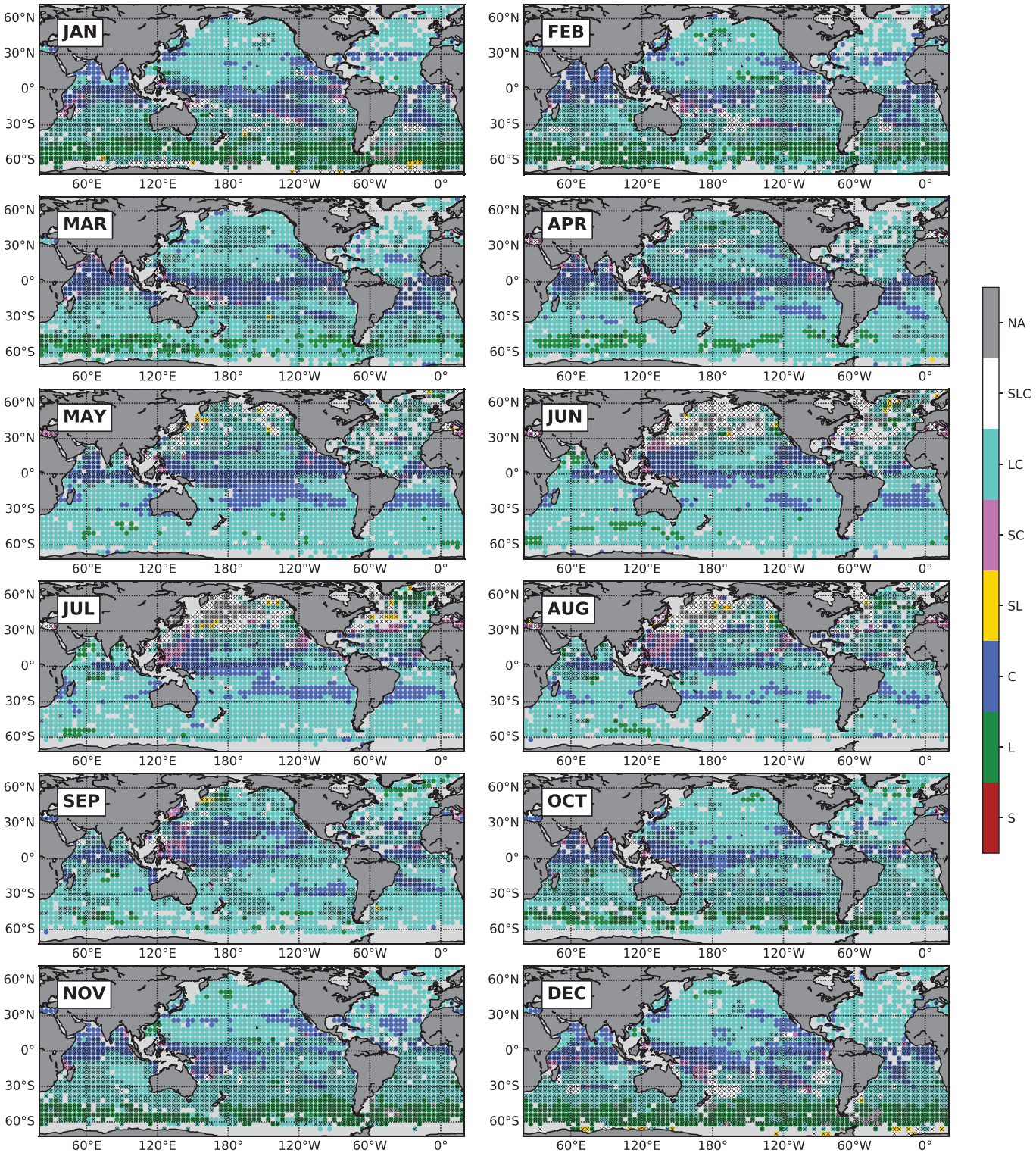

Figure 12. Maps of surface forcing regimes for each month, color coded according to the regime diagram in Figure 11, estimated from the JRA55-do forced GOTM5 simulations. Seven forcing regimes, shear turbulence (S), Langmuir turbulence (L), convective turbulence (C), combined shear and Langmuir turbulence (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 ×.

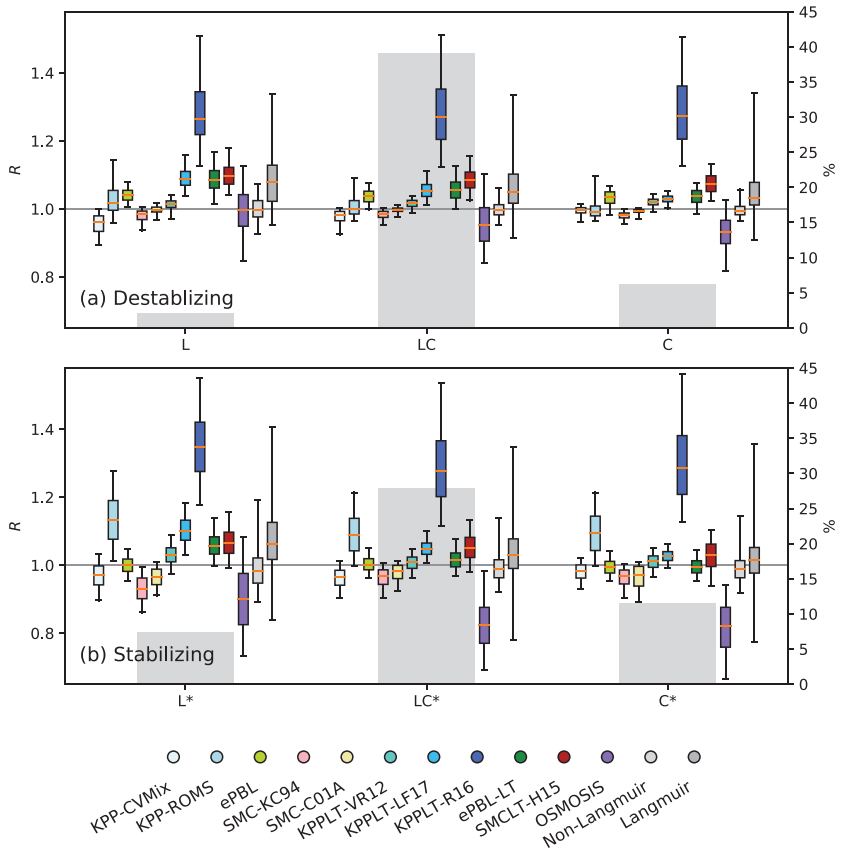

Figure 13. Distribution of the ratio of monthly mean MLD to the reference MLD, $R = \mathrm{MLD}/\mathrm{MLD}_{\mathrm{ref}}$, for different schemes (same as Figure 9c) in different surface forcing regimes. (a) Langmuir (L), Langmuir + Convection (LC), and Convection (C) regimes, with destabilizing monthly 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.

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

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

Consistently, the relative effects of Langmuir turbulence revealed by comparing individual Langmuir schemes to their non-Langmuir counterparts (Figure 14) are most prominent in L and L* regimes, with MLD deepening of 9% on average and 2% to 17% for 90% confidence range. It is also interesting to note that, in C and C* regimes, ePBL-LT converges to ePBL as expected; KPPLT-LF17 still gives deeper MLD than

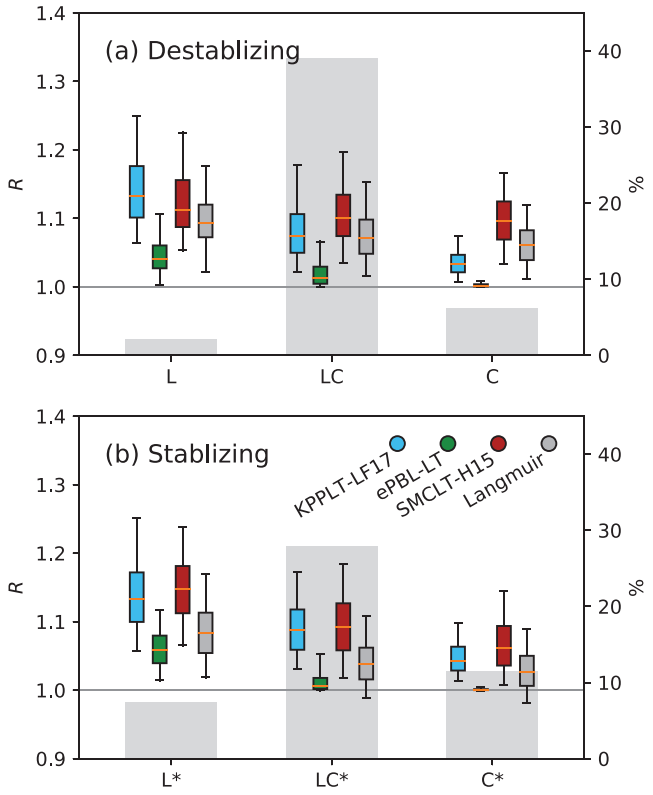

Figure 14. Same as Figure 9d but for different surface forcing regimes as described in the caption of Figure 13.

KPP-CVMix, but significantly less so than in other regimes, whereas SMCLT-H15 shows the least sensitivity to the transition from L and L* regimes to C and C* regimes (consistent with Figure 10).

4.4. Test Case 3: Evaluation Against Idealized LES

While the single-column simulations used here are not directly comparable to observations due to the lack of horizontal processes and vertical advection, they are directly comparable to many of the idealized LES results in which horizontally homogeneous turbulence is assumed. Here we choose two recent idealized LES studies of Langmuir turbulence, Li and Fox-Kemper (2017) and Reichl et al. (2016), to compare the behaviors of different schemes when simulating the entrainment process under (1) constant surface forcing of mixed Langmuir turbulence and convection and (2) transient surface forcing of typical wind and Langmuir turbulence during an idealized hurricane, respectively. Other LES studies of Langmuir turbulence under more realistic transient forcing, such as a mixed layer deepening event in Kukulka et al. (2009), diurnal restratification events in Kukulka et al. (2013), and at the Southern Ocean Flux Station in Large et al. (2019), are also excellent reference cases, but these more complex LES studies require more details than this comparison allows.

4.4.1. Entrainment

This section compares the 11 schemes to LES in simulating the entrainment by mixed Langmuir turbulence and convection in Li and Fox-Kemper (2017). Under constant destabilizing surface forcing conditions, the mixed layer keeps deepening as a result of the entrainment process, eroding the constant stratification below and converting TKE into potential energy. In a quasi-equilibrium state, the rate of mixed 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 sections 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 (Figure 15a; 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 is expected as ePBL-LT was tuned against the same set of LES to match both the rate of change in potential energy and the rate of change in SST (Reichl & Li, 2019).

