

# A Unified Data-Informed Model of Turbulence and Convection for Climate Prediction

Thesis by  
Ignacio Lopez Gomez

In Partial Fulfillment of the Requirements for the  
Degree of  
Doctor of Philosophy

The logo for the California Institute of Technology (Caltech), featuring the word "Caltech" in a bold, orange, sans-serif font.

CALIFORNIA INSTITUTE OF TECHNOLOGY  
Pasadena, California

2023  
Defended October 26, 2022

© 2023

Ignacio Lopez Gomez  
ORCID: 0000-0002-7255-5895

All rights reserved except where otherwise noted

*To the reader*

*Les liquides conduisent très-difficilement la chaleur; mais ils ont, comme les milieux aériformes, la propriété de la transporter rapidement dans certaines directions. C'est cette même propriété qui, se combinant avec la force centrifuge, déplace et mêle toutes les parties de l'atmosphère et celles de l'Océan; elle y entretient des courants réguliers et immenses.*

Jean-Baptiste Joseph Fourier

Mémoire sur les températures du globe terrestre et des espaces planétaires, 1827

## ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor and mentor, Tapio Schneider. His guidance over the past four years has had an impact on my approach to scientific research far beyond what can be inferred from these pages. I am truly grateful for his unconditional support, his encouragement, and his leadership.

I would also like to thank Kamran Mohseni, Ignacio Romero, and Laurent Joly, who prompted me to pursue a PhD and supported my application. Without them, I would not have been able to enter the graduate program at Caltech. I also thank my thesis committee members Jörn Callies, Andrew Stuart, and João Teixeira, for valuable feedback on my research.

At Caltech I learned about the power of collaboration as a science accelerator and a catalyst for critical thought. For that, I thank my collaborators Yair Cohen, Anna Jaruga, Clare Singer, Oliver Dunbar, Costa Christopoulos, Jia He, Haakon Ervik, Charles Kawczynski, Daniel Huang, and Zhaoyi Shen. I am also thankful for stimulating discussions about the magnitude and diversity of human impacts on Earth with Griffin Chure, Rachel Banks, Avi Flamholz, Nicholas Sarai, Mason Kamb, and Rob Phillips; and for the opportunity to work on extreme weather forecasting with Amy McGovern, Shreya Agrawal and Jason Hickey at Google Research.

Half of my time at Caltech was marked by the global COVID-19 epidemic, which had a considerable impact on my personal and academic life. I want to thank Zsófia Török, Tarun Sharma, Bryan Gerber, Sarah Wandelt, Konrad Pilch, Xuan Sun, and Lauren McAlpine for great conversations, trips, and adventures that made it easy to forget about it all. I also want to thank Ronan Le Floch-Best, Lenka Novak and Daniel Martí for exploring the great city of Los Angeles with me, and for their friendship.

Finally, I would like to acknowledge my family and friends from Spain, who supported and encouraged me from a distance, and occasionally came to visit me across the pond. I want to thank my mother, Ana María Gómez, and my father, José Cristóbal López, for always being there for me and for emphasizing the importance of academic education. I also want to thank my grandmother, Mercedes López, and my sister, Beatriz López, for their support over the years. Last but not least, I want to thank Javier Martí, Eduardo Fernández, and Ángel Ortega, for their everlasting

friendship, for some truly amazing trips, and for pushing the boundaries of Spanish lexicon.

## ABSTRACT

Resolving atmospheric turbulent and convective processes in global climate simulations is, and will remain for decades, an intractable computational problem. The strong influence of these processes on cloud formation and maintenance makes the task of modeling turbulence and convection one of the grand challenges in climate modeling, due to the outsized effect of clouds on climate. Current operational climate models fail to represent atmospheric turbulence and convection accurately and consistently across dynamical regimes and vertical levels; errors in the representation of these processes explain about half of the spread in climate projections. This dissertation seeks to reduce such representation errors by improving a recently proposed unified framework for modeling turbulence and convection, known as the extended eddy-diffusivity mass-flux scheme, in several ways. First, the framework is rederived by systematically coarse-graining the governing fluid equations, highlighting the assumptions about atmospheric motion that are necessary to yield the scheme. New terms related to turbulent entrainment processes are shown to arise from the derivation. Second, a generalized formulation of turbulent diffusion consistent with the framework is presented. This novel formulation is shown to accurately represent turbulent processes under statically stable and unstable conditions, including regimes with sharp lapse rate inversions such as the stratocumulus-topped boundary layer. Finally, a methodology to calibrate free parameters within the model from indirect data is proposed. The methodology, based on Kalman filtering, is shown to be efficient at calibrating imperfect black-box models from noisy data, and in its regularized unscented version approximately quantifies parametric uncertainty. The resulting unified data-informed model of turbulence and convection is shown to accurately represent a range of low-cloud regimes that are associated with the largest biases in current operational climate models. The response of the model to realistic climate perturbations is also shown to be consistent with the resolved climate response, although structural errors in the amount of condensate are still important at realistic vertical resolutions.

## PUBLISHED CONTENT AND CONTRIBUTIONS

Lopez-Gomez, Ignacio, Costa Christopoulos, Haakon Ludvig Langeland Ervik, Oliver R. A. Dunbar, Yair Cohen, and Tapio Schneider (2022). “Training physics-based machine-learning parameterizations with gradient-free ensemble Kalman methods”. In: *Journal of Advances in Modeling Earth Systems* 14, e2022MS003105. DOI: 10.1029/2022MS003105.

I.L.G. and T.S. conceived the study. I.L.G. led the research, the development of the associated software tools, analyzed the results, and led the writing of the paper.

Cohen, Yair, Ignacio Lopez-Gomez, Anna Jaruga, Jia He, Colleen M Kaul, and Tapio Schneider (2020). “Unified entrainment and detrainment closures for extended eddy-diffusivity mass-flux schemes”. In: *Journal of Advances in Modeling Earth Systems* 12, e2020MS002162. DOI: 10.1029/2020MS002162.

I.L.G. contributed to the research, the derivation of the equations, the evaluation tools, and the writing of the paper.

Lopez-Gomez, Ignacio, Yair Cohen, Jia He, Anna Jaruga, and Tapio Schneider (2020). “A generalized mixing length closure for eddy-diffusivity mass-flux schemes of turbulence and convection”. In: *Journal of Advances in Modeling Earth Systems* 12, e2020MS002161. DOI: 10.1029/2020MS002161.

I.L.G. and T.S. conceived the study. I.L.G. developed the presented model, analyzed the results, and led the writing of the paper.

## TABLE OF CONTENTS

Acknowledgements . . . . .	v
Abstract . . . . .	vii
Published Content and Contributions . . . . .	viii
Table of Contents . . . . .	viii
List of Illustrations . . . . .	xi
List of Tables . . . . .	xix
Chapter I: Introduction . . . . .	1
Chapter II: The Extended Eddy-Diffusivity Mass-Flux Scheme . . . . .	4
2.1 Modeling the effects of convection . . . . .	4
2.2 Modeling convection explicitly: Mass flux schemes . . . . .	5
2.3 Convection meets turbulence: Eddy-diffusivity mass-flux (EDMF) schemes . . . . .	9
2.4 Derivation of the extended EDMF scheme . . . . .	13
2.5 The full system of equations . . . . .	26
2.6 Similarities between the extended EDMF scheme and other approaches	29
2.7 Discussion . . . . .	34
Chapter III: A Generalized Mixing Length Closure for Eddy-Diffusivity Mass- Flux Schemes of Turbulence and Convection . . . . .	36
3.1 Abstract . . . . .	36
3.2 Introduction . . . . .	37
3.3 EDMF framework . . . . .	38
3.4 Mixing length formulation . . . . .	42
3.5 Results for single-column simulations . . . . .	53
3.6 Summary and discussion . . . . .	65
3.7 Appendix A: Smooth minimum function . . . . .	67
3.8 Appendix B: Subgrid kinetic energy in the extended EDMF scheme .	68
3.9 Appendix C: Probabilistic model for cloud fraction . . . . .	70
3.10 Appendix D: Results with alternative mixing length formulations . .	72
Chapter IV: Training Physics-Based Machine-Learning Parameterizations with Gradient-Free Ensemble Kalman Methods . . . . .	76
4.1 Abstract . . . . .	76
4.2 Introduction . . . . .	77
4.3 Learning about parameterizations as an inverse problem . . . . .	80
4.4 Ensemble Kalman methods . . . . .	88
4.5 Application to an atmospheric subgrid-scale model . . . . .	92
4.6 Discussion and conclusions . . . . .	104
4.7 Appendix A: Configuration-based principal component analysis . . .	107
4.8 Appendix B: Addressing model failures with unscented Kalman in- version . . . . .	108

4.9 Appendix C: Parameter transformation and prior . . . . .	109
4.10 Appendix D: Calibration using very noisy loss evaluations . . . . .	109
4.11 List of Symbols . . . . .	110
Chapter V: Skillful Representation of Low Cloud Regimes with the Extended Eddy-Diffusivity Mass-Flux Scheme . . . . .	113
5.1 Abstract . . . . .	113
5.2 Introduction . . . . .	114
5.3 Data and methods . . . . .	116
5.4 Dynamics of low cloud regimes: LES vs. EDMF . . . . .	121
5.5 Response to a uniform 4 K warming: LES vs. EDMF . . . . .	127
5.6 Discussion . . . . .	131
5.7 Appendix A: Stretched grid definition . . . . .	133
5.8 Appendix B: Results for miscalibrated models . . . . .	134
Chapter VI: Conclusions . . . . .	137
Bibliography . . . . .	139

## LIST OF ILLUSTRATIONS

<i>Number</i>	<i>Page</i>
2.1 Schematic of the mass flux scheme decomposition of Arakawa and Schubert (1974) within an atmospheric column, from the surface to some height below the highest cloud top. A spectrum of ascending saturated thermals ( $\Omega_1, \dots, \Omega_n$ ) occupying a cross-sectional area fraction $a_u$ (grey shading) is initialized atop a well-mixed turbulent boundary layer (height $z_{bl}$ ). Their upward mass flux (e.g., $M_1$ , thick solid arrows) is compensated by large-scale environmental subsidence ( $M_0$ , thin descending arrows). The spectrum of thermals aims to represent different clouds; the top height of each cloud depends on their mass flux at $z_{bl}$ and the net entrainment rate (blue arrows). Detrainment is concentrated at cloud top for each thermal (orange arrows). . . . .	8
2.2 Schematic of the EDMF scheme decomposition within an atmospheric column, from the surface to some height below the highest cloud top. Updrafts are initialized at the surface, and interact with the environment through entrainment (blue arrows) and detrainment fluxes (orange arrows). Updrafts are unsaturated below their lifting condensation level, which may be different for different plumes, and become saturated above. Turbulent mixing induces further vertical transport in the environment at all heights (green arrows). . . . .	12
2.3 Schematic of the extended EDMF scheme decomposition within an atmospheric column, from the surface to some height below the highest cloud top. Arrows as in Figure 2.2. The thermodynamic state of the updrafts at the surface, along with entrainment and detrainment processes, determine their area fraction, mass flux, and termination height. Transient terms allow mature convective clouds to dissipate over a finite amount of time. . . . .	21