While the scaling of entrainment buoyancy flux in KPPLT-LF17 was also derived from the same set of LES, KPPLT-LF17 suffers from the underlying strong sensitivities of KPP to numerical details as discussed in Van Roekel et al. (2018) and Appendix A1. Therefore, the results of KPPLT-LF17 (as well as KPPLT-VR12) shown in this figure depend strongly on how well the stratification at the base of the boundary layer is resolved and thereby the choice of numerical details in KPP. With the present KPP setup following the recommendation

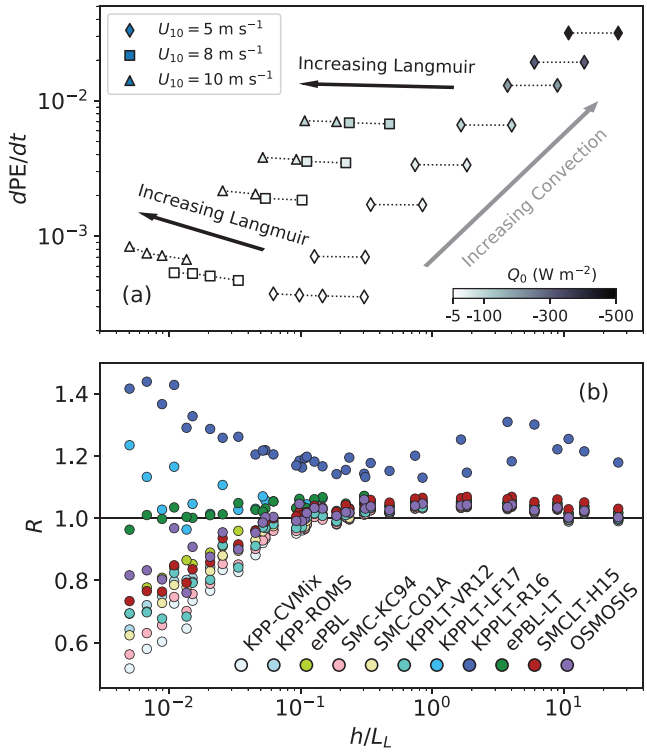

Figure 15. The rate of change in potential energy in the entrainment cases of Li and Fox-Kemper (2017) in LES and schemes. (a) The LES rate of change in potential energy (W/m^2) versus the parameter h/L_L. Different LES cases with the same surface cooling (color) and wind forcing (symbols) are grouped 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.

of Van Roekel et al. (2018), KPPLT-LF17 gives too strong entrainment as compared to LES. Sensitivity tests show that, with an alternative KPP setup (see the discussion in Appendix A1), KPPLT-LF17 may give weaker entrainment in Langmuir turbulence regime (e.g., seen in Reichl & Li, 2019). However, it is important to note that the scaling relationships used in both KPPLT-LF17 and KPPLT-VR12 were derived without any assumptions of such numerical details in KPP. The entrainment is generally too strong in KPPLT-R16 in both convection and Langmuir turbulence regimes, while too weak in KPPLT-VR12, SMCLT-H15, and OSMOSIS in the Langmuir turbulence regime.

4.4.2. Idealized Hurricane

This section compares the 11 schemes to LES in an idealized hurricane case of Reichl et al. (2016). The most significant feature in this case is rapid mixed layer deepening under strong transient hurricane wind (up to 65 m/s) and associated strong waves (La$_t$ → 0.3 during peak wind). The time series of the water-side friction velocity u_*, wind direction, and turbulent Langmuir number La$_t$ are shown in Figure 16a. The time evolution of the temperature profile is shown in Figure 16b. Note that the surface buoyancy flux was set to nearly neutral (5 W/m^2 surface cooling) throughout the run. So unlike the previous section, the convection does not play a significant role in driving the entrainment here.

Figures 16c and 16d show the time series of simulated SST and change in potential energy in LES (gray dashed lines) and GOTM5 simulations with different schemes (shown by the differences from the LES, solid lines in color), in which significant spread is seen for both. Except KPP-ROMS, all schemes underpredict the deepening of the mixed layer during the hurricane, resulting in warmer SST and smaller change of potential energy than LES after the hurricane. But the Langmuir schemes outperform the non-Langmuir turbulence ones. As expected, KPPLT-R16 performs the best in correctly simulating the SST, as it is tuned against a set of idealized hurricane LES of Reichl et al. (2016) including this one, though slightly underpredicts the change

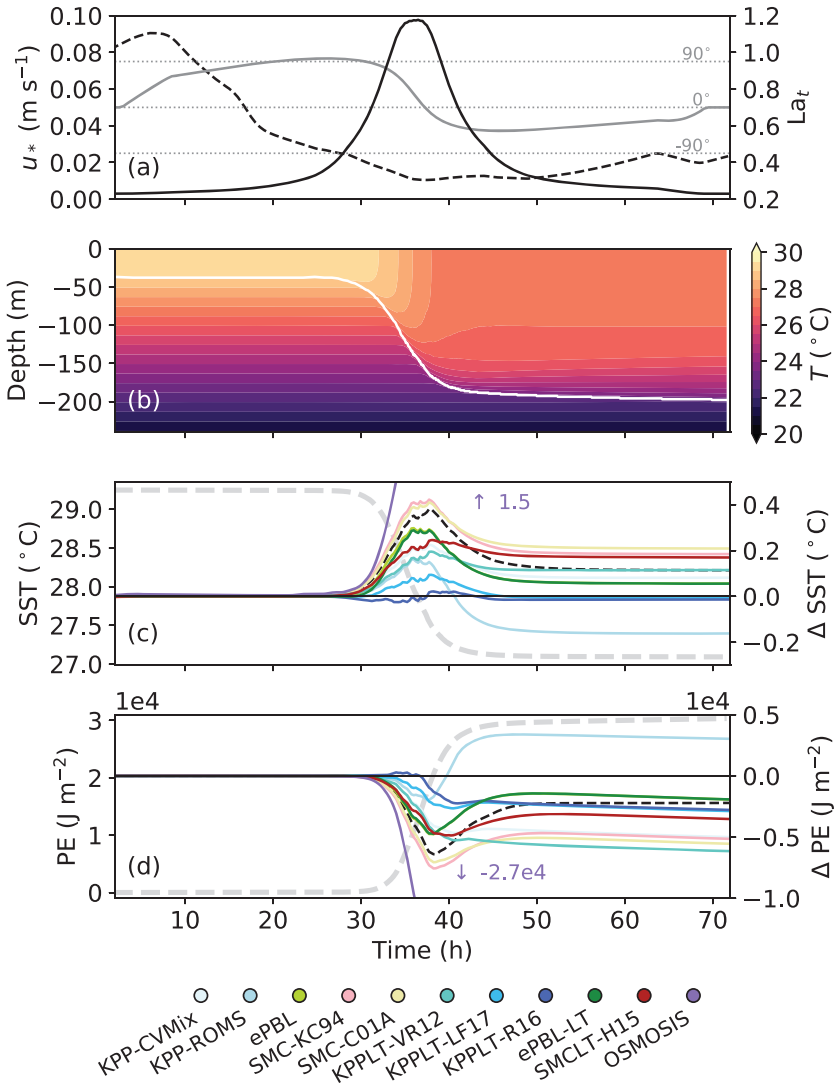

Figure 16. Comparison of different schemes to LES in an idealized hurricane case of Reichl et al. (2016). (a) Time series of friction velocity (solid line in black, vertical axis on the left), turbulent Langmuir number (dashed line in black, vertical axis on the right) and wind direction (solid line 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.

in potential energy. KPPLT-LF17 also correctly predicts the SST after the hurricane, but is slightly too warm in SST during the hurricane, and also underpredicts the change in potential energy. ePBL-LT performs better in predicting the change in potential energy after the hurricane than KPPLT-R16 and KPPLT-LF17, but performs less well during the mixed layer deepening phase, and also overpredicts the SST. SMCLT-H15 outperforms its non-Langmuir turbulence counterpart, SMC-KC94, but it is farther away from the LES solution than the non-Langmuir LES case. The OSMOSIS scheme performs poorly in this case, giving overly warm SST and too little change in potential energy, probably due to that the effects of resolved shear are not currently represented in this implementation of the OSMOSIS scheme. Note that these biases of the simulated SST and change in potential energy are almost as big as the overall change during the simulation and larger than the differences between LES with and without the wave forcing driving Langmuir turbulence (dashed line in black).

Overall, the effects of wave forcing are apparent, affecting both the schemes and the LES. Some schemes perform badly in this test (KPP-ROMS, OSMOSIS, KPPLT-VR12, and SMCLT-H15). KPPLT-LF17, ePBL-LT, and KPPLT-R16 schemes are able to capture 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 the 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; D'Asaro et al., 2014; Tseng & D'Asaro, 2004) and LES (Grant & Belcher, 2009; McWilliams et al., 1997). Refining the enhancement factor formula to account for additional properties of Langmuir turbulence, for example, penetration depth relative to the boundary layer depth (Harcourt & D'Asaro, 2008; Kukulka & Harcourt, 2017) and direction relative to the wind (McWilliams et al., 2014; Van Roekel et al., 2012), 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, that is, adding additional degrees of freedom, result in only small changes in the simulated monthly mean MLD (e.g., Li et al., 2017). Thus, other effects of Langmuir turbulence, such as entrainment, need to be considered on top of the enhanced turbulent diffusivity.

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

KPPLT-R16 also considers the influence of Langmuir turbulence on the entrainment separate from its enhancement effect on the turbulent diffusivity but in a different formulation than KPPLT-LF17. In particular, it explicitly uses the Lagrangian shear instead of the Eulerian shear to parameterize the entrainment due to Stokes drift modulated resolved shear and uses a different enhancement factor for the entrainment due to unresolved shear associated with Langmuir turbulence. It further retunes the critical Richardson number to remove any implicit effects of Langmuir turbulence previously included in the tuning parameters of KPP (Reichl et al., 2016). However, KPPLT-R16 was tuned against LES of idealized hurricanes with neutral surface stability conditions, and it lacks explicit considerations of scenarios where surface buoyancy flux is destabilizing. As a result, it greatly overestimates the deepening of mixed layer in most test cases shown here except the hurricane case. Improvements to KPPLT-R16 can be made by explicitly considering the effect of surface stability condition on its enhancement factor to the unresolved shear as in KPPLT-LF17.

In general, all the above three versions of KPPLT are expected to improve considering the other effects of Langmuir turbulence introduced in section 2. It should also be noted though that they can be affected by numerical details in a manner similar to the underlying KPP scheme (e.g., Van Roekel et al., 2018).

The strength of ePBL-LT is that it explicitly sets the rate of mechanical energy conversion to potential energy by mixing, which helps alleviate sensitivity to vertical resolution and time step. In comparison to LES here, ePBL-LT performs the best in the entrainment cases and reasonably well in the hurricane case. In other test cases, it gives similar results as KPPLT-LF17, though the difference between ePBL-LT and ePBL is less in magnitude than the differences between KPPLT-LF17 and KPP-CVMix, probably due to the different tuning strategies. ePBL-LT and ePBL were tuned to LES with and without Langmuir turbulence, respectively, whereas KPP-CVMix was tuned to observations that probably included Langmuir to some degree, and KPPLT-LF17 was derived by applying the Langmuir effects seen in LES to KPP-CVMix (see Appendices A1

and A2). ePBL-LT also generally performs better in numerical robustness than schemes based on KPP (see Appendix B).

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

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

KPP-CVMix and KPP-ROMS are in wide use in climate and regional modeling. The results here demonstrate that the numerical distinctions between these schemes, in particular the different algorithms to diagnose the boundary layer depth, are consequential and KPP-ROMS has difficulty in comparison to the LES cases chosen here, greater sensitivity to vertical and time step resolution, and exceptionally strong sensitivity to shear at the mixed layer base. Collectively, these results suggest that the KPP-CVMix formulation is more trustworthy, especially with the recommendations of Van Roekel et al. (2018). Furthermore, it is clear that non-Langmuir schemes are systematically inconsistent with LES including wave forcing; thus, KPP-ROMS needs a Langmuir version.

5.2. Limitations of the Comparisons

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

Consistent surface forcing data are used to force simulations with different Langmuir schemes. However, due to their different design, different Langmuir schemes may use the same surface forcing data differently. For example, the full profile of Stokes drift is used in KPPLT-R16 and SMCLT-H15, whereas only its surface value and decay depth are used in OSMOSIS, in which an exponential profile is assumed, even though the full profile tends to decay faster than an exponential with depth (Webb & Fox-Webb, 2011, 2015). The other three Langmuir schemes use the surface layer averaged Langmuir number, incorporating some integrated information of the profile shape of Stokes drift in addition to its surface value. The directional information of Stokes drift is only used in KPPLT-VR12 and partially in KPPLT-LF17. These inconsistencies may contribute to some of the discrepancies among schemes seen here, and a thoughtful refactoring of any of the schemes here to take advantage of good ideas in other schemes is likely to be beneficial. In addition, the limited quality of the surface forcing data may also complicate the interpretation of the differences in the comparison results. Comparison between JRA55-do and CORE-II forcing is underway to quantify these uncertainties. Agreement among schemes does indicate that the schemes perform similarly under this forcing but may also result from the lack of important features in the forcing data to reveal the different physics, such as sea state effects on stresses (e.g., Janssen, 1989; Shimura et al., 2017) and submesoscale-OSBL interactions

(e.g., Callies & Ferrari, 2018; Hamlington et al., 2014; Suzuki et al., 2016), which are known to be important but are outside of the forcing considered here. This comparison has nonetheless demonstrated why we need global atmospheric forcing data sets such as CORE-II and JRA55-do to cover typical surface forcing conditions over the global ocean. Continuing improvements in these data sets reduce these uncertainties, which is why we prefer the higher space and time resolution of JRA55-do over its predecessor CORE-II. That is, JRA55-do should have a better representation of high-frequency and extreme events.

5.3. Future Directions

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

An immediate question following the evaluation step is how to improve the parameterizations of Langmuir turbulence. The greater uncertainties in Figure 13b suggest that more LES and observational studies of Langmuir turbulence in regimes with various stabilizing surface buoyancy flux, such as diurnal heating, are needed following Min and Noh (2004), Pearson et al. (2015), and Walker et al. (2016). As shown in the regime diagram of Figure 1b, the parameter space covered by these equilibrium LES studies is rather limited. In addition, better understanding of the extradimensions that are not represented in the regime diagram of Figure 1 may also be important in explaining the discrepancies among these schemes. It is perhaps easiest to cross-pollinate these ideas among these schemes and study the consequences in comparison to studies quantifying their effects in LES, such as penetration depth of Stokes drift (Harcourt & D'Asaro, 2008; Kukulka & Harcourt, 2017), wind and wave misalignment (Van Roekel et al., 2012), Earth's rotation (e.g., Liu et al., 2018), and interactions between Langmuir turbulence with other processes such as submesoscale eddies and fronts (e.g., Hamlington et al., 2014; Suzuki et al., 2016), inertial currents (e.g., Wang et al., 2018), and combined effects of sensible, latent, and penetrative heat fluxes (e.g., Kukulka et al., 2013; Min & Noh, 2004; Pearson et al., 2015).

Another interesting question is related to the sensitivity of each scheme to the surface forcing. Preliminary comparisons between GOTM5 simulations forced by JRA55-do and CORE-II show that the relative discrepancies among schemes seen here are insensitive to the use of JRA55-do or CORE-II data sets. However, these two data sets do result in significant differences in the simulated mean state (e.g., the MLD), which is comparable to the discrepancies among schemes. A closer comparison between GOTM5 simulations forced by the two data sets is ongoing and will be helpful in this regard. Free from complicated interactions among processes as in an OGCM, such single-column simulations may also assist the analyses and interpretation of global OGCM simulations forced by CORE-II 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 nonlocal 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 as follows: (i) idealized conditions commonly used in LES studies of Langmuir turbulence; (ii) realistic conditions based on field measurements at ocean stations; and (iii) climatological forcings based on global atmospheric states targeted for use in forcing global ocean-sea ice models.

The extended GOTM5 code and test suite developed here (see Appendix D for more details) represent an initial effort toward 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 MLD (pointwise in monthly means, 6%, −1% to 14% for 90% confidence range or 5.2 m, −0.2 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.
- Correctly predicting the entrainment rate in the presence of Langmuir turbulence remains challenging for most of the Langmuir schemes, even under idealized steady forcing. It is even more difficult under transient surface forcing conditions.
- Many important regimes in the parameter space of surface forcing that are typical in the global ocean have not been well explored in LES. For example, previous LES studies of Langmuir turbulence have mostly focused on neutral surface forcing with only a few exceptions, whereas the real world appears to mostly lie in the combined Langmuir turbulence and convection, and Langmuir turbulence with stabilizing surface forcing regimes. More LES and observational studies in these regimes and the regions and seasons indicated in Figure 12 are required to better constrain the effects of Langmuir turbulence and boundary layer turbulence in general.
- The formulation and numerical implementation method of a scheme, for example, KPP-CVMix versus KPP-ROMS, can also lead to very different results or different numerical robustness even though the underlying physical foundation is the same.

Although offering more questions than answers, the results of this intercomparison provide guidance for further research into the parameterization of ocean boundary layer mixing. Future efforts will benefit from a focus on forcing regimes where the present suite of mixing schemes diverge, with an emphasis on strategies presented here to diagnose skills and differences. The disagreement of the schemes in even simple, idealized regimes suggests that the idealized process study approach is far from exhausted. Furthermore, the complexities of the realistic simulations suggest a broad parameter space that remains relatively unexplored. Continued development of robust and accurate parameterizations remains a critical endeavor for advancing regional and global model simulations.

Appendix A: Description and Implementation of Different Schemes

A1. KPP and KPPLT

Here we only briefly review the key ingredients of KPP that are relevant to the modifications to include the effects of Langmuir turbulence in the three KPPLT variants (KPPLT-VR12, KPPLT-LF17, and KPPLT-R16) in our comparison suite. Other components of KPP, for example, the nonlocal term, are treated the same among all the KPP variants in this study (KPP-CVMix, KPP-ROMS, and the above three versions of KPPLT) and therefore excluded in the discussion here for brevity. The full description of KPP can be found in Large et al. (1994) and Van Roekel et al. (2018), with the specific adaptation in KPP-CVMix detailed in Griffies et al. (2015).

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

$$K_\lambda(\sigma) = h_b w_\lambda(\sigma) G_\lambda(\sigma), \tag{A1}$$

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

KPP-CVMix follows the bulk Richardson number calculation of Large et al. (1994) and Griffies et al. (2015) to diagnose the boundary layer depth h_b. In this approach, the boundary layer depth is diagnosed by finding the shallowest depth where the bulk Richardson number,

$$\mathrm{Ri}_b(z) = \frac{z\left[b_r - \bar{b}(z)\right]}{\left[\boldsymbol{u}_r - \overline{\boldsymbol{u}}(z)\right]^2 + U_t^2(z)}, \tag{A2}$$

reaches a critical value, $\mathrm{Ri}_c = 0.3$. The reference velocity u_r and buoyancy b_r are found 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.

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

$$\mathrm{Cr}(z) = \int_z^0 J(z')\left[\left[\partial_z \overline{\boldsymbol{u}}\right]^2 - \frac{N^2(z')}{\mathrm{Ri}_c} - C_{\mathrm{ek}} f^2\right] dz' + \frac{U_t^2(z)}{|z|}, \tag{A3}$$

with $N^2 = \partial_z \bar{b}$ the square of the Brunt-Väisälä frequency, $J(z) = |z|/(|z| + 0.1h_b)$ a weighting function, and f the Coriolis parameter. Note that the last term in the integral with C_{ek} a constant represents the stabilizing effect of rotation, which is absent in the bulk Richardson number formula in equation (A2). See Appendix B of McWilliams et al. (2009) for more discussion on this formula.

In accordance with the observation that the vertical velocity variance within the OSBL is enhanced in the presence of Langmuir turbulence, the most straightforward modification of KPP is applying an enhancement factor $\mathcal{E}$ to the turbulent velocity scale w_λ, or $w_{\lambda L} = \mathcal{E} w_\lambda$. Following the idea of McWilliams and Sullivan (2000), various formulas of $\mathcal{E}$ as a function of Langmuir number have been proposed (e.g., Li et al., 2016; McWilliams & Sullivan, 2000; Smyth et al., 2002; Takaya et al., 2010). For simplicity, only one such KPPLT model is presented here (KPPLT-VR12). The relative differences among the different formulas of $\mathcal{E}$ are illustrated in Figure A1. In KPPLT-VR12, the enhancement factor is written as a function of the surface layer averaged and projected Langmuir number $\mathrm{La}_{\mathrm{SLP}}$ defined in (6) based on the LES work of Van Roekel et al. (2012),

$$\mathcal{E} = \left|\cos\theta_{\mathrm{wl}}\right| \left[1 + \left(3.1\,\mathrm{La}_{\mathrm{SLP}}\right)^{-2} + \left(5.4\,\mathrm{La}_{\mathrm{SLP}}\right)^{-4}\right]^{1/2}, \tag{A4}$$

where θ_{wl} is the angle between wind and Langmuir cells. It is expected from Figure A1 that enhancement factors based on McWilliams and Sullivan (2000), Smyth et al. (2002), or Takaya et al. (2010) will lead to much stronger enhanced vertical mixing than equation (A4), as shown in Li et al. (2016) and Ali et al. (2019). We note, however, that the VR12 case in Ali et al. (2019) is different from KPPLT-VR12 detailed here by the use of La_t and the additional Stokes drift term in the bulk Richardson number in KPP (see more discussions on this term in Li et al., 2016).

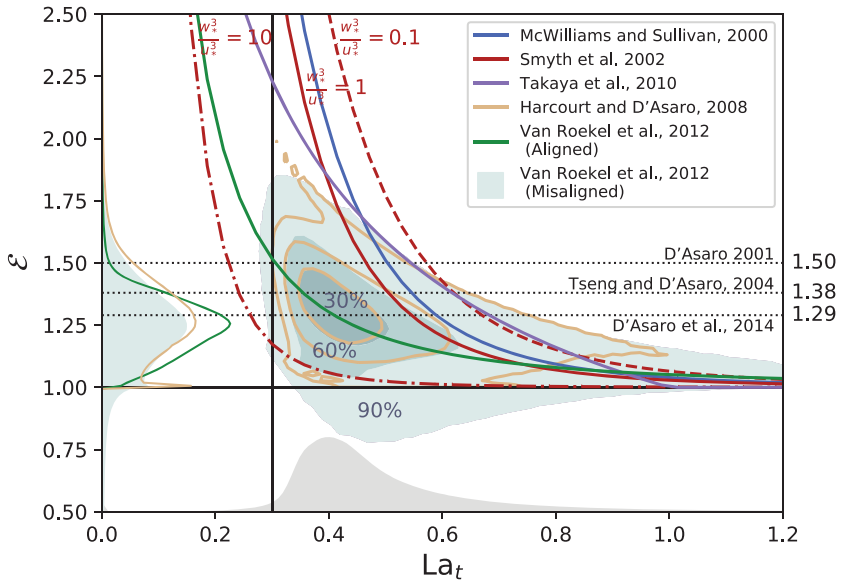

Figure A1. Comparison among the formulas of enhancement factor $\mathcal{E}$ from McWilliams and Sullivan (2000), Smyth et al. (2002), Takaya et al. (2010), Harcourt and D'Asaro (2008), aligned and misaligned versions of Van Roekel et al. (2012). Three cases with $w_*^3/u_*^3 = [0.1, 1, 10]$, with $w_* = (-B_0 h)^{1/3}$ the convective velocity scale, are shown for Smyth et al. (2002). Since the conversion from La_{SL}, or La_{SLP}, to 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 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 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 $La_t = 0.3$, respectively.

In addition to the direct effect of enhanced w_λ that increases the eddy diffusivity according to equation (A1), the entrainment at the base of the OSBL is also enhanced according to equation (A2), where the unresolved shear-term U_t^2 is a function of w_λ. Therefore, this approach simply assumes that the entrainment is affected by Langmuir turbulence in the same way as the vertical turbulent diffusion, which turns out to be insufficient (Li & Fox-Kemper, 2017). Improvements are possible by separately considering the effects of Langmuir turbulence on those processes.

Li and Fox-Kemper (2017) show that the entrainment buoyancy flux is affected differently by the presence of Langmuir turbulence and follows a different scaling law than the vertical velocity variance within the OSBL. KPPLT-LF17 therefore further incorporates this new scaling law of entrainment buoyancy flux into KPP in addition to the enhancement factor as in KPPLT-VR12 by modifying the unresolved shear term in equation (A2) to

$$U_{tL}^2(z) = \frac{C_v N(z) w_\lambda(z) |z|}{\mathrm{Ri}_c} \left[\frac{0.15 {w_*}^3 + 0.17 {u_*}^3 (1 + 0.49 \mathrm{La}_{\mathrm{SL}}^{-2})}{w_\lambda(z)^3} \right]^{1/2}, \tag{A5}$$

where C_v is a dimensionless coefficient and $w_* = (-B_0 h)^{1/3}$ is the convective velocity scale. Note that, unlike in KPPLT-VR12, in KPPLT-LF17 the enhancement factor in equation (A4) is only applied to the eddy diffusivity in equation (A1) and thereby w_λ in equation (A5) is not enhanced. It is also important to note that all coefficients in equations (A4) and (A5) are derived from LES and no tuning is required when applied in KPP.