- 3.1 Profiles of (a) potential temperature and (b) horizontal velocity averaged over the 9th hour of the GABLS simulation. Results are shown for LES and for the EDMF-based SCM with  $\Delta z = 3.125$  m,  $\Delta z = 12.5$  m, and  $\Delta z = 50$  m. The shaded region represents the spread of LES results with  $\Delta z=3.125$  m reported in Beare et al. (2006). 56
- 3.2 (a) Contours of eddy diffusivity  $K_m$  ( $\text{m}^2 \text{s}^{-1}$ ) as a function of time and height for the GABLS simulation using the SCM with  $\Delta z = 12.5$  m. Colors show the dominant (minimum) mixing length. (b) Profiles of eddy diffusivity averaged over the 9th hour. LES diffusivity is diagnosed from the shear production term  $\mathcal{S}$  and the grid-mean shear. (c) Time series of vertically integrated TKE  $E_{int}$ . (d) Profiles of  $u_*$ -normalized TKE averaged over the 9th hour. In (d),  $u_*$  is the average friction velocity during the 9th hour, with  $u_* = 0.25 \text{ m s}^{-1}$  for LES and  $u_* = 0.26, 0.26$  and  $0.30 \text{ m s}^{-1}$  for the SCM cases, in order of increasing  $\Delta z$ . In (b), (c), and (d), results are shown for LES (solid line), EDMF with  $\Delta z = 3.125$  m (dashed line),  $\Delta z = 12.5$  m (dash-dotted line), and  $\Delta z = 50$  m (dotted line). The shaded region represents the spread of LES results with  $\Delta z=3.125$  m reported in Beare et al. (2006). . . . . 57
- 3.3 Profiles of TKE budget terms averaged over the 9th hour of GABLS simulation. Profiles shown in (a) are for shear production  $\mathcal{S}$ , dissipation  $\mathcal{D}$  and the pressure term  $\mathcal{P}$ . Shown in (b) are the buoyant production term  $\mathcal{B}$  and turbulent transport  $\mathcal{T}$ . (c) Time evolution of the vertically integrated TKE ( $E_{int}$ ) production and dissipation terms. The rate of change of  $E_{int}$  in LES is included as a reference. Results are shown for LES (solid line) and for EDMF with  $\Delta z = 3.125$  m (dashed line),  $\Delta z = 12.5$  m (dash-dotted line), and  $\Delta z = 50$  m (dotted line). The shaded region represents the spread of LES results with  $\Delta z=3.125$  m reported in Beare et al. (2006). . . . . 58

- 3.4 Profiles of (a) liquid water potential temperature, (b) total water specific humidity, (c) liquid water specific humidity, (d) vertical buoyancy flux, (e) vertical transport of  $q_t$  and (f) vertical transport of  $\theta_l$ . Profiles averaged over the 4th hour of the DYCOMS-II RF01 simulation. In (e) and (f), the eddy diffusivity ( $SCM_{ED}$ ) and mass flux ( $SCM_{MF}$ ) components of the vertical flux are shown (plotting conventions follow the legend in panel (a)). The shaded region represents the spread of LES results reported in Stevens et al. (2005). Observations are also reported in Stevens et al. (2005). . . . . 62
- 3.5 (a) Contours of eddy diffusivity  $K_m$  ( $m^2 s^{-1}$ ) as a function of time and height for the DYCOMS-II RF01 simulation using the SCM with  $\Delta z = 5$  m. Colors show the dominant (minimum) mixing length. (b) Time series of vertically integrated TKE  $E_{int}$ . (c) Time series of liquid water path (LWP). In (b) and (c), results are shown for LES (solid line), EDMF with  $\Delta z = 5$  m (dashed line),  $\Delta z = 20$  m (dash-dotted line),  $z = 50$  m (dotted line) and  $z = 75$  m ( $\times$ ). . . . . 63
- 3.6 Profiles of (a) normalized turbulence kinetic energy and (b) eddy diffusivity. Profiles averaged over the 4th hour of the DYCOMS-II RF01 simulation. In (a),  $u_*$  is the average friction velocity during the 4th hour, with  $u_* = 0.22$   $m s^{-1}$  for LES and  $u_* = 0.26, 0.25, 0.24,$  and  $0.24$   $m s^{-1}$  for the SCM cases, in order of increasing  $\Delta z$ . In (b), the eddy diffusivity diagnosed from LES follows (3.45). The shaded region represents the spread of LES results reported in Stevens et al. (2005). Plotting conventions follow the legend in panel (a). . . . . 64
- 3.7 Profiles of (a) potential temperature, (b) vertical buoyancy flux, and (c) vertical flux of potential temperature. Profiles averaged over the 5th hour of dry convection simulation. In (c), the eddy diffusivity ( $SCM_{ED}$ ) and mass flux ( $SCM_{MF}$ ) components of the vertical flux are shown separately. (d) Time evolution of the boundary layer depth  $z_{bl}$ , computed as the level of minimum buoyancy flux  $\langle w^* b^* \rangle$ . Plotting conventions follow the legend in panel (a). . . . . 65

3.8	Profiles of (a) normalized turbulence kinetic energy and (b) eddy diffusivity for the dry convective boundary layer. Profiles averaged over the 5th hour of simulation. In (a), $u_*$ is the average friction velocity during the 5th hour, with $u_* = 0.18 \text{ m s}^{-1}$ for LES and $u_* = 0.24 \text{ m s}^{-1}$ for all SCM cases. In (b), the eddy diffusivity diagnosed from LES follows (3.45). Plotting conventions follow the legend in panel (a). . . . .	66
3.9	Schematic of subgrid kinetic energy reservoirs and pathways in the extended EDMF scheme under the net dissipation closure. Notation follows equation (3.56). Dashed lines represent energy pathways that result in implicit grid-mean TKE dissipation under the net dissipation closure (3.13). Summation over $i = 0, \dots, n$ is implied. . . . .	70
3.10	Profiles of (a) potential temperature, (b-c) horizontal velocity, (d) buoyancy flux, (e) normalized TKE, and (f) eddy diffusivity for the 9th hour of the GABLS simulation. Results shown for LES and for the SCM ( $\Delta z = 50 \text{ m}$ ) with alternative mixing length formulations: original (OG), no $l_w$ (NLW), and no $l_b$ (NLB). Shaded region as in Figure 3.1. . . . .	73
3.11	Profiles of (a) liquid water potential temperature, (b) total water specific humidity, (c) liquid water specific humidity, (d) buoyancy flux, (e) normalized TKE and (f) eddy diffusivity for the 4th hour of the DYCOMS-II RF01 simulation. The time evolution of the liquid water path is shown in (g). Results shown for LES and for the SCM ( $\Delta z = 50 \text{ m}$ ) using alternative mixing length formulations: original (OG), no $l_w$ (NLW), no $l_b$ (NLB) and no entrainment term $\mathcal{I}$ in equation (3.41) (NED). Shaded region as in Figure 3.4. . . . .	74
3.12	Profiles of (a) potential temperature, (b) buoyancy flux, (c) normalized TKE, and (d) eddy diffusivity for the 5th hour of the dry convection simulation. Results shown for LES and for the SCM ( $\Delta z = 150 \text{ m}$ ) with alternative mixing length formulations. . . . .	75

- 4.1 Schematic of the strategy used to construct a regularized inverse problem from observations of a dynamical system  $\zeta$ . The two branches represent different configurations of the dynamical system. From left to right: (a) the observed state is obtained following Section 4.3 or from any observable fields for each configuration  $c$ ; (b) the observed state is normalized; (c) mean and covariance of the normalized state are computed; (d)  $\tilde{y}_c$  and  $\tilde{\Gamma}_c$  are projected onto a lower dimension and regularized; (e) the statistical summaries of each configuration are aggregated, defining the global inverse problem (4.3)–(4.5). . . . 87
- 4.2 Batch (a) training and (b) validation MSE as defined in equation (4.34). Lines represent the error of the ensemble mean  $\bar{\theta}$ ,  $\text{MSE}(\bar{\theta}; \tilde{y}_B)$ , and the shading represents the ensemble standard deviation of  $\text{MSE}(\theta; \tilde{y}_B)$  around the optimal point estimate  $\bar{\theta}$ . All errors are normalized with respect to the largest initial  $\text{MSE}_v(\bar{\theta}; \tilde{y}_B)$ , so they can be compared. Results are shown for EKI and UKI, using  $J = 2p + 1$  and training batch sizes  $|B| = 5, 20$ . Errors for  $|B| = 5$  are averaged using a rolling mean of 20 configurations to enable comparison with  $|B| = 20$ . In (b), the inset focuses on the validation error evolution for a longer training period. . . . . 96
- 4.3 Parameter evolution of the turbulent dissipation (a), updraft advection damping (b), and updraft surface area fraction (c). All values are given in physical space. The solid lines describe the trajectories of the mean estimate,  $\mathcal{T}^{-1}(m_n)$ . For UKI, the marginal  $\pm\sigma$  uncertainty band is included in shading. This uncertainty is equal to  $\pm\mathcal{T}^{-1}(\sqrt{(\Sigma_n)_{i,i}})$  for the  $i$ -th parameter. The black dashed lines are the  $\pm\sigma$  uncertainty bands of the prior used for regularization. Legend as in Figure 4.2. . . . 98
- 4.4 Parameter correlations estimated from UKI using  $|B| = 20$  (a), and evolution of the total parameter variance from UKI using  $|B| = 20, 10$  and  $5$ , normalized by the prior variance  $\text{tr}(\Lambda) = 16$  (b). Note that the initial covariance estimate used in UKI (with  $\text{tr}(\Sigma_0) = 1$ ) is decoupled from the prior. Symbols follow Table 4.1. . . . . 98