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

$$\begin{aligned} \mathcal{E}_K &= 1 + (\mathcal{E}' - 1) G_\lambda(\sigma) / \max\left[G_\lambda(\sigma)\right], \\ \mathcal{E}' &= \min\left(2.25, 1 + \mathrm{La}_{\mathrm{SLP}}^{-1}\right), \\ \mathcal{E}_{U_t^2} &= 1 + 2.3 \mathrm{La}_{\mathrm{SLP}}^{-1/2}, \end{aligned} \tag{A6}$$

where the enhancement of the eddy diffusivity is concentrated near its peak value by the weighting function $G_\lambda(\sigma)/\max\left[G_\lambda(\sigma)\right]$. Note that here a slightly different definition of $\mathrm{La}_{\mathrm{SLP}}$ than equation (6) is used (see equation 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, $\mathrm{Ri}_c' = 0.235$, 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 mixing on the entrainment. The criterion to find h_b in KPPLT-R16 is therefore the shallowest depth where

$$\mathrm{Ri}_b'(z) = \frac{z\left[b_r - \bar{b}(z)\right]}{\left[\boldsymbol{u}_r^L - \overline{\boldsymbol{u}}^L(z)\right]^2 + U_t^2(z)\mathcal{E}_{U_t^2}} \le \mathrm{Ri}_c' = 0.235. \tag{A7}$$

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

Note that the Coriolis-Stokes force is explicitly written out in the horizontal momentum equation in KPPLT-R16 but is implicit in KPPLT-VR12 and KPPLT-LF17 where the simulated $\boldsymbol{u}$ is treated as the Lagrangian flow. In this sense, the usage of $\boldsymbol{u}$ in the bulk Richardson number in KPPLT-VR12 and KPPLT-LF17 is generally consistent with the usage of $\boldsymbol{u}^L$ in KPPLT-R16.

It should also be noted that parameters in KPPLT-R16 were tuned against LES of idealized hurricanes in Reichl et al. (2016) with nearly neutral surface conditions, in contrast to KPPLT-LF17, which were tuned against LES with weak to moderate wind and various destabilizing surface buoyancy fluxes. This may explain some of the different behaviors between these two, especially under strong destabilizing surface forcing conditions.

Consistent with Van Roekel et al. (2018), we found that the results of KPP-CVMix (thereby all the KPPLT schemes here) are quite sensitive to the choice of where N (defined at cell interfaces) is evaluated in the unresolved shear-term U_t^2 in the bulk Richardson number Ri_b (both defined at cell centers) in equation (A2). Default setup in CVMix uses the values at the interfaces below the cell centers following Danabasoglu et al. (2006), which does not behave very well with the relatively high vertical resolution ($\Delta z = 1$ m) used in our runs here. Using this setup generally results in too small N (and thereby U_t^2) and strongly affects the apparent effects of Langmuir turbulence enhanced entrainment in KPPLT-LF17 and KPPLT-R16. Ultimately, this strong sensitivity to the numerical details is related to how well the stratification at the base of the boundary layer is resolved. In theory, the maximum stratification defines the boundary layer base in the limit of pure convection, from which U_t^2 is derived. We therefore follow the recommendation of Van Roekel et al. (2018) by using the maximum value of N above and below the cell center.

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

KPP-ROMS uses a different critical condition for determining the boundary layer depth, replacing the original critical Richardson number condition (A2) by the condition (A3). As yet it includes no Langmuir turbulence parameterization effects. It is described in McWilliams et al. (2009) and Lemarié et al. (2012). It also uses the particular space-time algorithms and discretizations employed in the UCLA version of the Regional Oceanic Modeling System (ROMS; Shchepetkin & McWilliams, 2005, 2009). KPP-ROMS was implemented in GOTM5 directly as a subroutine following the KPP subroutine of UCLA ROMS. The time stepping and conservation laws are all controlled by GOTM5.

A2. ePBL and ePBL-LT

The ePBL parameterization is the OSBL turbulent mixing parameterization as described in Reichl and Hallberg (2018). It is constructed as a framework for climate simulations by introducing turbulent mixing with a relatively weak dependence on model vertical resolution and time step. The technical details that provide such capability exploit the fact that the change in column potential energy can be exactly computed from the linear integral of the change in specific volume through a vertical coordinate transformation to pressure. The buoyancy diffusion problem, assuming a known turbulent diffusion coefficient, is also linear to the close approximation that the equation of state is nearly linear over small changes in temperature and salinity. This linearity makes it possible to track the integrated change of potential energy of the entire water column as a result of changing the diffusion coefficient anywhere within the water column. The algorithm's implicit diffusion solver (e.g., the tridiagonal solver) can thus be implicitly coupled to the energetic considerations within the ePBL turbulence parameterization.

Similar to equation (A1), ePBL defines the turbulent diffusivity profile $K_\lambda(z)$ from profiles of an empirical velocity $w(z)$ and a length scale $L(z)$,

$$K_\lambda(z) = C_\lambda w(z) L(z), \tag{A8}$$

where C_λ is a dimensionless coefficient. There is no restriction in ePBL on the specific form used to estimate $w(z)$ and $L(z)$. The primary emphasis of ePBL is in setting the integral of the vertical turbulence buoyancy flux that describes potential energy change associated with turbulent mixing,

$$M_e = -\int_{-h_b}^{0} K_b(z) \max(0, \partial_z \bar{b}) dz = -\int_{-h_b}^{0} C_b w(z) L(z) \max(0, \partial_z \bar{b}) dz, \tag{A9}$$

where $\bar{b}$ is the mean buoyancy (the calling model's resolved buoyancy) and the nonlocal flux is zero. This constraint is imposed by predefining M_e via various parameterizations and then seeking solutions for the boundary layer depth h_b and the corresponding $K_b(z)$ profile that satisfy equation (A9). In practice this strategy is applied for temperature and salinity. Here we describe the specific forms of $w(z)$, $L(z)$, and M_e adopted for this study, which are based on the version of ePBL described in Reichl and Hallberg (2018) and Reichl and Li (2019).

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

$$L(z) = (z_0 + |z|) \max\left[\frac{l_b}{h_b}, \left(\frac{h_b - |z|}{h_b}\right)^\gamma\right], \tag{A10}$$

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

$$w(z) = C_{w_*}\left(\int_z^0 \overline{w'b'} dz\right)^{1/3} + (c_\mu^0)^{1/3} u_* \left[1 - a \cdot \min\left(1, \frac{|z|}{h_b}\right)\right], \tag{A11}$$

where C_{w_*} and c_μ^0 are empirical coefficients and a is a fixed vertical decay scale.

M_e is parameterized from the surface buoyancy flux B_0, the surface friction velocity u_*, the boundary layer depth h_b, the Coriolis parameter f, and for ePBL-LT, the surface layer averaged Langmuir number $\mathrm{La}_{\mathrm{SL}}$. The formula of M_e was originally optimized to represent integral properties of the turbulent mixing inferred from k-ϵ (with Schumann & Gerz, 1995, stability functions) simulations of buoyancy and shear driven turbulent boundary layers (see Reichl & Hallberg, 2018). When applied to a global ocean model, this version of ePBL yields shallow-mixing biases in mechanically driven turbulence 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 $\mathrm{La}_{\mathrm{SL}}$ was thus introduced to represent the effect of surface waves and ultimately reduce this bias (Reichl & 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

$$M_e = (m_* + m_{*\mathrm{LT}}) u_*^3 + n_* \int_{-h_b}^{0} \max(0, \overline{w'b'}) dz, \tag{A12}$$

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

$$m_* = m_{*N}\Psi + m_{*S}, \tag{A13}$$

which separates the stabilizing (S) and nonstabilizing (N) regimes as

$$m_{*N} = c_{N1}\left\{1 - \left[1 + c_{N2}\exp(-c_{N3}h_b|f|/u_*)\right]^{-1}\right\}, \tag{A14a}$$

$$m_{*S} = c_{S1}\left[\max\left(0, B_0\right)^2 h_b/u_*^5|f|\right]^{c_{S2}}, \tag{A14b}$$

$$\Psi = 1 - c_\Psi \frac{\max\left(0, -B_0\right)}{\max\left(0, -B_0\right) + 2m_{*N}u_*^3/h_b}, \tag{A14c}$$

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 $f \to 0$. The coefficient for Langmuir mixing $m_{*\mathrm{LT}}$ used here is given by

$$m_{*\mathrm{LT}} = 0.105\left[La_{\mathrm{SL}}\left(1 + 0.8\frac{|B_0|}{fu_*^2}\right)\right]^{-1}. \tag{A15}$$

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 nondimensional 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.

A3. 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 nondimensional vertical coordinates in addition to $\sigma = z/h_b$ (equation (A1)), 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_λ and nonlocal fluxes Γ_λ of scalars ($\lambda = T, S$, etc.) and of momentum ($\lambda = u, v$) include Stokes drift effects. Specifically, the diffusivity is

$$K_\lambda = \mathcal{F}_{\lambda,\phi}(u_*, w_{*L}, w_*, \sigma_i, h_b, h_m, \phi), \tag{A16}$$

where u_*, $w_{*L} = (u_*^2 u_0^S)^{1/3} = (-B_0 L_L)^{1/3}$ and w_* are the respective velocity scales of shear-driven, Langmuir, and convective turbulence; σ_i denotes multiple nondimensional 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 (A16) the functional form of $\mathcal{F}_{\lambda,\phi}$ varies for different diffused variables (λ) and between stable and unstable conditions (ϕ).

The nonlocal fluxes in the OSMOSIS scheme can be decomposed as

$$\Gamma_\lambda = \Gamma_{\text{st}} + \Gamma_{\text{buoy}} + \Gamma_{\text{trnsp}} + \Gamma_{\text{pyc}} + \Gamma_{\text{ent}}, \tag{A17}$$

where the terms on the right, respectively, denote fluxes caused by Stokes drift, buoyancy, nonlocal transport, pycnocline structure, and entrainment at the base of the boundary layer. Like the diffusivity, the functional form of the fluxes depends on the variable being transported and the stability of the boundary layer. Notably, the mechanism driving Γ_{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 (A17) are functions of velocity scales, surface fluxes, and nondimensional coordinates (σ_i) including z/δ^S where δ^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_λ and Γ_λ on the stability (ϕ), 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).

A4. 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-$q^2 l$ SMC of Mellor and Yamada (1974). These SMC models combine two nonequilibrium prognostic equations for the TKE (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 this "Level 2 1/4" model from MY2.5 and several other weak equilibrium formulations also implemented in GOTM.

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

$$\overline{\boldsymbol{u}' w'} = -S_M q l \partial_z \bar{\boldsymbol{u}} - S_M^S q l \partial_z \boldsymbol{u}^S, \tag{A18}$$

$$\overline{w' \theta'} = -S_H q l \partial_z \bar{\theta}, \tag{A19}$$

with corresponding expressions for other scalar components. Here the new eddy coefficient $K_M^S = S_M^S q l$ directs momentum flux down the Stokes gradient $\partial_z \boldsymbol{u}^S$ and the numerators and denominators of the stability functions are now polynomials in new nondimensional forcing $G_V = l^2 q^{-2} \partial_z \bar{\boldsymbol{u}} \cdot \partial_z \boldsymbol{u}^S$ and $G_S = l^2 q^{-2} |\partial_z \boldsymbol{u}^S|^2$ as well as on G_H. An important modification in Harcourt (2015) near the surface is that all appearances of the Stokes shear in G_V, G_S, and $\overline{\boldsymbol{u}' w'}$ are modified by $\partial_z \boldsymbol{u}^S \to (1 - f_z^S) \partial_z \boldsymbol{u}^S$, where f_z^S is a surface proximity

function that decays with depth from $f_z^S = 1$ at the surface. Expressions for f_z^S, S_H, S_M, and S_M^S are those provided in Harcourt (2015).