- 4.5 Prior, posterior and LES profiles of liquid water specific humidity ( $\bar{q}_l$ ), subgrid-scale moisture flux ( $\overline{w'q'_l}$ ) and zonal velocity ( $\bar{u}$ ) for cfSites 5 (top) and 14 (bottom) using July forcing from the AMIP4K experiment as in Shen et al. (2022). The gray shading represents the internal variability of the LES simulations over 6 days of steady forcing, and the full lines represent 3-hour time-averaged profiles. EKI prior and posterior results are point estimates evaluated at the parameter vector closest to the ensemble mean. The UKI posterior shading spans the central 68% of the profile posterior distribution. All Kalman methods used  $|B| = 5$  and  $J = 2p + 1$ . . . . . 99
- 4.6 Evolution of the validation error (a) and estimates of the turbulent dissipation (b) and entrainment coefficient (c) for calibration processes using observations of the state (4.33) at 50 m resolution (UKI<sub>f</sub>, EKI<sub>f</sub>), or from  $\bar{\theta}_l$ ,  $\bar{q}_t$  and  $\bar{q}_l$  at 200 m resolution (UKI<sub>o</sub>, EKI<sub>o</sub>). All inversion processes use  $|B| = 20$ . Shading is defined as in Figures 4.2 and 4.3. . . . . 100
- 4.7 Prior, posterior and LES profiles of liquid water specific humidity ( $\bar{q}_l$ ), vertical moisture flux ( $\overline{w'q'_l}$ ) and zonal velocity ( $\bar{u}$ ) for cfSite 3 using July forcing (top) and cfSite 14 using January forcing (bottom) from the AMIP4K experiment (Shen et al., 2022). Posterior results are shown for a model calibrated using the high-resolution state (4.33) (Full Obs.), and coarse-resolution observations of  $\bar{\theta}_l$ ,  $\bar{q}_t$  and  $\bar{q}_l$  (Partial Obs.). Shadings and legend as in Figure 4.5. Results obtained using UKI with  $|B| = 20$ . . . . . 101
- 4.8 Batch (a) training and (b) validation normalized MSE for the hybrid (EDMF+NN) and empirical (EDMF) models. Lines, shading and inset as in Figure 4.2. Results are shown for calibration with EKI, using  $J = 50, 100$  and  $2p + 1 = 159$  ensemble members for the hybrid model. The empirical model training uses  $J = 2p + 1 = 33$ . All inversion processes use batch size  $|B| = 10$ . . . . . 104

4.9	Prior, posterior and LES profiles of liquid water specific humidity ( $\bar{q}_l$ ), total water specific humidity ( $\bar{q}_t$ ) and vertical moisture flux ( $\overline{w'q'_t}$ ) for cfSite 14 using July forcing (top) and cfSite 8 using January forcing (bottom) from the AMIP4K experiment (Shen et al., 2022). Definitions of prior, posterior and shading as in Figure 4.5. Posterior results are shown for the EDMF model with empirical closures (EDMF), and with the neural network entrainment closure (4.35) (EDMF+NN), using early stopping and 25 epochs of training. Results obtained using EKI with $ B  = 10$ . . . . .	105
4.10	(a) Scatter plot of covariance eigenvalues $\mu_{c,i}$ and the number of zero-crossings of their corresponding eigenmode for three different configurations of an LES solver. (b) The same plot, with eigenvalues normalized by the leading eigenvalue of each configuration ( $\mu_{c,1}$ ). Trailing eigenvalues are associated with high-wavenumber oscillatory modes with frequent sign changes. . . . .	108
4.11	Evolution of the validation error (a) and estimates of the entrainment (b), and eddy diffusivity (c) coefficients. Results shown for UKI using batch sizes of 10 and 1, and EKI using a batch size of 1. Parameter uncertainty only shown for UKI <sub>10</sub> and UKI <sub>1</sub> , $\Delta t = 1/60$ for clarity. All results shown use $\Delta t =  C / B $ unless otherwise specified. Shading as in Figures 4.2 and 4.3. . . . .	110
5.1	Location of cfSites considered in this study (black dots), and average observed low-level cloud fraction over the ocean during the period 2007-2012. Low-level cloud fraction is derived from CALIPSO-GOCCP for clouds below 680 hPa (Chepfer et al., 2010). . . . .	117
5.2	Bin-averaged values (left) and biases per forcing model (right) of the vertically integrated subgrid-scale fluxes of (a, b) moisture and (c, d) entropy, (e, f) cloud-top height, and (g, h) cloud thickness. Results for all simulations in the library are shown (dots), binned in $30 \text{ W m}^{-2}$ intervals of surface latent heat flux (LHF). Error bars define the central 68% of the bin distribution, using the median-unbiased estimator of Hyndman and Fan (1996). Shading shows the bootstrapped standard deviation of the bin-averaged bias. . . . .	122
5.3	Same as Figure 5.2, but for the ISO-EDMF configuration. . . . .	124

5.4	Bin-averaged values (left) and biases per forcing model (right) of (a, b) liquid water path and (c, d) average cloud fraction. Individual simulations (dots) are binned in decoupling parameter intervals with bounds $[-1.15, 0, 1, 2, 3, 4]$ . Bin-averaged values from the forcing climate models are also shown (GCM). Error bars and shading as in Figure 5.2. . . . .	125
5.5	Same as Figure 5.4, but for the ISO-EDMF configuration. . . . .	126
5.6	Bin-averaged biases in (a) liquid water path and (b) average cloud fraction, computed over the entire library. Results shown for the forcing climate models (GCM), and the extended EDMF scheme in ESM and ISO configurations. Shading as in Figure 5.4. . . . .	127
5.7	Vertical profiles of (a,d) liquid water potential temperature, (b,e) total water specific humidity, and (c,f) liquid water specific humidity from LES, the EDMF scheme at ESM resolution, and HadGEM2-A. Results shown for cfSites 17 (top) and 23 (bottom) of the AMIP4K experiment run with HadGEM2-A and with July forcing. . . . .	128
5.8	Cloud fraction and decoupling phase space of low cloud equilibrium states simulated by (a) LES, (b) ESM-EDMF and (c) ISO-EDMF. Arrows connect the AMIP and AMIP4K states of each cfSite and season. Simulations forced by all three climate models in the library are shown. . . . .	129
5.9	Scatter plot of AMIP4K-AMIP differences in KPI values as simulated by LES (x-axis), and ESM-EDMF (y-axis). The legend of each panel shows the coefficient of determination of the EDMF change in terms of the LES change; the linear regression model is shown as a blue line. Dashed lines show the zero-line and the identity. . . . .	130
5.10	Same as Figure 5.9, but for the ISO-EDMF configuration. . . . .	131
5.11	Same as Figure 5.2, but for a miscalibrated extended EDMF scheme with a random feature closure for entrainment fluxes, at ESM resolution. . . . .	135
5.12	Same as Figure 5.9, but for a miscalibrated extended EDMF scheme at ESM resolution. . . . .	136

## LIST OF TABLES

<i>Number</i>	<i>Page</i>
3.1 Parameters in the mixing length closure and values used in this study.	53
4.1 Parameters $\phi$ considered for calibration in this study. The prior mean values are taken from LG2020 (Lopez-Gomez et al., 2020), C2020 (Cohen et al., 2020) and H2021 (J. He et al., 2021), where a physical description of the parameters may be found. . . . .	94
4.2 List of all symbols used in the chapter, their dimensions when applicable, and a short description. . . . .	111
4.3 (cont.) List of all symbols used in the chapter. . . . .	112
5.1 Symbol and description of key performance indicators (KPIs) of the extended EDMF scheme in the stratocumulus-to-cumulus transition. Each KPI is associated with a predictor that explains a large fraction of its variance across regimes. LHF stands for latent heat flux, and $D$ is the decoupling parameter (Bretherton and Wyant, 1997). All KPIs computed over the lower 4 km of atmosphere. . . . .	121

*Chapter 1*

## INTRODUCTION

Climate models are the best tool at our disposal to analyze the possible trajectories of the climate system given a specific anthropogenic forcing. Roughly 60 years since their inception, they have evolved from models of a dry<sup>1</sup> atmosphere over a shallow quiescent ocean, to Earth system models that simulate aerosols (Kim et al., 2008), mesoscale ocean eddies (Adcroft et al., 2019), and even plant stomatal physiology (Bonan et al., 2014). Granted, their use as harbingers of anthropogenic climate change has been rendered obsolete by humanity's relentless effort to make the evidence of global warming inescapable (H.-M. Zhang et al., 2019; Chure et al., 2022). Nevertheless, they remain our only means to quantitatively compare future climate adaptation, mitigation and geoengineering strategies. For this reason, the need for accurate climate projections with quantified uncertainties is more pressing than ever (Goddard et al., 2009; Hope, 2015). In order to improve climate projections, we need to understand their strengths, as well as where their uncertainties stem from.

Our understanding of the greenhouse effect, the fundamental process driving global warming, is robust and has been for over a century. Fourier (1827) first noted that higher atmospheric absorptivity in the infrared than in the visible range resulted in warmer surface temperatures than those expected in a planet without an atmosphere. Following Fourier's work, Pouillet (1838) presented the first idealized analytical model of the greenhouse effect, showing that the surface would become warmer if the infrared absorptivity of the atmosphere were to increase. More accurate estimates of the greenhouse effect would become possible in the coming decades through the work of Kirchhoff (1860), Tyndall (1865), Stefan (1879), Boltzmann (1884), and Planck (1901), on thermal radiation.

The now routine analysis of the equilibrium climate sensitivity (ECS), or forecast of the mean surface temperature increase in equilibrium with a doubling of atmospheric CO<sub>2</sub>, was first considered by Arrhenius (1896). His estimate assumed global radiative equilibrium between the Earth's surface, the atmosphere and its outermost layer. Notably, he considered the most significant of climate feedbacks: the strengthening of the greenhouse effect due to an increase in atmospheric water

---

<sup>1</sup>Without a hydrologic cycle (Smagorinsky et al., 1965).

vapor, governed by the Clausius-Clapeyron relation. Little progress was made in the next 60 years, with many analyses of the ECS ignoring the water vapor feedback that Arrhenius recognized (Callendar, 1938; Plass, 1956), or even questioning the sign of the ECS (Möller, 1963).

Climate change research gained significant momentum in the 1960s through the seminal work of Syukuro Manabe, enabled by the creation of the first general circulation model at the Geophysical Fluid Dynamics Laboratory (Smagorinsky et al., 1965). The Nobel laureate recognized the important role of convection (Manabe and Strickler, 1964), reintroduced the water vapor feedback (Manabe and Wetherald, 1967), estimated the high-latitude albedo feedback, and studied the changes in meridional transport and the hydrologic cycle stemming from a change in CO<sub>2</sub> concentration (Manabe and Wetherald, 1975). The largest sources of error in his analysis of the ECS were, to a large extent, the same problems that haunt climate models today: an imperfect treatment of fluid transport, a failure to predict the response of clouds, and a lack of data-based constraints on the physical process parameterizations within his model.