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

To improve SMCLT-H15 performance under variable and convectively unstable forcing, the crude limiters (e.g., $G_H \leq G_H^{\max}$) have been replaced by formulations more specific to the functional dependence of the stability functions on the turbulence time scale l/q. Under both unstable and stable conditions, the limiters are applied consistently across nondimensional forcings G_H, G_V, G_S, and G_M by limiting the turbulence timescale l/q to values for which the ARSM is able to provide realizable values for Reynolds tensor components and associated stability functions that are both physically possible and numerically stable. To do this, l^2/q^2 is rescaled to the equilibrium curve as $r_{\lim} l^2/q^2$ when the distance from the origin to $[G_H, G_V]$ exceeds the distance in the same direction to the equilibrium curve for $G_S = G_M = 0$ in the G_H - G_V plane. Here $r_{\lim}$ is a coefficient between 0 and 1. The equilibrium curve is approximated by shifting the curve defined by the zero of the S_H denominator $\mathrm{Den}\{S_H\} = f(G_H, G_V)$ in equation 33b of Harcourt (2015) by $\Delta G_H = 0.003$ and $\Delta G_V = 0.006$ and solving for $r_{\lim}$ as the smallest positive root of $f(r_{\lim} G_H + \Delta G_H, r_{\lim} G_V + \Delta G_V) = 0$. An additional limiter is applied for unstable $G_H > 0$ to control oscillations that arise when the closure is able to support the same scalar flux with two different scalar gradients. To avoid this, the requirement $\partial(S_H/G_H)/\partial G_H > -1$ of Burchard and Deleersnijder (2001) is approximated by rescaling the nondimensional forcing by $r_{\lim}$ whenever the solution to $f(2r_{\lim} G_H, r_{\lim} G_V) = 0$ gives $0 < r_{\lim} < 1$. This limiter unfortunately constrains the nondimensional forcing away from the equilibrium 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 nonlocal buoyancy fluxes. It appears likely that adding nonlocal 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 time scale 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^2l 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^2l 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^2l equation for both weak and quasi-equilibrium q^2-q^2l SMC models in GOTM5: Using a predicted value of q^2 to determine the dissipation term for the concurrent advance of the q^2l equation generated artificial instabilities requiring very short time steps for stability.

Appendix 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 B1. The same analyses were performed for OSMOSIS-Winter and

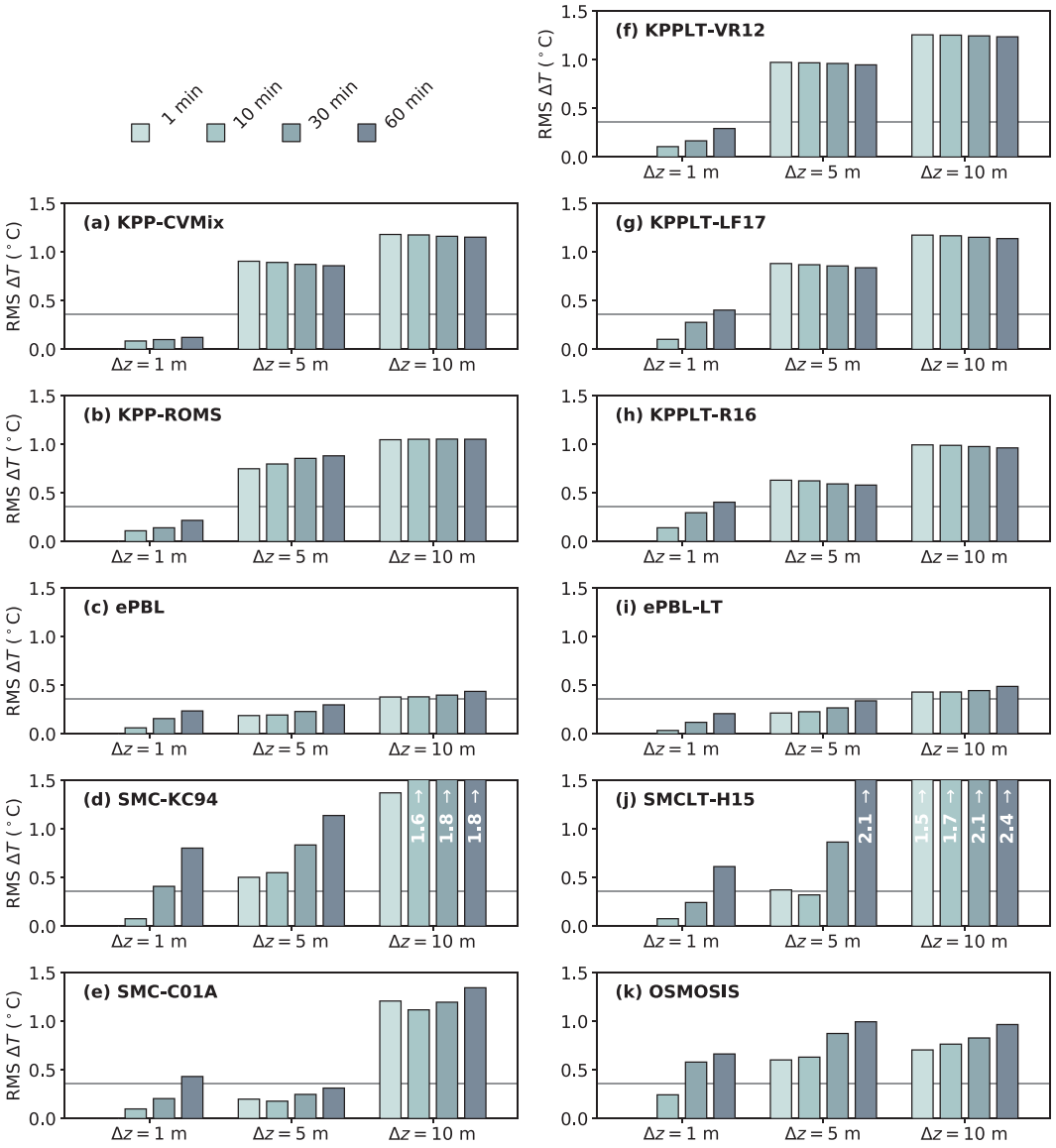

Figure B1. Sensitivity to vertical resolution and time step for different schemes in the OCS-Papa case. Each panel shows the root-mean-square (RMS) differences of the simulated temperature (° C) using different vertical resolutions and time steps as compared to the finest resolution simulation ($\Delta t = 1$ min, $\Delta z = 1$ m). Results with three different vertical grid spacings ($\Delta z = [1, 5, 10]$ m, x-axis) and four different time steps ($\Delta t = [1, 10, 30, 60]$ min, bars) are shown. The gray line in each panel marks the standard deviation of the simulated temperature (0.47 ° C) across all schemes with the finest resolution.

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 B1 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 MLD at OCS-Papa ranges from around 20 to 60 m (Figure 4) and coarsening the vertical resolution from 1 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 & Hallberg, 2018).

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

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

A second potential point of confusion is how the probabilities are calculated from the area and what the notion of distance means in Figure 1. Figure C1 shows a linear space version of Figure 1, with black points scattered through the domain to illustrate each of the 3-hourly JRA55-do data. A joint probability (p) of attaining particular values of $La_t \in [a_1, a_2]$ and $h/L_L \in [b_1, b_2]$ is the integral in linear space of the joint probability density function (ρ):

$$p\left(a_1 \le La_t \le a_2, \quad b_1 \le \frac{h}{L_L} \le b_2\right) = \int_{a_1}^{a_2} \int_{b_1}^{b_2} \rho\left(La_t, \frac{h}{L_L}\right) \quad \mathrm{d}La_t \quad \mathrm{d}\left(\frac{h}{L_L}\right). \tag{C1}$$

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

$$\begin{aligned} &p\left(c_1 \le \log_{10}(La_t) \le c_2, \quad d_1 \le \log_{10}\left(\frac{h}{L_L}\right) \le d_2\right) \\ &= \int_{c_1}^{c_2} \int_{d_1}^{d_2} \rho\left(\log_{10}(La_t), \log_{10}\left(\frac{h}{L_L}\right)\right) \quad \mathrm{d}\log_{10}(La_t) \quad \mathrm{d}\log_{10}\left(\frac{h}{L_L}\right). \end{aligned} \tag{C2}$$

However, when taking appropriate bounds, that is, $[c_1, c_2] = [\log_{10}(a_1), \log_{10}(a_2)]$ and $[d_1, d_2] = [\log_{10}(b_1), \log_{10}(b_2)]$, these two expressions are equivalent. It is clear from Figure C1 that the interpretation of whether or not values at near-neutral buoyancy forcing are within the highly probable range depends sensitively on this choice. On the other hand, as the logarithm is a monotonic function of positive real numbers, there is good agreement on the location of the joint probability density contours on the large magnitude ends of the distribution, the trouble is on the low magnitude of h/L_L ends. Relatedly, considering a point as "distant" from the high probability range depends sensitively on whether one uses log axes or not. The scatter points with h/L_L near 100 or 10^2 are clearly many, but are they close or far to those with a value at 1,000 or 10^3?

A third point of confusion is that negative (stable) h/L_L values are not geometrically or statistically connected to positive (unstable) values in the expected manner, where the negative values are located below the positive ones on a coordinate axis and the joint probability density functions are calculated including all of the data

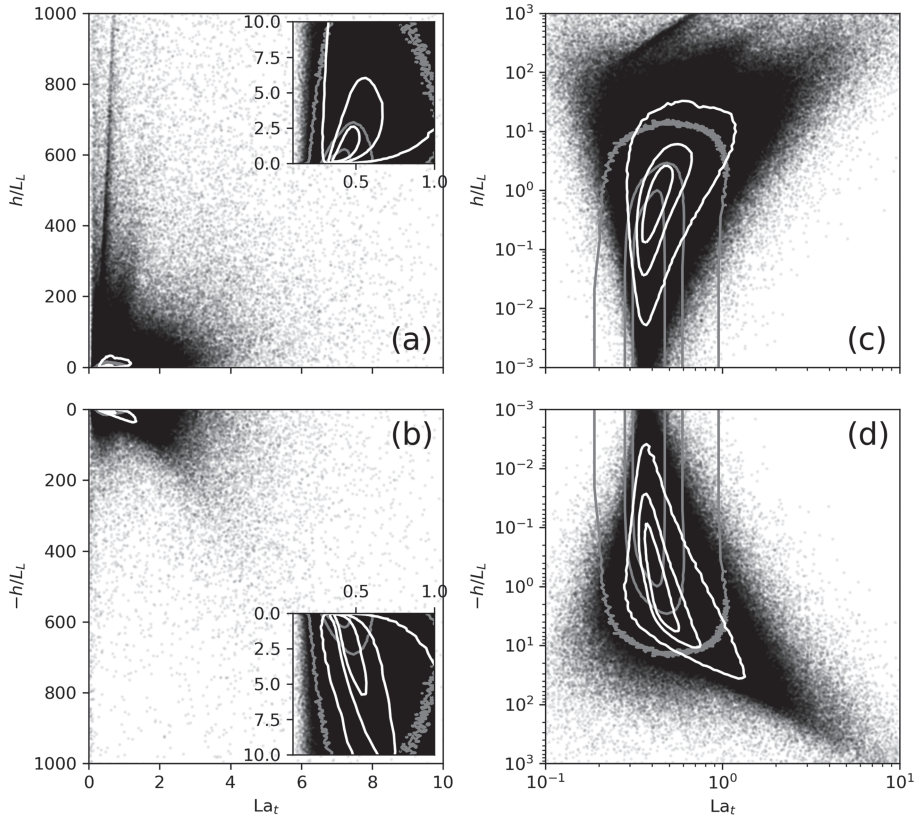

Figure C1. Joint distribution of La_t and h/L_L in (a and b) linear and (c and d) log-log spaces, estimated from 3-hourly output of JRA55-do forced GOTM5 simulations over 12 months. Black scatters show the raw data points with opacity 0.05. White contours show the joint PDF of $\log_{10}(La_t)$ and $\log_{10}(h/L_L)$ (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 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.

rather than the positive and negative subsets separately. Instead, they are laid out side by side in two panels with the absolute value of h/L_L increasing upward in both cases. This choice is made for easier comparison between the stable and unstable cases, although it is unnecessary in the linear axis case (see Figure C1).

A fourth point of confusion is why there is any lower bound to h/L_L in the log-log version of the figure. Why not values near 10^{-6}? There are many physical reasons perhaps, 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 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 data set 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? It is because we are building upon the asymptotic multiscale framework traditional in Langmuir studies (e.g., Craik & Leibovich, 1976; Grant & Belcher, 2009; McWilliams et al., 1997; 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.

Appendix D: Source Code and Data

The source code used here is hosted on GitHub (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 A1 have been merged into version v0.94b-beta of CVMix on GitHub (github.com/CVMix/CVMix-src). The test suite is also hosted on GitHub (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 online (at https://doi.org/10.26300/mknw-3842).