This thesis seeks to address several necessarily focused aspects of these problems that remain relevant to climate models today. The first one is the need for an accurate and physically consistent parameterization of unresolved vertical fluid transport in the troposphere; a need that arises from the scale separation between computationally affordable resolutions and the eddies and updrafts that control the fate of clouds on Earth (Schneider et al., 2017a). Uncertainties about vertical mixing in the troposphere alone account for about half of the variance in ECS forecasts by current climate models (Sherwood et al., 2014).

In this context, Chapter 2 reviews the development of a unified modeling framework for atmospheric fluid flow known as the eddy-diffusivity mass-flux (EDMF) scheme (Siebesma and Teixeira, 2000; Soares et al., 2004; Siebesma et al., 2007), outlining its scientific background, its underlying assumptions, and presenting a self-contained derivation of a generalized time-dependent version of the scheme from the Navier-Stokes equations. A discussion of similarities and differences between current implementations of the EDMF framework in the literature is also offered. The use of this scheme reduces the unresolved turbulence and convection closure problem to finding parameterizations for a set of elementary fluid processes, like small-scale mixing and entrainment by ascending plumes. Chapter 3 then derives a generalized parameterization of small-scale turbulent mixing in the atmosphere,

based on a quasi-balanced budget of turbulence kinetic energy and universal physical constraints that arise naturally in stratified boundary layer flows. The reader will not find references in the derivation of this closure to properties of emergent atmospheric flow patterns like cumulus or stratocumulus clouds that, although critical for Earth's climate, have in my opinion become an obstacle to unification of turbulence and convection models in recent years.

The second problem covered in this thesis is the lack of efficient and flexible tools to calibrate and quantify the uncertainty of climate process parameterizations from data. Although training algorithms exist in cases where the model output is differentiable with respect to model parameters, this represents a strong constraint on parameterization architectures and the data used for training them (Bretherton et al., 2022; Lopez-Gomez et al., 2022b). As a solution to this problem, Chapter 4 presents a methodology to systematically calibrate physical process parameterizations with unknown error structures from possibly noisy and indirect data. The solutions proposed to solve this problem draw heavily on the growing body of work on ensemble Kalman processes for inverse problems (Iglesias et al., 2013; Schillings and Stuart, 2017; D. Z. Huang et al., 2022a; D. Z. Huang et al., 2022b). This methodology is then applied to train physical parameterizations within the EDMF scheme derived in Chapters 2 and 3, using data from large-eddy simulations of cloud and boundary layer flows typical of the stratocumulus-to-cumulus transition regions in the eastern Pacific Ocean (Teixeira et al., 2011; Shen et al., 2022).

Finally, Chapter 5 is devoted to the physical analysis of the dynamics and climate change response of a data-informed EDMF scheme trained on a wide range of current-climate atmospheric conditions, spanning the stratocumulus to shallow cumulus transition found off the coasts of California and Peru. The response of the parameterized dynamics is compared, on a single-column basis, to the response of large-eddy simulations with perfectly-resolved turbulent and convective dynamics. This analysis is necessarily incomplete due to the focus on a single atmospheric column, which neglects interactions with mesoscale flow patterns and the general circulation of the atmosphere. Nevertheless, it enables a controlled study of the structural errors within the scheme, its generalization to unseen large-scale forcing conditions, and the relation between the changes predicted by the model and those resolved by large-eddy simulations under global warming. An assessment of the performance of the parameterized dynamics within a global Earth System Model, beyond the scope of this thesis, is left for future work.

## Chapter 2

### THE EXTENDED EDDY-DIFFUSIVITY MASS-FLUX SCHEME

#### 2.1 Modeling the effects of convection

In his 1827 *Mémoire* on the temperature of the Earth, and before discussing the greenhouse effect induced by infrared atmospheric absorbers, Fourier highlighted the important role of fluid transport in setting the temperature distribution of the planet. He remarked that the presence of a gaseous atmosphere tends to homogenize temperatures through buoyancy-driven transport, or convection (Fourier, 1827). Because of convection, he wrote, the actual temperature lapse rate of the lower atmosphere must be lower than the one resulting from radiative balance. The modification of Earth's equilibrium climate by convection remained unaddressed beyond similar qualitative statements for the next century; most climate sensitivity studies routinely chose to ignore atmospheric motion for analysis tractability (e.g., Plass, 1956).

The first attempt to quantitatively assess the effect of convection on the temperature distribution of the atmosphere is due to Manabe and Strickler (1964) and Manabe et al. (1965). Manabe and Strickler (1964) proposed to model convection as a stabilizing mechanism that completely balances the destabilizing effect of upwelling longwave radiation through an atmosphere. In their model, the atmosphere simply undergoes instantaneous convective adjustments that allow its lapse rate to remain below a critical value of 6.5 K/km, chosen to match the observed average tropospheric lapse rate. The introduction of this simple radiative-convective equilibrium proved crucial to reproduce realistic climatologies in computational models. Earth's atmosphere in pure radiative equilibrium would have surface temperatures exceeding 50 °C in the subtropics and tropopause temperatures below -80 °C, far from the observed values (Manabe and Strickler, 1964). In contrast, imposing radiative-convective equilibrium leads to subtropical surface temperatures of  $\sim 27$  °C, and more realistic temperatures in the upper troposphere.

Nevertheless, several unphysical phenomena were soon identified in models using this simple convection scheme. It was shown that instantaneous convective adjustment leads to large precipitation rates, and to saturation of the atmosphere after convection is triggered, which is uncommon in nature (Emanuel and Raymond,

1993; Villalba-Pradas and Tapiador, 2022). It also results in cold biases in the tropical upper troposphere due to the absence of penetrative convection into conditionally stable atmospheric layers (G. J. Zhang and McFarlane, 1995). The convective adjustment approach would be revisited a few decades later by Betts (1986) and Betts and Miller (1986), who proposed a lagged convective adjustment model with two parameters dictating the adjustment timescales for the temperature and specific humidity lapse rates, respectively. In this model, convective fluxes are the result of finite-time relaxation processes toward reference thermodynamic profiles specified a priori. Due to its simplicity, the Betts-Miller scheme is still a popular choice in low-complexity atmospheric models (Beucler et al., 2018; Dunbar et al., 2021). A modified version of the scheme introduced by Janjić (1994), in which the relaxation timescales and profiles are a function of the thermodynamic state of the convecting clouds, is still used in some weather models today (Skamarock et al., 2019; J. Li et al., 2018).

Convective adjustment schemes model convection implicitly, in terms of a desired outcome: a realistic thermodynamic structure of Earth's atmosphere that is specified a priori. They do not attempt to model clouds or convective motion explicitly, nor their relation to the environment<sup>1</sup>. Their simplicity limits their usefulness if one is interested in understanding how convection and clouds may change with increasing levels of CO<sub>2</sub>, or how these changes feed back on the ECS.

## 2.2 Modeling convection explicitly: Mass flux schemes

Early attempts to model convection explicitly include the mass flux schemes of Ooyama (1971), Yanai et al. (1973), and Arakawa and Schubert (1974). These schemes aim to represent *moist* convection as the aggregate effect of an ensemble of slender saturated updrafts carrying buoyant moist air to the upper layers of the troposphere; exchanging mass, moisture and energy with their environment along the way. This decomposition of the convecting atmosphere into updrafts and a surrounding environment was proposed based on the observed bimodal structure of the joint probability density function of vertical velocity and moist conserved variables within convective clouds (Yanai et al., 1973; Arakawa and Schubert, 1974; Siebesma et al., 2007). Moist convection refers to convective processes that result in phase changes of water within the rising air parcels, which lead to latent

---

<sup>1</sup>In fact, according to Betts (1986), "it is clearly impossible to attempt to integrate at each grid-point in a global model, a cloud-scale model of much realism." Hopefully the reader is convinced this is no longer the case by the end of this thesis, at least for cloud ensembles.

heat release and further increase the vertical transport of energy and moisture to upper layers of the troposphere. Dry convection, which does not result in latent heat release, is represented in these schemes through a separate boundary layer parameterization (Arakawa and Schubert, 1974).

Analytically, mass flux schemes are derived by considering a statistical partition of the atmosphere above the boundary layer into a set of saturated thermals, and an unsaturated environment in which the thermals are embedded. From this decomposition arises a parameterization of moist convection in the equations of motion. To see this, consider the evolution equation for a thermodynamic field  $\phi$  averaged over an atmospheric volume  $\Omega$ ,

$$\frac{\partial \langle \phi \rangle}{\partial t} + \langle \mathbf{u}_h \rangle \cdot \nabla_h \langle \phi \rangle + \langle w \rangle \frac{\partial \langle \phi \rangle}{\partial z} = -\frac{1}{\rho} \left[ \nabla_h \cdot (\rho \langle \mathbf{u}_h^* \phi^* \rangle) + \frac{\partial (\rho \langle w^* \phi^* \rangle)}{\partial z} \right] + S_{\langle \phi \rangle}, \quad (2.1)$$

where  $\langle \cdot \rangle$  denotes a density-weighted average over volume  $\Omega$ ,  $(\cdot)^*$  denotes deviations about this average,  $w$  is the vertical velocity,  $\mathbf{u}_h$  is the horizontal velocity vector,  $\rho$  is the volume-averaged density, and  $S_{\langle \phi \rangle}$  represents all sources and sinks of  $\langle \phi \rangle$ . The term  $\partial(\rho \langle w^* \phi^* \rangle)/\partial z$  encodes all unresolved vertical transport processes within the atmospheric volume  $\Omega$ , which for typical climate model resolutions include the effects of turbulence and convection in the troposphere (Schneider et al., 2017b).