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 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 NSF1258907 and 1350795 and ONRN00014-17-1-2393 to B. F.-K., and Q. L. is gratefully acknowledged. B. G. R. and A. J. A. acknowledge support from the Cooperative Institute for Climate Science and the Cooperative Institute for Modeling the Earth's System at Princeton University. B. G. R. also acknowledges support from the Carbon Mitigation Initiative through Princeton's Environmental Institute at Princeton University. J. C. M. and P. W. acknowledge support from ONRN00014-14-1-0626 and N00014-15-1-2645. T. K. acknowledges support from NSF Grant OCE-1634578. R. R. H. and Z. Z. acknowledge support from NSF1558459 and 1756115, and from ONRN00014-15-1-2308 and N00014-17-1-2859. B. P. acknowledges support from ONRN00014-17-1-2963. The National Center for Atmospheric Research is a major facility sponsored by the National Science Foundation under Cooperative Agreement 1852977. All source code and data required to set up the simulations and reconstruct the figures in this paper are publicly available online (at https://doi.org/10.26300/mknw-3842).

References

Agrawal, Y. C., Terray, E. A., Donelan, M. A., Hwang, P. A., Williams, A. J., Drennan, W. M., et al. (1992). Enhanced dissipation of kinetic energy beneath surface waves. *Nature*, *359*(6392), 219–220. https://doi.org/10.1038/359219a0

Ali, A., Christensen, K. H., Breivik, Ø., Malila, M., Raj, R. P., Bertino, L., et al. (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. https://doi.org/10.1016/j.ocemod.2019.02.005

Argo (2000). *Argo float data and metadata from Global Data Assembly Centre (Argo GDAC)*. Issy-les-Moulineaux, France: SEANOE. http://doi.org/10.17882/42182

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. https://doi.org/10.1029/2001JC000922

Ayotte, K. W., Sullivan, P. P., Andrén, A., Doney, S. C., Holtslag, A. A. M., Large, W. G., et al. (1996). An Evaluation of Neutral and Convective Planetary Boundary-Layer Parameterizations Relative to Large Eddy Simulations. *Boundary-Layer Meteorology*, *79*(1–2), 131–175. https://doi.org/10.1007/BF00120078

Babanin, A. V., & Haus, B. K. (2009). On the Existence of Water Turbulence Induced by Nonbreaking Surface Waves. *Journal of Physical Oceanography*, *39*, 2675–2679. https://doi.org/10.1175/2009JPO4202.1

Beare, R. J., Macvean, M. K., Holtslag, A. A. M., Cuxart, J., Esau, I., Golaz, J.-C., et al. (2006). An Intercomparison of Large-Eddy Simulations of the Stable Boundary Layer. *Boundary-Layer Meteorology*, *118*(2), 247–272. https://doi.org/10.1007/s10546-004-2820-6

Belcher, S. E., Grant, A. L. M., Hanley, K. E., Fox-Kemper, B., Van Roekel, L., Sullivan, P. P., et al. (2012). A Global Perspective on Langmuir Turbulence in the Ocean Surface Boundary Layer. *Geophysical Research Letters*, *39*, L18605. https://doi.org/10.1029/2012GL052932

Boyer, T. P., Antonov, J. I., Baranova, O. K., Coleman, C., Garcia, H. E., Grodsky, A., et al. (2013). World Ocean Database 2013. In S. Levitus (Ed.) & A. Mishonov (Technical Ed.), NOAA Atlas NESDIS (Vol. *72*, p. 209). https://repository.library.noaa.gov/view/noaa/1291

Breivik, Ø., Bidlot, J.-R., & Janssen, P. A. E. M. (2016). A Stokes Drift Approximation Based on the Phillips Spectrum. *Ocean Modelling*, *100*, 49–56. https://doi.org/10.1016/j.ocemod.2016.01.005

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

Burchard, H., Craig, P. D., Gemmrich, J. R., van Haren, H., Mathieu, P.-P., Meier, H. E. M., et al. (2008). Observational and Numerical Modeling Methods for Quantifying Coastal Ocean Turbulence and Mixing. *Progress in Oceanography*, *76*(4), 399–442. https://doi.org/10.1016/j.pocean.2007.09.005

Burchard, H., & Deleersnijder, E. (2001). Stability of Algebraic Non-Equilibrium Second-Order Closure Models. *Ocean Modelling*, *3*(1), 33–50. https://doi.org/10.1016/S1463-5003(00)00016-0

Callies, J., & Ferrari, R. (2018). Baroclinic Instability in the Presence of Convection. *Journal of Physical Oceanography*, *48*(1), 45–60. https://doi.org/10.1175/JPO-D-17-0028.1

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

Cavaleri, L., Fox-Kemper, B., & Hemer, M. (2012). Wind Waves in the Coupled Climate System. *Bulletin of the American Meteorological Society*, *93*(11), 1651–1661. https://doi.org/10.1175/BAMS-D-11-00170.1

Chen, S., Qiao, F., Huang, C., & Song, Z. (2018). Effects of the Non-Breaking Surface Wave-Induced Vertical Mixing on Winter Mixed Layer Depth in Subtropical Regions. *Journal of Geophysical Research: Oceans*, *123*, 2934–2944. https://doi.org/10.1002/2017JC013038

Craik, A. D. D., & Leibovich, S. (1976). A Rational Model for Langmuir Circulations. *Journal of Fluid Mechanics*, *73*(3), 401–426. https://doi.org/10.1017/S0022112076001420

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. https://doi.org/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. https://doi.org/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. https://doi.org/10.1146/annurev-marine-010213-135138

D'Asaro, E. A., Thomson, J., Shcherbina, A. Y., Harcourt, R. R., Cronin, M. F., Hemer, M. A., & Fox-Kemper, B. (2014). Quantifying Upper Ocean Turbulence Driven by Surface Waves. *Geophysical Research Letters*, *41*, 102–107. https://doi.org/10.1002/2013GL058193

Damerell, G. M., Heywood, K. J., Thompson, A. F., Binetti, U., & Kaiser, J. (2016). The Vertical Structure of Upper Ocean Variability at the Porcupine Abyssal Plain during 2012–2013. *Journal of Geophysical Research: Oceans*, *121*, 3075–3089. https://doi.org/10.1002/2015JC011423

Danabasoglu, G., Large, W. G., Tribbia, J. J., Gent, P. R., Briegleb, B. P., & McWilliams, J. C. (2006). Diurnal Coupling in the Tropical Oceans of CCSM3. *Journal of Climate*, *19*(11), 2347–2365. https://doi.org/10.1175/JCLI3739.1

de Boyer Montégut, C., Madec, C., Fischer, A. S., Lazar, A., & Iudicone, D. (2004). Mixed Layer Depth over the Global Ocean: An Examination of Profile Data and a Profile-Based Climatology. *Journal of Geophysical Research*, *109*, C12003. https://doi.org/10.1029/2004JC002378

Fairall, C. W., Bradley, E. F., Rogers, D. P., Edson, J. B., & Young, G. S. (1996). Bulk Parameterization of Air-Sea Fluxes for Tropical Ocean-Global Atmosphere Coupled-Ocean Atmosphere Response Experiment. *Journal of Geophysical Research*, *101*(C2), 3747–3764. https://doi.org/10.1029/95JC03205

Fan, Y., & Griffies, S. M. (2014). Impacts of Parameterized Langmuir Turbulence and Nonbreaking Wave Mixing in Global Climate Simulations. *Journal of Climate*, *27*(12), 4752–4775. https://doi.org/10.1175/JCLI-D-13-00583.1

Galperin, B., Kantha, L. H., Hassid, S., & Rosati, A. (1988). A Quasi-Equilibrium Turbulent Energy Model for Geophysical Flows. *Journal of the Atmospheric Sciences*, *45*(1), 55–62. https://doi.org/10.1175/1520-0469(1988)045<0055:AQETEM>2.0.CO;2

Gargett, A., Wells, J., Tejada-Martínez, A. E., & Grosch, C. E. (2004). Langmuir Supercells: A Mechanism for Sediment Resuspension and Transport in Shallow Seas. *Science*, *306*(5703), 1925–1928. https://doi.org/10.1126/science.1100849

Grant, A. L. M., & Belcher, S. E. (2009). Characteristics of Langmuir Turbulence in the Ocean Mixed Layer. *Journal of Physical Oceanography*, *39*(8), 1871–1887. https://doi.org/10.1175/2009JPO4119.1

Griffies, S. M., Biastoch, A., Böning, C. W., Bryan, F. O., Danabasoglu, G., Chassignet, E. P., et al. (2009). Coordinated Ocean-Ice Reference Experiments (COREs). *Ocean Modelling*, *26*(1–2), 1–46. https://doi.org/10.1016/j.ocemod.2008.08.007

Griffies, S. M., Danabasoglu, G., Durack, P. J., Adcroft, A. J., Balaji, V., Böning, C. W., et al. (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. https://doi.org/https://doi.org/10.5194/gmd-9-3231-2016

Griffies, S. M., Levy, M., Adcroft, A. J., Danabasoglu, G., Hallberg, R. W., Jacobsen, D., et al. (2015). Theory and Numerics of the Community Ocean Vertical Mixing (CVMix) Project. (Tech. Rep.)

Hamlington, P. E., Van Roekel, L. P., Fox-Kemper, B., Julien, K., & Chini, G. P. (2014). Langmuir-Submesoscale Interactions: Descriptive Analysis of Multiscale Frontal Spindown Simulations. *Journal of Physical Oceanography*, *44*(9), 2249–2272. https://doi.org/10.1175/JPO-D-13-0139.1

Haney, S., Fox-Kemper, B., Julien, K., & Webb, A. (2015). Symmetric and Geostrophic Instabilities in the Wave-Forced Ocean Mixed Layer. *Journal of Physical Oceanography*, *45*(12), 3033–3056. https://doi.org/10.1175/JPO-D-15-0044.1

Hanley, K. E., Belcher, S. E., & Sullivan, P. P. (2010). A Global Climatology of Wind-Wave Interaction. *Journal of Physical Oceanography*, *40*(6), 1263–1282. https://doi.org/10.1175/2010JPO4377.1

Harcourt, R. R. (2013). A Second-Moment Closure Model of Langmuir Turbulence. *Journal of Physical Oceanography*, *43*(4), 673–697. https://doi.org/10.1175/JPO-D-12-0105.1

Harcourt, R. R. (2015). An Improved Second-Moment Closure Model of Langmuir Turbulence. *Journal of Physical Oceanography*, *45*(4), 84–103. https://doi.org/10.1175/JPO-D-14-0046.1

Harcourt, R. R., & D'Asaro, E. A. (2008). Large-Eddy Simulation of Langmuir Turbulence in Pure Wind Seas. *Journal of Physical Oceanography*, *38*(7), 1542–1562. https://doi.org/10.1175/2007JPO3842.1

Harvey, N. J., Hogan, R. J., & Dacre, H. F. (2015). Evaluation of boundary-layer type in a weather forecast model utilizing long-term Doppler lidar observations. *Quarterly Journal of the Royal Meteorological Society*, *141*, 1345–1353.