Mass flux schemes decompose the volume  $\Omega$  into  $n + 1$  subdomains  $\Omega_0, \dots, \Omega_n$ . Here,  $\Omega_0$  denotes the environment, which occupies a large fraction  $a_0 \approx 1$  of the volume of  $\Omega$  (Bjerknes, 1938; Ogura and Cho, 1973). The other subdomains  $\Omega_i$  represent thermals or updrafts. Some mass flux models retain a single bulk updraft ( $n = 1$ , Yanai et al., 1973; Tiedtke, 1989; Gregory and Rowntree, 1990), while others consider the effect of an ensemble of thermals ( $n > 1$ , Arakawa and Schubert, 1974; G. J. Zhang and McFarlane, 1995; Grell and Dévényi, 2002), each of them occupying an area fraction<sup>2</sup>  $a_i$ . In any case, all schemes satisfy the volume partition

$$\sum_{i=1}^n a_i + a_0 = a_u + a_0 = 1, \quad (2.2)$$

where  $n \geq 1$  is the number of thermals or updrafts considered. The volume decomposition (2.2) also implies the Favre average (Favre, 1965)

$$\rho \langle \phi \rangle = \sum_{i=0}^n \rho a_i \bar{\phi}_i, \quad (2.3)$$

---

<sup>2</sup>Many studies envision this partition in the limit  $\Delta z \rightarrow 0$ , which is why  $a_i$  is denoted area fraction in the literature. In practice, all schemes are implemented in grids with a finite  $\Delta z$  due to limited computational power.

where  $\overline{(\cdot)}_i$  denotes the density-weighted average over the subdomain  $\Omega_i$ . Several approximations are added to the updraft/environment decomposition in classic mass flux schemes for tractability,

- Fields within each subdomain (e.g.,  $\bar{\phi}_i$ ) are approximated as uniform, so turbulent fluxes or covariances within the subdomains are neglected. Then, the grid-averaged covariance can be written as

$$\rho \langle w^* \phi^* \rangle \approx \sum_{i=0}^n \rho a_i (\bar{w}_i - \langle w \rangle) (\bar{\phi}_i - \langle \phi \rangle); \quad (2.4)$$

- Updrafts are assumed to be in dynamic equilibrium, such that transient terms like  $\partial_t a_i$ ,  $\partial_t \bar{\phi}_i$  or  $\partial_t \bar{w}_i$  are neglected;
- Environmental averages roughly match the global average, since the updraft area fraction is assumed negligible ( $a_u \ll 1$ );
- Unresolved horizontal fluxes in equation (2.1) are much smaller than the vertical fluxes, and so they are neglected;
- Density variations between updrafts and environment are neglected, except in buoyancy terms. This is implied in the Favre average (2.3), where the density  $\rho$  is taken to be constant across subdomains.

The usefulness of the mass flux framework becomes apparent when applying the decomposition (2.4) to equation (2.1). Using the approximations listed above, this decomposition implies

$$\frac{\partial \langle \phi \rangle}{\partial t} + \langle \mathbf{u}_h \rangle \cdot \nabla \langle \phi \rangle + \langle w \rangle \frac{\partial \langle \phi \rangle}{\partial z} = -\frac{1}{\rho} \frac{\partial}{\partial z} \left[ \sum_{i=1}^n M_i (\bar{\phi}_i - \langle \phi \rangle) \right] + S_{\langle \phi \rangle}, \quad (2.5)$$

where we have neglected the mass flux contribution from the environment, consistent with the approximation  $\bar{\phi}_0 \sim \langle \phi \rangle$ . In equation (2.5),  $M_i$  is the mass flux due to updraft  $i$ ,

$$M_i = \rho a_i (\bar{w}_i - \langle w \rangle), \quad \sum_{i=1}^n M_i = -M_0. \quad (2.6)$$

The mass flux  $M_i$ , multiplied by the subdomain anomaly  $\bar{\phi}_i - \langle \phi \rangle$ , represents the vertical convective transport of  $\phi$  due to updraft  $i$ . The magnitude of this transport is regulated in mass flux schemes through entrainment of unsaturated environmental air and detrainment of saturated air,

$$\frac{\partial M_i}{\partial z} = E_i - D_i, \quad (2.7)$$

$$\frac{\partial}{\partial z}(M_i \bar{\phi}_i) = E_i \langle \phi \rangle - D_i \bar{\phi}_i + S_{\bar{\phi}_i}, \quad (2.8)$$

where  $E_i$  is the entrainment of environmental air into the  $i$ -th updraft,  $D_i$  is the detrainment, and  $S_{\bar{\phi}_i}$  is the net source or sink (Tiedtke, 1989). Equation (2.7) controls the intensity of convection as a function of height, and equation (2.8) modifies the thermodynamic properties of the updraft. Note that entrainment and detrainment may modify the thermodynamic state of the updraft without changing the mass flux, if  $E_i - D_i = 0$ .

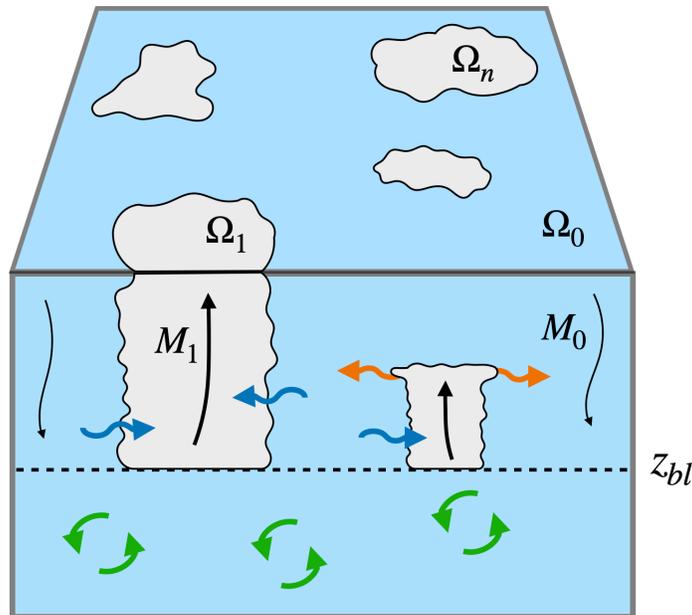


Figure 2.1: Schematic of the mass flux scheme decomposition of Arakawa and Schubert (1974) within an atmospheric column, from the surface to some height below the highest cloud top. A spectrum of ascending saturated thermals ( $\Omega_1, \dots, \Omega_n$ ) occupying a cross-sectional area fraction  $a_u$  (grey shading) is initialized atop a well-mixed turbulent boundary layer (height  $z_{bl}$ ). Their upward mass flux (e.g.,  $M_1$ , thick solid arrows) is compensated by large-scale environmental subsidence ( $M_0$ , thin descending arrows). The spectrum of thermals aims to represent different clouds; the top height of each cloud depends on their mass flux at  $z_{bl}$  and the net entrainment rate (blue arrows). Detrainment is concentrated at cloud top for each thermal (orange arrows).

Equation (2.7) uniquely specifies the mass flux given closures for entrainment and detrainment; independent equations for the area fraction  $a_i$  and the updraft vertical velocity  $\bar{w}_i$  are not necessary in classic mass flux schemes (Tiedtke, 1989). Nevertheless, modern mass flux schemes typically employ an additional equation for

updraft vertical velocity, which can be used to obtain better estimates of updraft top height (de Roode et al., 2012). In the majority of these schemes, the updraft vertical velocity equation follows Simpson and Wiggert (1969),

$$\frac{1}{2} \frac{\partial \bar{w}_i^2}{\partial z} = c_1 \bar{b}_i - c_2 \epsilon_i \bar{w}_i^2, \quad (2.9)$$

where  $\epsilon_i = E_i/M_i$  is the fractional entrainment rate,  $\bar{b}_i$  is the updraft buoyancy, and  $c_1, c_2$  are tunable parameters.

A schematic of the mass flux decomposition within an atmospheric column is shown in Figure 2.1. Although equation (2.7) does not prescribe the closures for entrainment and detrainment rates, classic mass-flux schemes concentrate all net detrainment ( $E_i - D_i < 0$ ) around the thermal termination height, which is diagnosed as the height of neutral buoyancy (Arakawa and Schubert, 1974; Tiedtke, 1989), or as the height of neutral vertical velocity in models that use equation (2.9). Therefore, the integrated vertical transport of each thermal depends on its cloud base mass flux and entrainment processes. Many parameterizations of entrainment have been proposed (Houghton and Cramer, 1951; Squires, 1958; Turner, 1962; Squires and Turner, 1962; Randall and Huffman, 1982; Kain and Fritsch, 1990; de Rooy et al., 2013), and finding optimal formulations continues to be an active area of research (Savre and Herzog, 2019; Cohen et al., 2020; Shin and Baik, 2022).

The simple set of equations (2.2)–(2.8) provides a framework to represent explicitly many of the physical phenomena observed in convective regions of Earth’s atmosphere. For instance, equation (2.6) implies that enhanced convective transport  $M_i$  by the updrafts results in stronger adiabatic heating and drying of the environment through compensating subsidence (Ooyama, 1971; Tiedtke, 1989). Considering the total specific humidity  $\phi = q_t$ , entrainment of environmental air dries the saturated updrafts through equation (2.8), reducing their buoyancy; detrainment moistens the environment, cooling it through evaporation (Nitta, 1975).

### 2.3 Convection meets turbulence: Eddy-diffusivity mass-flux (EDMF) schemes

Despite representing a step change in the parameterization of convection with respect to convective adjustment (e.g., Tiedtke, 1989), mass flux schemes still rely on several unrealistic approximations. As sketched in Figure 2.1, they assume that organized convective motion is initiated directly above the boundary layer, while in nature thermals typically originate close to the heat source; in this case Earth’s surface (Couvreur et al., 2010; Davini et al., 2017). This further results in large sensitivities

to the trigger function used to initiate convection (Stensrud and Fritsch, 1994; T. Wu, 2012; Saint-Martin et al., 2021). In addition, moist convective fluxes are considered independent from turbulence and dry convection within the boundary layer, which can lead to unrealistic discontinuities in the parameterized vertical fluxes (Siebesma et al., 2007; Villalba-Pradas and Tapiador, 2022).

To overcome these inconsistencies, Siebesma and Teixeira (2000) and Siebesma et al. (2007) proposed a new scheme that combines the parameterizations of convection and boundary layer turbulence in a single unified scheme, known as the eddy-diffusivity mass-flux (EDMF) scheme. Originally proposed to avoid the need for ad hoc countergradient fluxes in simulations of dry boundary layer convection (Siebesma et al., 2007), the scheme has since been successfully applied to a wide range of convective regimes, including stratocumulus and shallow cumulus clouds (Soares et al., 2004; Neggers, 2009).

The EDMF scheme relies on the same statistical decomposition into subdomains  $\Omega_0, \dots, \Omega_n$  of each atmospheric volume  $\Omega$  as classic mass flux schemes, with a few generalizations. First, EDMF schemes lift the assumption of field homogeneity within the environment. This allows them to capture environmental turbulent fluxes within the same framework. The covariance decomposition (2.4) then becomes

$$\rho \langle w^* \phi^* \rangle \approx \rho a_0 \overline{w'_0 \phi'_0} + \sum_{i=0}^n \rho a_i (\bar{w}_i - \langle w \rangle) (\bar{\phi}_i - \langle \phi \rangle), \quad (2.10)$$

where  $(\cdot)'_i$  denotes deviations about the subdomain average  $(\bar{\cdot})_i$ , and  $\overline{w'_0 \phi'_0}$  represents the vertical turbulent flux of  $\phi$  in the environment. Contrary to the mass flux decomposition, the EDMF statistical decomposition does not solve the closure problem in the grid-averaged equation (2.1) given entrainment and detrainment rates; a closure is still required for unresolved fluxes within the environment (i.e.,  $\overline{w'_0 \phi'_0}$ ). In EDMF schemes, these turbulent fluxes are parameterized as diffusive,

$$\overline{w'_0 \phi'_0} = -K_\phi \frac{\partial \bar{\phi}_0}{\partial z}, \quad (2.11)$$

where  $K_\phi$  is the environmental eddy diffusivity of field  $\phi$ . Retaining the diffusive term (2.11) in the definition of the unresolved flux (2.10) allows EDMF schemes to decompose the atmospheric column into updrafts and environment from the ground up, instead of relying on a different turbulence parameterization within the boundary layer.

Using the EDMF decomposition, the equation for  $\langle\phi\rangle$  can be written as

$$\frac{\partial\langle\phi\rangle}{\partial t} + \langle\mathbf{u}_h\rangle \cdot \nabla\langle\phi\rangle + \langle w\rangle \frac{\partial\langle\phi\rangle}{\partial z} = \frac{1}{\rho} \frac{\partial}{\partial z} \left( \rho a_0 K_\phi \frac{\partial\bar{\phi}_0}{\partial z} - \sum_{i=1}^n M_i (\bar{\phi}_i - \langle\phi\rangle) \right) + S_{\langle\phi\rangle}. \quad (2.12)$$

Some EDMF schemes make the approximation  $a_0 \approx 1$ , such that the area fraction does not enter the equation (2.12). In cases where the environmental area fraction is retained,  $a_0$  can be diagnosed from equation (2.2) if the updraft area fractions are known. Common implementations of the EDMF scheme set each updraft area fraction to a constant (Suselj et al., 2013; Neggers, 2015; Suselj et al., 2019b), and diagnose the mass flux by solving an equilibrium equation for the updraft vertical velocity similar to (2.9). Note that this procedure implicitly prescribes the detrainment rate through equation (2.7).

A schematic of the EDMF scheme decomposition is shown in Figure 2.2. In this case, mass flux terms in equation (2.12) represent both dry and moist convection; there is no discontinuity at cloud base. The EDMF decomposition circumvents the need to implement empirical countergradient transport terms below the lifting condensation level (Deardorff, 1966). It also eliminates moist convection triggers, an important advantage given that the physical basis for commonplace trigger functions (Kuo, 1974) is strongly questioned (Emanuel, 1991; Emanuel et al., 1994). Moist convective processes arise naturally when moisture within updrafts condenses (Neggers, 2009). Further, the continuous parameterization of the subgrid-scale flux  $\langle w^* \phi^* \rangle$  ensures a consistent representation of transport across vertical levels.

EDMF schemes such as the ones proposed by Siebesma et al. (2007) and Soares et al. (2004) have been extensively tested in a wide range of atmospheric regimes (Neggers, 2009; Suselj et al., 2013; Suselj et al., 2019a; Smalley et al., 2022), and recently in several configurations of oceanic convection (Giordani et al., 2020). They have been used in operational weather models to some extent for over 15 years now. The implementation of the EDMF scheme in the European Centre for Medium-Range Weather Forecasts (ECMWF) operational model led to drastic improvements in the representation of marine stratocumulus and winter stratus, although it is still not used operationally above the boundary layer in said model (Köhler, 2005; Köhler et al., 2011). Its implementation in the operational Navy Global Environmental Model led to reduced cold biases in the low troposphere over oceans, and an increase in geopotential height prediction skill (Hogan et al., 2014; Suselj et al., 2014).

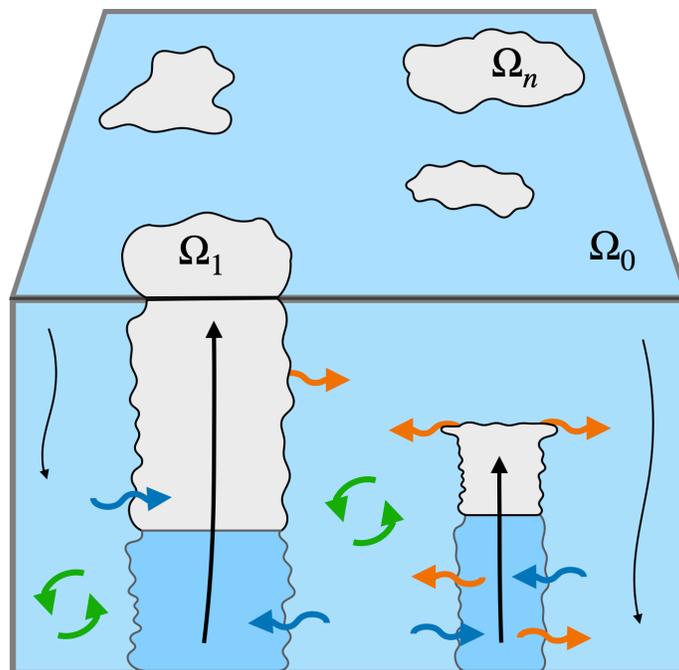


Figure 2.2: Schematic of the EDMF scheme decomposition within an atmospheric column, from the surface to some height below the highest cloud top. Updrafts are initialized at the surface, and interact with the environment through entrainment (blue arrows) and detrainment fluxes (orange arrows). Updrafts are unsaturated below their lifting condensation level, which may be different for different plumes, and become saturated above. Turbulent mixing induces further vertical transport in the environment at all heights (green arrows).

The demonstrated effectiveness of the EDMF scheme, together with its appealing physical explainability, have sparked considerable interest in the scientific community over the last decade (Witek et al., 2011a; Suselj et al., 2013; Neggers, 2015; Kurowski and Teixeira, 2018; Tan et al., 2018; Suselj et al., 2019a; E. Wu et al., 2020; Lopez-Gomez et al., 2020; Cohen et al., 2020; Q. Li et al., 2021), and a push toward the implementation of this scheme in operational weather and climate models (Kurowski et al., 2019; Suselj et al., 2021). As the horizontal resolution of weather and climate models approaches the 10–100 km scale, some of the assumptions of EDMF schemes inherited from earlier mass flux parameterizations are no longer justifiable. Plumes and convective clouds may cover time-varying and non-negligible fractions of each computational grid-box, and equilibrium assumptions about updraft properties may not hold when the ensemble of thermals they represent decreases in size (Randall, 2013; Neggers, 2015; Neggers and Griewank, 2022).

These challenges have prompted recent work on the derivation and analysis of

EDMF-like schemes through a systematic coarse-graining of the governing equations of fluid motion. The objective is to rigorously trace, and potentially revisit, all approximations used in the resulting parameterizations of atmospheric turbulence and convection. Such studies include the multi-fluid approach of Thuburn et al. (2018), Weller and McIntyre (2019), and Thuburn et al. (2022a); and the extended EDMF scheme of Tan et al. (2018), Cohen et al. (2020) and Lopez-Gomez et al. (2020). Section 2.4 focuses on the derivation of the extended EDMF scheme, following the formulation presented in Cohen et al. (2020). The full set of equations employed by the scheme is outlined in Section 2.5. Similarities and differences with respect to the multi-fluid approach and steady-state EDMF schemes are discussed in Section 2.6, followed by some concluding remarks. The extended EDMF scheme is then used as the basis for the turbulence closures and simulation results presented in Chapters 3–5.

## 2.4 Derivation of the extended EDMF scheme

The dynamics of atmospheric fluid flow are governed by the compressible Navier-Stokes equations, an equation defining the thermodynamic state of air, and the conservation of its energy and constituents (i.e., water species). Moist adiabatic processes due to the condensation, evaporation and freezing of water have a considerable impact on atmospheric dynamics, so a prognostic equation for the total water specific humidity  $q_t$  is retained. Conservation of energy is imposed through a thermodynamic balance law that defines the energy of the fluid once the specific humidity  $q_t$  is known. The exact thermodynamic balance law used is a design choice; formulations in terms of dry static energy (Tiedtke, 1989), entropy (Pressel et al., 2015; Thuburn et al., 2022a), or a moist adiabatic potential temperature (Tan et al., 2018; Cohen et al., 2020) may all be found in the literature. Here, we choose to retain an equation for the liquid-ice potential temperature  $\theta_l$  (Deardorff, 1976),

$$\theta_l = \frac{T}{\pi} \exp\left(\frac{-L_v(q_l + q_i)}{c_p T}\right), \quad (2.13)$$

where  $\pi$  is the Exner function,  $L_v$  is the effective latent heat of water (Pressel et al., 2015),  $q_l$  and  $q_i$  are the liquid and ice water specific humidities,  $c_p$  is the isobaric specific heat of air and  $T$  is the temperature. We approximate  $c_p$  as constant and, consistent with Kirchhoff's law, the latent heat of vaporization  $L_v$  to be a linear function of temperature (Romps, 2008).

The full set of equations in conservation form reads

$$\frac{\partial \rho}{\partial t} + \nabla_h \cdot (\rho \mathbf{u}_h) + \frac{\partial(\rho w)}{\partial z} = 0, \quad (2.14)$$

$$\frac{\partial(\rho \mathbf{u}_h)}{\partial t} + \nabla_h \cdot (\rho \mathbf{u}_h \otimes \mathbf{u}_h) + \frac{\partial(\rho w \mathbf{u}_h)}{\partial z} = -\nabla_h p^\dagger + \rho S_{\mathbf{u}_h}, \quad (2.15)$$

$$\frac{\partial(\rho w)}{\partial t} + \nabla_h \cdot (\rho \mathbf{u}_h w) + \frac{\partial(\rho w w)}{\partial z} = \rho b - \frac{\partial p^\dagger}{\partial z} + \rho S_w, \quad (2.16)$$

$$\frac{\partial(\rho \theta_l)}{\partial t} + \nabla_h \cdot (\rho \mathbf{u}_h \theta_l) + \frac{\partial(\rho w \theta_l)}{\partial z} = \rho S_{\theta_l}, \quad (2.17)$$

$$\frac{\partial(\rho q_t)}{\partial t} + \nabla_h \cdot (\rho \mathbf{u}_h q_t) + \frac{\partial(\rho w q_t)}{\partial z} = \rho S_{q_t}, \quad (2.18)$$

$$p = \rho R_d T_v. \quad (2.19)$$

The momentum sources  $S_{\mathbf{u}_h}$  and  $S_w$  include the effect of molecular viscous stress and Coriolis forces. In equations (2.17) and (2.18), the sources  $S_{\theta_l}$  and  $S_{q_t}$  represent the effect of molecular diffusion, microphysical processes, and radiation. In the momentum equations, a reference pressure profile  $p_h(z)$  in hydrostatic balance with a reference density  $\rho_h(z)$  is removed to improve numerical stability:

$$\frac{\partial p_h}{\partial z} = -\rho_h g,$$

where  $g$  is the gravitational acceleration. Therefore, the perturbation pressure

$$p^\dagger = p - p_h$$

and the buoyancy

$$b = -g \frac{\rho - \rho_h}{\rho}$$

replace the full pressure  $p$  and gravitational acceleration  $g$  in the momentum equations. In the equation of state (2.19),  $T_v$  is the virtual temperature,

$$T_v = \frac{R_m}{R_v} T, \quad (2.20)$$

where  $R_m$  and  $R_d$  are the gas constants for moist and dry air, respectively. The temperature  $T$  and the distribution of  $q_t$  among water species are obtained from the thermodynamic variables  $\theta_l$ ,  $\rho$ , and  $q_t$  through a saturation adjustment procedure (Bryan and Fritsch, 2002); we do not consider deviations from thermodynamic equilibrium in our formulation, although this assumption can easily be relaxed by adding prognostic equations for the condensed species.