Jackson, L., Hallberg, R., & Legg, S. (2008). A Parameterization of Shear-Driven Turbulence for Ocean Climate Models. *Journal of Physical Oceanography*, *38*(5), 1033–1053. https://doi.org/10.1175/2007JPO3779.1

Janssen, P. A. E. M. (1989). Wave-Induced Stress and the Drag of Air Flow over Sea Waves. *Journal of Physical Oceanography*, *19*(6), 745–754. https://doi.org/10.1175/1520-0485(1989)019<0745:WISATD>2.0.CO;2

Japan Meteorological Agency/Japan (2013). JRA-55: Japanese 55-Year Reanalysis, Daily 3-Hourly and 6-Hourly Data, Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder, Colo. (Updated monthly). https://doi.org/10.5065/D6HH6H41

Kantha, L. H., & Clayson, C. A. (1994). An Improved Mixed Layer Model for Geophysical Applications. *Journal of Geophysical Research*, *99*(C12), 25,235–25,266. https://doi.org/10.1029/94JC02257

Kantha, L. H., & Clayson, C. A. (2004). On the Effect of Surface Gravity Waves on Mixing in the Oceanic Mixed Layer. *Ocean Modelling*, *6*(2), 101–124. https://doi.org/10.1016/S1463-5003(02)00062-8

Kenyon, K. E. (1969). Stokes Drift for Random Gravity Waves. *Journal of Geophysical Research*, *74*(28), 6991. https://doi.org/10.1029/JC074i028p06991

Kudryavtsev, V., Shrira, V., Dulov, V., & Malinovsky, V. (2008). On the Vertical Structure of Wind-Driven Sea Currents. *Journal of Physical Oceanography*, *38*(10), 2121–2144. https://doi.org/10.1175/2008JPO3883.1

Kukulka, T., & Harcourt, R. R. (2017). Influence of Stokes Drift Decay Scale on Langmuir Turbulence. *Journal of Physical Oceanography*, *47*(7), 1637–1656. https://doi.org/10.1175/JPO-D-16-0244.1

Kukulka, T., Plueddemann, A. J., & Sullivan, P. P. (2013). Inhibited Upper Ocean Restratification in Nonequilibrium Swell Conditions. *Geophysical Research Letters*, *40*, 3672–3676. https://doi.org/10.1002/grl.50708

Kukulka, T., Plueddemann, A. J., Trowbridge, J. H., & Sullivan, P. P. (2009). Significance of Langmuir Circulation in Upper Ocean Mixing: Comparison of Observations and Simulations. *Geophysical Research Letters*, *36*, L10603. https://doi.org/10.1029/2009GL037620

Kukulka, T., Plueddemann, A. J., Trowbridge, J. H., & Sullivan, P. P. (2010). Rapid Mixed Layer Deepening by the Combination of Langmuir and Shear Instabilities: A Case Study. *Journal of Physical Oceanography*, *40*(11), 2381–2400. https://doi.org/10.1175/2010JPO4403.1

Langmuir, I. (1938). Surface Motion of Water Induced by Wind. *Science*, *87*(2250), 119–123.

Large, W. G. (1996). An Observational and Numerical Investigation of the Climatological Heat and Salt Balances at OWS Papa. *Journal of Climate*, *9*(8), 1856–1876. https://doi.org/10.1175/1520-0442(1996)009<1856:AOANIO>2.0.CO;2

Large, W. G., & Caron, J. M. (2015). Diurnal Cycling of Sea Surface Temperature, Salinity, and Current in the CESM Coupled Climate Model. *Journal of Geophysical Research: Oceans*, *120*, 3711–3729. https://doi.org/10.1002/2014JC010691

Large, W. G., Danabasoglu, G., Doney, S. C., & McWilliams, J. C. (1997). Sensitivity to surface forcing and boundary layer mixing in a global ocean model: Annual-mean climatology. *Journal of Physical Oceanography*, *27*(11), 2418–2447.

Large, W. G., Mcwilliams, J. C., & Doney, S. C. (1994). Oceanic Vertical Mixing: A Review and a Model with a Nonlocal Boundary Layer Parameterization. *Reviews of Geophysics*, *32*(4), 363–403.

Large, W. G., Patton, E. G., DuVivier, A. K., & Sullivan, P. P. (2019). Similarity Theory in the Surface Layer of Large-Eddy Simulations of the Southern Ocean with Waves. *Journal of Physical Oceanography*, *49*(8), 2165–2187. https://doi.org/10.1175/JPO-D-18-0066.1

Large, W. G., & Yeager, S. G. (2009). The Global Climatology of an Interannually Varying Air-sea Flux Data Set. *Climate Dynamics*, *33*(2-3), 341–364. https://doi.org/10.1007/s00382-008-0441-3

Lemarié, F., Kurian, J., Shchepetkin, A. F., Jeroen Molemaker, M., Colas, F., & McWilliams, J. C. (2012). Are There Inescapable Issues Prohibiting the Use of Terrain-Following Coordinates in Climate Models? *Ocean Modelling*, *42*, 57–79. https://doi.org/10.1016/j.ocemod.2011.11.007

Li, Q. (2018). Langmuir Turbulence and Its Effects on Global Climate (Ph.D. thesis), Brown University, Providence, RI. https://doi.org/doi.org/10.26300/mj20-3x15

Li, Q., & Fox-Kemper, B. (2017). Assessing the Effects of Langmuir Turbulence on the Entrainment Buoyancy Flux in the Ocean Surface Boundary Layer. *Journal of Physical Oceanography*, *47*(12), 2863–2886. https://doi.org/10.1175/JPO-D-17-0085.1

Li, Q., Fox-Kemper, B., Breivik, Ø., & Webb, A. (2017). Statistical Models of Global Langmuir Mixing. *Ocean Modelling*, *113*, 95–114. https://doi.org/10.1016/j.ocemod.2017.03.016

Li, Q., Webb, A., Fox-Kemper, B., Craig, A., Danabasoglu, G., Large, W. G., & Vertenstein, M. (2016). Langmuir Mixing Effects on Global Climate: WAVEWATCH III in CESM. *Ocean Modelling*, *103*, 145–160. https://doi.org/10.1016/j.ocemod.2015.07.020

Liu, J., Liang, J.-H., McWilliams, J. C., Sullivan, P. P., Fan, Y., & Chen, Q. (2018). Effect of Planetary Rotation on Oceanic Surface Boundary Layer Turbulence. *Journal of Physical Oceanography*, *48*(9), 2057–2080. https://doi.org/10.1175/JPO-D-17-0150.1

Mazloff, M. R., Heimbach, P., & Wunsch, C. (2010). An Eddy-Permitting Southern Ocean State Estimate. *Journal of Physical Oceanography*, *40*(5), 880–899. https://doi.org/10.1175/2009JPO4236.1

McWilliams, J. C., & Fox-Kemper, B. (2013). Oceanic Wave-Balanced Surface Fronts and Filaments. *Journal of Fluid Mechanics*, *730*, 464–490. https://doi.org/10.1017/jfm.2013.348

McWilliams, J. C., Huckle, E., Liang, J.-H., & Sullivan, P. P. (2012). The Wavy Ekman Layer: Langmuir Circulations, Breaking Waves, and Reynolds Stress. *Journal of Physical Oceanography*, *42*(11), 1793–1816. https://doi.org/10.1175/JPO-D-12-07.1

McWilliams, J. C., Huckle, E., Liang, J.-H., & Sullivan, P. P. (2014). Langmuir Turbulence in Swell. *Journal of Physical Oceanography*, *44*(3), 870–890. https://doi.org/10.1175/JPO-D-13-0122.1

McWilliams, J. C., Huckle, E., & Shchepetkin, A. F. (2009). Buoyancy Effects in a Stratified Ekman Layer. *Journal of Physical Oceanography*, *39*, 2581–2599. https://doi.org/10.1175/2009JPO4130.1

McWilliams, J. C., & Sullivan, P. P. (2000). Vertical Mixing by Langmuir Circulations. *Spill Science and Technology Bulletin*, *6*(3), 225–237. https://doi.org/10.1016/S1353-2561(01)00041-X

McWilliams, J. C., Sullivan, P. P., & Moeng, C.-H. (1997). Langmuir Turbulence in the Ocean. *Journal of Fluid Mechanics*, *334*(1), 1–30.

Mellor, G. L., & Yamada, T. (1974). A Hierarchy of Turbulence Closure Models for Planetary Boundary Layers. *Journal of the Atmospheric Sciences*, *31*, 1791–1806.

Mellor, G. L., & Yamada, T. (1982). Development of a Turbulence Closure Model for Geophysical Fluid Problems. *Reviews of Geophysics and Space Physics*, *20*(4), 851–875.

Min, H. S., & Noh, Y. (2004). Influence of the Surface Heating on Langmuir Circulation. *Journal of Physical Oceanography*, *34*(12), 2630–2641. https://doi.org/10.1175/JPOJPO-2654.1

Moore, J. K., Lindsay, K., Doney, S. C., Long, M. C., & Misumi, K. (2013). Marine Ecosystem Dynamics and Biogeochemical Cycling in the Community Earth System Model [CESM1(BGC)]: Comparison of the 1990s with the 2090s under the RCP4.5 and RCP8.5 Scenarios. *Journal of Climate*, *26*(23), 9291–9312. https://doi.org/10.1175/JCLI-D-12-00566.1

Myrhaug, D., Wang, H., & Holmedal, L. E. (2014). Stokes Drift Estimation for Deep Water Waves Based on Short-Term Variation of Wave Conditions. *Coastal Engineering*, *88*, 27–32. https://doi.org/10.1016/j.coastaleng.2014.01.014

Noh, Y., & Lee, W. S. (2008). Mixed and Mixing Layer Depths Simulated by an OGCM. *Journal of Oceanography*, *64*(2), 217–225. https://doi.org/10.1007/s10872-008-0017-1

Noh, Y., Ok, H., Lee, E., Toyoda, T., & Hirose, N. (2016). Parameterization of Langmuir Circulation in the Ocean Mixed Layer Model Using LES and Its Application to the OGCM. *Journal of Physical Oceanography*, *46*(1), 57–78. https://doi.org/10.1175/JPO-D-14-0137.1

Pearson, B. (2018). Turbulence-Induced Anti-Stokes Flow and the Resulting Limitations of Large-Eddy Simulation. *Journal of Physical Oceanography*, *48*(1), 117–122. https://doi.org/10.1175/JPO-D-17-0208.1

Pearson, B., Grant, A. L. M., Polton, J. A., & Belcher, S. E. (2015). Langmuir Turbulence and Surface Heating in the Ocean Surface Boundary Layer. *Journal of Physical Oceanography*, *45*, 2897–2911. https://doi.org/10.1175/JPO-D-15-0018.1

Polton, J. A., & Belcher, S. E. (2007). Langmuir Turbulence and Deeply Penetrating Jets in an Unstratified Mixed Layer. *Journal of Geophysical Research*, *112*, C09020. https://doi.org/10.1029/2007JC004205

Prandtl, L. (1925). Bericht über die Entstehung der Turbulenz. *Zeitschrift f¨r angewandte Mathematik und Mechanik*, *5*, 136–139.

Price, J. F., Weller, R. A., & Pinkel, R. (1986). Diurnal Cycling: Observations and Models of the Upper Ocean Response to Diurnal Heating, Cooling, and Wind Mixing. *Journal of Geophysical Research*, *91*(C7), 8411–8427. https://doi.org/10.1029/JC091iC07p08411

Qiao, F., Yuan, Y., Deng, J., Dai, D., & Song, Z. (2016). Wave-Turbulence Interaction-Induced Vertical Mixing and Its Effects in Ocean and Climate Models. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, *374*(2065), 20150201. https://doi.org/10.1098/rsta.2015.0201

Qiao, F., Yuan, Y., Yang, Y., Zheng, Q., Xia, C., & Ma, J. (2004). Wave-Induced Mixing in the Upper Ocean: Distribution and Application to a Global Ocean Circulation Model. *Geophysical Research Letters*, *31*, L11303. https://doi.org/10.1029/2004GL019824

Reichl, B. G., & Hallberg, R. (2018). A Simplified Energetics Based Planetary Boundary Layer (ePBL) Approach for Ocean Climate Simulations. *Ocean Modelling*, *132*, 112–129. https://doi.org/10.1016/j.ocemod.2018.10.004

Reichl, B. G., & Li, Q. (2019). A Parameterization with a Constrained Potential Energy Conversion Rate of Vertical Mixing Due to Langmuir Turbulence. *Journal of Physical Oceanography*, *49*(11), 2935–2959. https://doi.org/10.1175/JPO-D-18-0258.1

Reichl, B. G., Wang, D., Hara, T., Ginis, I., & Kukulka, T. (2016). Langmuir Turbulence Parameterization in Tropical Cyclone Conditions. *Journal of Physical Oceanography*, *46*(3), 863–886. https://doi.org/10.1175/JPO-D-15-0106.1

Rodgers, K. B., Aumont, O., Mikaloff Fletcher, S. E., Plancherel, Y., Bopp, L., de Boyer Montégut, C., et al. (2014). Strong Sensitivity of Southern Ocean Carbon Uptake and Nutrient Cycling to Wind Stirring. *Biogeosciences*, *11*, 4077–4098. https://doi.org/10.5194/bg-11-4077-2014

Rodi, W. (1987). Examples of Calculation Methods for Flow and Mixing in Stratified Fluids. *Journal of Geophysical Research*, *92*(C5), 5305–5328. https://doi.org/10.1029/JC092iC05p05305