### Grid-averaged equations of motion

In practice, a discrete resolution  $(\Delta x, \Delta y, \Delta z)$  must be used when the system of equations (2.14)–(2.18) is solved numerically. Analytically, this amounts to solving the weak form of equations (2.14)–(2.18), averaged over some volume  $\Omega$ . Denoting Favre averages over this volume as  $\langle \cdot \rangle$ , deviations about the Favre average as  $(\cdot)^*$ , Reynolds averages as  $\langle \cdot \rangle_R$ , and the Reynolds-averaged fluid density as  $\rho := \langle \rho \rangle_R$ , the resulting equations are

$$\frac{\partial \rho}{\partial t} + \nabla_h \cdot (\rho \langle \mathbf{u}_h \rangle) + \frac{\partial (\rho \langle w \rangle)}{\partial z} = 0, \quad (2.21)$$

$$\begin{aligned} \frac{\partial (\rho \langle \mathbf{u}_h \rangle)}{\partial t} + \nabla_h \cdot (\rho \langle \mathbf{u}_h \rangle \otimes \langle \mathbf{u}_h \rangle) + \frac{\partial (\rho \langle w \rangle \langle \mathbf{u}_h \rangle)}{\partial z} = \\ - \nabla_h \cdot (\rho \langle \mathbf{u}_h^* \otimes \mathbf{u}_h^* \rangle) - \frac{\partial (\rho \langle w^* \mathbf{u}_h^* \rangle)}{\partial z} - \nabla_h \langle p^\dagger \rangle_R + \rho \mathcal{S}_{\langle \mathbf{u}_h \rangle}, \end{aligned} \quad (2.22)$$

$$\begin{aligned} \frac{\partial (\rho \langle w \rangle)}{\partial t} + \nabla_h \cdot (\rho \langle \mathbf{u}_h \rangle \langle w \rangle) + \frac{\partial (\rho \langle w \rangle \langle w \rangle)}{\partial z} = \\ - \nabla_h \cdot (\rho \langle \mathbf{u}_h^* w^* \rangle) - \frac{\partial (\rho \langle w^* w^* \rangle)}{\partial z} + \rho \langle b \rangle - \frac{\partial \langle p^\dagger \rangle_R}{\partial z} + \rho \mathcal{S}_{\langle w \rangle}, \end{aligned} \quad (2.23)$$

$$\begin{aligned} \frac{\partial (\rho \langle \theta_l \rangle)}{\partial t} + \nabla_h \cdot (\rho \langle \mathbf{u}_h \rangle \langle \theta_l \rangle) + \frac{\partial (\rho \langle w \rangle \langle \theta_l \rangle)}{\partial z} = \\ - \nabla_h \cdot (\rho \langle \mathbf{u}_h^* \theta_l^* \rangle) - \frac{\partial (\rho \langle w^* \theta_l^* \rangle)}{\partial z} + \rho \mathcal{S}_{\langle \theta_l \rangle}, \end{aligned} \quad (2.24)$$

$$\begin{aligned} \frac{\partial (\rho \langle q_t \rangle)}{\partial t} + \nabla_h \cdot (\rho \langle \mathbf{u}_h \rangle \langle q_t \rangle) + \frac{\partial (\rho \langle w \rangle \langle q_t \rangle)}{\partial z} = \\ - \nabla_h \cdot (\rho \langle \mathbf{u}_h^* q_t^* \rangle) - \frac{\partial (\rho \langle w^* q_t^* \rangle)}{\partial z} + \rho \mathcal{S}_{\langle q_t \rangle}, \end{aligned} \quad (2.25)$$

$$\langle p \rangle_R = \rho R_d \langle T_v \rangle. \quad (2.26)$$

In equations (2.22)–(2.25), none of the subgrid-scale (SGS) covariances can be resolved numerically, they must be parameterized. As discussed in earlier sections, our focus here is on obtaining a unified scheme to represent vertical SGS fluxes arising from turbulence and convection; horizontal SGS fluxes are typically represented by eddy diffusion. To this end, we follow a statistical decomposition of the fluid flow into subdomains, as done by earlier EDMF schemes.

### The subdomain-averaged equations

The extended EDMF scheme decomposes each grid-cell volume  $\Omega$  into subdomains  $\Omega_0, \dots, \Omega_n$ , each of them occupying a volume fraction  $a_i$  that, importantly, is not necessarily negligible nor stationary. The sum of volume fractions satisfies equation (2.2). We make the following approximations regarding the subdomains in our derivation:

- The fluid is approximated as anelastic, meaning that the Reynolds-averaged subdomain density  $\rho_i$  is taken to be a function of height and constant across subdomains,  $\rho_i \approx \rho$ . This approximation is used everywhere except in buoyancy terms, where the true subdomain densities are retained through the definition of the subdomain buoyancy  $\bar{b}_i$ ,

$$\bar{b}_i = -g \frac{\rho_i - \rho_h}{\rho}; \quad (2.27)$$

- The horizontal velocity is assumed constant across subdomains,  $\bar{\mathbf{u}}_{h,i} = \langle \mathbf{u}_h \rangle$ . This essentially enforces convection to be a vertical process at the subdomain scale, limiting horizontal transport processes to diffusion and large-scale advection.

In practical implementations of the extended EDMF scheme, other approximations are made, the most important being the neglect of subgrid-scale variability within the updrafts  $\Omega_1, \dots, \Omega_n$  (Cohen et al., 2020). This is an approximation done a posteriori; here we retain all subgrid-scale variability when deriving the equations.

The equations of motion for subdomain  $\Omega_i$  are obtained by averaging equations (2.14) – (2.18) over its volume. Considering the total water specific humidity equation (2.18) as an example,

$$\int_{\Omega_i(t)} \frac{\partial(\rho q_t)}{\partial t} dV + \int_{\Omega_i(t)} \nabla \cdot (\rho q_t \mathbf{u}) dV = \int_{\Omega_i(t)} \rho S_{q_t} dV. \quad (2.28)$$

We do not assume steady state of each subdomain  $\Omega_i$  as done in earlier EDMF schemes, so the domain of integration in equation (2.28) generally has moving boundaries. Let us express the subdomain boundary  $\partial\Omega_i$  as the union  $\partial\Omega_i = \partial\Omega_i^g \cup \partial\Omega_i^{sg}$ , where  $\partial\Omega_i^g = \partial\Omega_i \cap \partial\Omega$  is the part of the subdomain boundary  $\partial\Omega_i$  that coincides with the grid-box boundary  $\partial\Omega$ , and  $\partial\Omega_i^{sg}$  the subdomain boundary that remains within the volume  $\Omega$ . The domain and subdomain boundaries are related through  $\sum_i \partial\Omega_i^g = \partial\Omega$ . The subgrid boundary  $\partial\Omega_i^{sg}$  is a free moving surface with

velocity  $\mathbf{u}_b$ , while boundary  $\partial\Omega_i^g$  may only move within  $\partial\Omega$ , by definition. Using the Reynolds transport theorem for the transient term in equation (2.28),

$$\frac{\partial}{\partial t} \int_{\Omega_i(t)} \rho q_t dV + \int_{\Omega_i(t)} \nabla \cdot (\rho q_t \mathbf{u}) dV = \int_{\partial\Omega_i^{sg}(t)} \rho q_t \mathbf{u}_b \cdot \mathbf{n} dS + \int_{\Omega_i(t)} \rho S_{q_t} dV, \quad (2.29)$$

where  $\mathbf{n}$  is the outward-pointing unit vector normal to the surface over which the integration is performed. Since  $\mathbf{u}_b \cdot \mathbf{n} = 0$  over  $\partial\Omega_i^g$ , no relative transport term through this boundary appears in equation (2.29). Then, using the Gauss-Ostrogradsky theorem for the divergence term, and rearranging the surface integrals yields

$$\frac{\partial}{\partial t} \int_{\Omega_i(t)} \rho q_t dV + \int_{\partial\Omega_i^g} \rho q_t \mathbf{u} \cdot \mathbf{n} dS = - \int_{\partial\Omega_i^{sg}(t)} \rho q_t (\mathbf{u} - \mathbf{u}_b) \cdot \mathbf{n} dS + \int_{\Omega_i(t)} \rho S_{q_t} dV. \quad (2.30)$$

The first term on the right-hand side is the flux out of subdomain  $\Omega_i$  into other subdomains within the same grid box, and the second term on the left-hand side is the flux out of subdomain  $\Omega_i$  into a neighboring grid-box. The total grid-scale divergence equals the sum of fluxes from all subdomains across the grid box,

$$\int_{\Omega} \nabla \cdot (\rho q_t \mathbf{u}) dV = \sum_{i \geq 0} \int_{\partial\Omega_i^g} \rho q_t \mathbf{u} \cdot \mathbf{n} dS. \quad (2.31)$$