Sallée, J.-B., Shuckburgh, E., Bruneau, N., Meijers, A. J. S., Bracegirdle, T. J., & Wang, Z. (2013b). Assessment of Southern Ocean Mixed-Layer Depths in CMIP5 Models: Historical Bias and Forcing Response. *Journal of Geophysical Research: Oceans*, *118*, 1845–1862. https://doi.org/10.1002/jgrc.20157

Sallée, J.-B., Shuckburgh, E., Bruneau, N., Meijers, A. J. S., Bracegirdle, T. J., Wang, Z., & Roy, T. (2013a). Assessment of Southern Ocean Water Mass Circulation and Characteristics in CMIP5 Models: Historical Bias and Forcing Response. *Journal of Geophysical Research: Oceans*, *118*, 1830–1844. https://doi.org/10.1002/jgrc.20135

Schudlich, R. R., & Price, J. F. (1998). Observations of Seasonal Variation in the Ekman Layer. *Journal of Physical Oceanography*, *28*(6), 1187–1204. https://doi.org/10.1175/1520-0485(1998)028<1187:OOSVIT>2.0.CO;2

Schulz, E. W., Josey, S. A., & Verein, R. (2012). First Air-Sea Flux Mooring Measurements in the Southern Ocean. *Geophysical Research Letters*, *39*, L16606. https://doi.org/10.1029/2012GL052290

Schumann, U., & Gerz, T. (1995). Turbulent Mixing in Stably Stratified Shear Flows. *Journal of Applied Meteorology*, *34*(1), 33–48. https://doi.org/10.1175/1520-0450-34.1.33

Shchepetkin, A. F., & McWilliams, J. C. (2005). The Regional Oceanic Modeling System (ROMS): A Split-Explicit, Free-Surface, Topography-Following-Coordinate Oceanic Model. *Ocean Modelling*, *9*(4), 347–404. https://doi.org/10.1016/j.ocemod.2004.08.002

Shchepetkin, A. F., & McWilliams, J. C. (2009). Computational Kernel Algorithms for Fine-Scale, Multiprocess, Longtime Oceanic Simulations. In R. M. Temam, & J. J. Tribbia (Eds.), *Handbook of Numerical Analysis* (Vol. 14, pp. 121–183): Elsevier. https://doi.org/10.1016/S1570-8659(08)01202-0

Shimura, T., Mori, N., Takemi, T., & Mizuta, R. (2017). Long-term Impacts of Ocean Wave-dependent Roughness on Global Climate Systems. *Journal of Geophysical Research: Oceans*, *122*, 1995–2011. https://doi.org/10.1002/2016JC012621

Sinha, N., Tejada-Martínez, A. E., Akan, C., & Grosch, C. E. (2015). Toward a K-Profile Parameterization of Langmuir Turbulence in Shallow Coastal Shelves. *Journal of Physical Oceanography*, *45*(12), 2869–2895. https://doi.org/10.1175/JPO-D-14-0158.1

Skyllingstad, E. D., Smyth, W. D., & Crawford, G. B. (2000). Resonant Wind-Driven Mixing in the Ocean Boundary Layer. *Journal of Physical Oceanography*, *30*(8), 1866–1890. https://doi.org/10.1175/1520-0485(2000)030<1866:RWDMIT>2.0.CO;2

Smyth, W. D., Skyllingstad, E. D., Crawford, G. B., & Wijesekera, H. (2002). Nonlocal Fluxes and Stokes Drift Effects in the K-Profile Parameterization. *Ocean Dynamics*, *52*(3), 104–115. https://doi.org/10.1007/s10236-002-0012-9

Stokes, G. G. (1847). On the Theory of Oscillatory Waves. *Cambridge Philosophical Society*, *8*, 197–237. https://doi.org/10.1017/CBO9780511702242.016

Stommel, H. (1979). Determination of Water Mass Properties of Water Pumped down from the Ekman Layer to the Geostrophic Flow Below. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(7), 3051–3055. https://doi.org/10.1073/pnas.76.7.3051

Sullivan, P. P., McWilliams, J. C., & Melville, W. K. (2007). Surface Gravity Wave Effects in the Oceanic Boundary Layer: Large-Eddy Simulation with Vortex Force and Stochastic Breakers. *Journal of Fluid Mechanics*, *593*, 405–452. https://doi.org/10.1017/S002211200700897X

Sullivan, P. P., Romero, L., McWilliams, J. C., & Melville, W. K. (2012). Transient Evolution of Langmuir Turbulence in Ocean Boundary Layers Driven by Hurricane Winds and Waves. *Journal of Physical Oceanography*, *42*, 1959–1980. https://doi.org/10.1175/JPO-D-12-025.1

Sullivan, P. P., Weil, J. C., Patton, E. G., Jonker, H. J. J., & Mironov, D. V. (2016). Turbulent Winds and Temperature Fronts in Large-Eddy Simulations of the Stable Atmospheric Boundary Layer. *Journal of the Atmospheric Sciences*, *73*(4), 1815–1840. https://doi.org/10.1175/JAS-D-15-0339.1

Sutherland, G., Christensen, K. H., & Ward, B. (2014). Evaluating Langmuir Turbulence Parameterizations in the Ocean Surface Boundary Layer. *Journal of Geophysical Research: Oceans*, *119*, 1899–1910. https://doi.org/10.1002/2013JC009537

Suzuki, N., & Fox-Kemper, B. (2016). Understanding Stokes Forces in the Wave-Averaged Equations. *Journal of Geophysical Research: Oceans*, *121*, 3579–3596. https://doi.org/10.1002/2015JC011566

Suzuki, N., Fox-Kemper, B., Hamlington, P. E., & Van Roekel, L. P. (2016). Surface Waves Affect Frontogenesis. *Journal of Geophysical Research: Oceans*, *121*, 3597–3624. https://doi.org/10.1002/2015JC011563

Takaya, Y., Bidlot, J.-R., Beljaars, A. C. M., & Janssen, P. A. E. M. (2010). Refinements to a Prognostic Scheme of Skin Sea Surface Temperature. *Journal of Geophysical Research*, *115*, C06009. https://doi.org/10.1029/2009JC005985

Teixeira, M. A. C. (2012). The Influence of Langmuir Turbulence on the Scaling for the Dissipation Rate in the Oceanic Boundary Layer. *Journal of Geophysical Research*, *117*, C05015. https://doi.org/10.1029/2011JC007235

Teixeira, M. A. C. (2018). A Model for the Wind-Driven Current in the Wavy Oceanic Surface Layer: Apparent Friction Velocity Reduction and Roughness Length Enhancement. *Journal of Physical Oceanography*, *48*, 2721–2736. https://doi.org/10.1175/JPO-D-18-0086.1

Teixeira, M. A. C., & Belcher, S. E. (2010). On the Structure of Langmuir Turbulence. *Ocean Modelling*, *31*(3–4), 105–119. https://doi.org/10.1016/j.ocemod.2009.10.007

Terray, E. A., Donelan, M. A., Agrawal, Y. C., Drennan, W. M., Kahma, K. K., Williams, A. J., et al. (1996). Estimates of Kinetic Energy Dissipation under Breaking Waves. *Journal of Physical Oceanography*, *26*(5), 792–807. https://doi.org/10.1175/1520-0485(1996)026<0792:EOKEDU>2.0.CO;2

Thompson, A. F., Lazar, A., Buckingham, C., Naveira Garabato, A. C., Damerell, GillianM., & Heywood, K. J. (2016). Open-Ocean Submesoscale Motions: A Full Seasonal Cycle of Mixed Layer Instabilities from Gliders. *Journal of Physical Oceanography*, *46*(4), 1285–1307. https://doi.org/10.1175/JPO-D-15-0170.1

Thorpe, S. A. (2004). Langmuir Circulation. *Annual Review of Fluid Mechanics*, *36*(1), 55–79. https://doi.org/10.1146/annurev.fluid.36.052203.071431

Troen, I. B., & Mahrt, L. (1986). A Simple Model of the Atmospheric Boundary Layer; Sensitivity to Surface Evaporation. *Boundary-Layer Meteorology*, *37*(1-2), 129–148. https://doi.org/10.1007/BF00122760

Tseng, R.-S., & D'Asaro, E. A. (2004). Measurements of Turbulent Vertical Kinetic Energy in the Ocean Mixed Layer from Lagrangian Floats. *Journal of Physical Oceanography*, *34*(9), 1984–1990. https://doi.org/10.1175/1520-0485(2004)034<1984:MOTVKE>2.0.CO;2

Tsujino, H., Urakawa, S., Nakano, H., Justin Small, R., Kim, W. M., Yeager, S. G., et al. (2018). JRA-55 Based Surface Dataset for Driving Ocean - Sea-Ice Models (JRA55-Do). *Ocean Modelling*, *130*, 79–139. https://doi.org/10.1016/j.ocemod.2018.07.002

Uchiyama, Y., McWilliams, J. C., & Shchepetkin, A. F. (2010). Wave-Current Interaction in an Oceanic Circulation Model with a Vortex-Force Formalism: Application to the Surf Zone. *Ocean Modelling*, *34*(1), 16–35. https://doi.org/10.1016/j.ocemod.2010.04.002

Umlauf, L., & Burchard, H. (2005). Second-Order Turbulence Closure Models for Geophysical Boundary Layers. A Review of Recent Work. *Continental Shelf Research*, *25*, 795–827. https://doi.org/10.1016/j.csr.2004.08.004

Umlauf, L., Burchard, H., & Bolding, K. (2014). GOTM Sourcecode and Test Case Documentation.

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

Van Roekel, L., Fox-Kemper, B., Sullivan, P. P., Hamlington, P. E., & Haney, S. R. (2012). The Form and Orientation of Langmuir Cells for Misaligned Winds and Waves. *Journal of Geophysical Research*, *117*, C05001. https://doi.org/10.1029/2011JC007516

van den Bremer, T. S., & Breivik, Ø. (2018). Stokes Drift. *Philosophical Transactions of the Royal Society A*, *376*(2111), 20170104. https://doi.org/10.1098/rsta.2017.0104

Walker, R., Tejada-Martínez, A. E., & Grosch, C. E. (2016). Large-Eddy Simulation of a Coastal Ocean under the Combined Effects of Surface Heat Fluxes Full-Depth Langmuir Circulation. *Journal of Physical Oceanography*, *46*(8), 2411–2436. htpps://doi.org/10.1175/JPO-D-15-0168.1

Wang, D., Kukulka, T., Reichl, B. G., Hara, T., Ginis, I., & Sullivan, P. P. (2018). Interaction of Langmuir Turbulence and Inertial Currents in the Ocean Surface Boundary Layer under Tropical Cyclones. *Journal of Physical Oceanography*, *48*(9), 1921–1940. htpps://doi.org/10.1175/JPO-D-17-0258.1

Warner, J. C., Sherwood, C. R., Arango, H. G., & Signell, R. P. (2005). Performance of Four Turbulence Closure Models Implemented Using a Generic Length Scale Method. *Ocean Modelling*, *8*(1), 81–113. htpps://doi.org/10.1016/j.ocemod.2003.12.003

Webb, A., & Fox-Kemper, B. (2011). Wave Spectral Moments and Stokes Drift Estimation. *Ocean Modelling*, *40*(3-4), 273–288. htpps://doi.org/10.1016/j.ocemod.2011.08.007

Webb, A., & Fox-Kemper, B. (2015). Impacts of Wave Spreading and Multidirectional Waves on Estimating Stokes Drift. *Ocean Modelling*, *96, Part 1*, 49–64. htpps://doi.org/10.1016/j.ocemod.2014.12.007

Yang, D., Chen, B., Chamecki, M., & Meneveau, C. (2015). Oil Plumes and Dispersion in Langmuir, Upper-Ocean Turbulence: Large-Eddy Simulations and K-Profile Parameterization. *Journal of Geophysical Research: Oceans*, *120*, 4729–4759. htpps://doi.org/10.1002/2014JC010542

Zhang, Z., Chini, G. P., Julien, K., & Knobloch, E. (2015). Dynamic patterns in the reduced Craik-Leibovich equations. *Physics of Fluids (1994-present)*, *27*(4), 46605.