Commutativity of the divergence and the volume average is exact for uniform grids and results in a small error otherwise (Fureby and Tabor, 1997). Using this property,

$$\nabla \cdot \int_{\Omega} (\rho q_t \mathbf{u}) dV = \nabla \cdot (\rho V_T \langle q_t \mathbf{u} \rangle) = \sum_{i \geq 0} \int_{\partial\Omega_i^g} \rho q_t \mathbf{u} \cdot \mathbf{n} dS, \quad (2.32)$$

where  $V_T$  is the total volume of  $\Omega$ . The divergence term in equation (2.32) can be written in terms of the sum of the subdomain-mean values through the domain decomposition (2.3),

$$\sum_{i \geq 0} \nabla \cdot [\rho V_i (\overline{q_t \mathbf{u}})_i] = \sum_{i \geq 0} \int_{\partial\Omega_i^g} \rho q_t \mathbf{u} \cdot \mathbf{n} dS, \quad (2.33)$$

where  $V_i$  is the volume of subdomain  $\Omega_i$ , and (2.33) holds generally. The divergence in equation (2.33) acts on the grid scale. The diagnosis of the contribution of each subdomain flux on the right-hand side of equation (2.33) to the grid-mean divergence requires an assumption regarding the fraction of the grid-cell boundary  $\partial\Omega$  covered by each subdomain,  $\partial\Omega_i^g$ . Here, we assume that  $A_i^g = a_i A_T^g$ , where  $A_i^g$  and  $A_T^g$  are the areas of surfaces  $\partial\Omega_i^g$  and  $\partial\Omega$ , respectively. We further assume that for each  $\Omega_i$  the average over  $\partial\Omega_i^g$  equals the subdomain mean. From this it follows that

$$\int_{\partial\Omega_i^g} \rho q_t \mathbf{u} \cdot \mathbf{n} dS = \nabla \cdot [\rho V_i (\overline{q_t \mathbf{u}})_i] = \nabla \cdot [\rho V_i (\bar{q}_{t,i} \bar{\mathbf{u}}_i + \overline{q'_{t,i} \mathbf{u}'_i})]. \quad (2.34)$$

Equation (2.34) cannot be obtained from the divergence theorem, since  $\partial\Omega_i^{sg}$  is not a closed surface. Using equality (2.34) and dividing by the grid-box volume  $V_T$ , we can rewrite (2.30) as

$$\frac{\partial(\rho a_i \bar{q}_{t,i})}{\partial t} = -\nabla \cdot [\rho a_i (\bar{q}_{t,i} \bar{\mathbf{u}}_i + \overline{q'_{t,i} \mathbf{u}'_i})] - \frac{1}{V_T} \int_{\partial\Omega_i^{sg}(t)} \rho q_t \mathbf{u}_r \cdot \mathbf{n} dS + \rho a_i \bar{S}_{q_{t,i}}, \quad (2.35)$$

where  $\mathbf{u}_r = \mathbf{u} - \mathbf{u}_b$ . Since the vertical extent of the volumes is fixed at the model vertical resolution,  $V_i/V_T = A_i/A_T = a_i$ , with  $a_i$  as the area fraction,  $A_T = \Delta x \Delta y$ , and  $A_i$  the average horizontal cross-sectional area of subdomain  $\Omega_i$ .

Thus, the subdomain humidity equation (2.35) can be obtained from a statistical domain decomposition and the subdomain interior-boundary homogeneity assumption (2.34). The divergence of the flux  $\overline{q'_{t,i} \mathbf{u}'_i}$  represents within-subdomain turbulent transport, and the integral term represents transport across the subdomain boundary to other subdomains, which we refer to as entrainment and detrainment processes. These two terms require closures in terms of subdomain-mean variables for equation (2.35) to be solvable.

### Entrainment and detrainment fluxes

The net entrainment flux in equation (2.35) can be expressed as the sum of a bulk entrainment flux, denoted *dynamical* entrainment, and an entrainment flux due to turbulent correlations between the across-subdomain velocity  $\mathbf{u}_r$  and the prognostic variable, denoted *turbulent* entrainment,

$$\frac{1}{V_T} \int_{\partial\Omega_i^{sg}(t)} \rho q_t \mathbf{u}_r \cdot \mathbf{n} dS = \rho \frac{A_i^{sg}}{V_T} \left( \underbrace{\widehat{q_t u_{r,n}}}_{\text{dynamical}} + \underbrace{\widehat{q'_t u'_{r,n}}}_{\text{turbulent}} \right)_i, \quad (2.36)$$

where symbol  $\widehat{(\cdot)}$  represents the average over the surface  $\partial\Omega_i^{sg}$ ,  $u_{r,n} = \mathbf{u}_r \cdot \mathbf{n}$ , and  $A_i^{sg}$  is the total area of surface  $\partial\Omega_i^{sg}$ .

The net dynamical entrainment flux is taken to be the sum of mass entrainment and detrainment processes that carry the mean properties of the subdomain they detrain from (de Rooy et al., 2013),

$$-\frac{A_i^{sg}}{V_T} (\rho \widehat{q_t u_{r,n}})_i = \sum_{j \neq i} (E_{ij} \bar{q}_{t,j} - \Delta_{ij} \bar{q}_{t,i}), \quad (2.37)$$

where  $E_{ij}$  is the mass entrainment into subdomain  $\Omega_i$  from  $\Omega_j$ , and  $\Delta_{ij}$  is the detrainment out of  $\Omega_i$  into  $\Omega_j$ . Both of these terms are positive semidefinite, due to our upwind model for  $\hat{q}_t$ : expression (2.37) assumes that the average specific

humidity over  $\partial\Omega_i^{sg}$  depends on the sign of  $\widehat{u_{r,n}}$ : it is equal to the subdomain-mean  $\bar{q}_{t,i}$  if air is being detrained from  $\Omega_i$ , and equal to  $\bar{q}_{t,j}$  if air is being entrained from a different subdomain  $\Omega_j$ . Other models of dynamical entrainment, based on preferential sorting arguments, have been proposed in the literature (Thuburn et al., 2022b).

Since density is taken constant across subdomains, the dynamical entrainment and detrainment fluxes satisfy

$$-\frac{A_i^{sg}}{V_T}(\rho\widehat{u_{r,n}})_i = \sum_{j \neq i} (E_{ij} - \Delta_{ij}). \quad (2.38)$$

The turbulent entrainment flux does not involve net mass exchange between subdomains, and it is modeled as a downgradient diffusive flux across the subdomain interface (Asai and Kasahara, 1967; Cohen et al., 2020),

$$-\frac{A_i^{sg}}{V_T}(\rho\widehat{q'_t u'_{r,n}})_i = \sum_{j \neq i} \hat{E}_{ij}(\bar{q}_{t,j} - \bar{q}_{t,i}). \quad (2.39)$$

where  $\hat{E}_{ij} = \hat{E}_{ji}$  is the turbulent entrainment rate between  $\Omega_j$  and  $\Omega_i$ . Note that in regions where both  $E_{ij}$  and  $\Delta_{ij}$  are positive, the entrainment offset by detrainment acts as a turbulent entrainment term similar to (2.39) (Tiedtke, 1989).

Using the decomposition (2.36), expressions (2.37) and (2.39) for the entrainment fluxes, and taking the large-scale horizontal velocity to be constant across subdomains, equation (2.35) may be written as

$$\begin{aligned} \frac{\partial(\rho a_i \bar{q}_{t,i})}{\partial t} + \nabla_h \cdot (\rho a_i \bar{q}_{t,i} \langle \mathbf{u}_h \rangle) + \frac{\partial(\rho a_i \bar{q}_{t,i} \bar{w}_i)}{\partial z} = -\nabla \cdot (\rho a_i \overline{q'_{t,i} \mathbf{u}'_i}) \\ + \sum_{j \neq i} (E_{ij} \bar{q}_{t,j} - \Delta_{ij} \bar{q}_{t,i}) + \sum_{j \neq i} \hat{E}_{ij}(\bar{q}_{t,j} - \bar{q}_{t,i}) + \rho a_i \bar{S}_{q_{t,i}}. \end{aligned} \quad (2.40)$$

Equation (2.40) is the subdomain-averaged balance law for total water specific humidity. The sum of this equation over all subdomains yields the grid-mean equation (2.25). The equations for all other subdomain-averaged thermodynamic fields, such as the liquid-water potential temperature  $\bar{\theta}_{l,i}$ , follow an identical formulation,

$$\begin{aligned} \frac{\partial(\rho a_i \bar{\theta}_{l,i})}{\partial t} + \nabla_h \cdot (\rho a_i \bar{\theta}_{l,i} \langle \mathbf{u}_h \rangle) + \frac{\partial(\rho a_i \bar{\theta}_{l,i} \bar{w}_i)}{\partial z} = -\nabla \cdot (\rho a_i \overline{\theta'_{l,i} \mathbf{u}'_i}) \\ + \sum_{j \neq i} (E_{ij} \bar{\theta}_{l,j} - \Delta_{ij} \bar{\theta}_{l,i}) + \sum_{j \neq i} \hat{E}_{ij}(\bar{\theta}_{l,j} - \bar{\theta}_{l,i}) + \rho a_i \bar{S}_{\theta_{l,i}}. \end{aligned} \quad (2.41)$$

Since density is taken to be constant across subdomains, except in terms accounting for buoyancy effects, the subdomain-averaged continuity equation becomes a balance law for the subdomain area fraction  $a_i$ ,

$$\frac{\partial(\rho a_i)}{\partial t} + \nabla_h \cdot (\rho a_i \langle \mathbf{u}_h \rangle) + \frac{\partial(\rho a_i \bar{w}_i)}{\partial z} = \sum_{j \neq i} (E_{ij} - \Delta_{ij}). \quad (2.42)$$

Equation (2.42) follows the same derivation as (2.40), substituting  $q_t \rightarrow 1$ , so all covariances in the equation vanish. Note that this balance law is a generalized version of the steady-state mass flux equation (2.7) used in mass flux schemes and steady EDMF schemes.

A sketch of the extended EDMF decomposition is shown in Figure 2.3, including updrafts for deep, shallow, and decaying convection. Contrary to steady EDMF schemes, net entrainment processes and large-scale horizontal convergence may modify the updraft area fraction and mass flux both in time and space. These additions enable a more realistic representation of cloud cover, endow the extended EDMF scheme with convective memory, and enable a consistent representation of transient processes such as the diurnal cycle of convection (Tan et al., 2018; Cohen et al., 2020).

Finally, the subdomain-averaged vertical momentum equation includes two additional terms with respect to the thermodynamic balance laws, related to buoyancy forces and pressure gradients. Buoyancy forces are computed using an SGS anelastic approximation, following Pauluis (2008). Within this approximation, the subdomain density  $\bar{\rho}_i$  used to compute the subdomain buoyancy (2.27) is defined as

$$\rho_i = \frac{\langle p \rangle_R}{R_d \bar{T}_{v,i}}, \quad (2.43)$$

which ensures that the virtual temperature  $T_v$  satisfies the Favre average

$$\rho \langle T_v \rangle = \sum_{i=0}^n \rho_i a_i \bar{T}_{v,i}. \quad (2.44)$$

The Reynolds-averaged subdomain perturbation pressure  $\bar{p}_i^\dagger$  is defined consistently with this definition of buoyancy, and following hydrostatic balance, as shown in Appendix C of Cohen et al. (2020). From the SGS anelastic approximation and

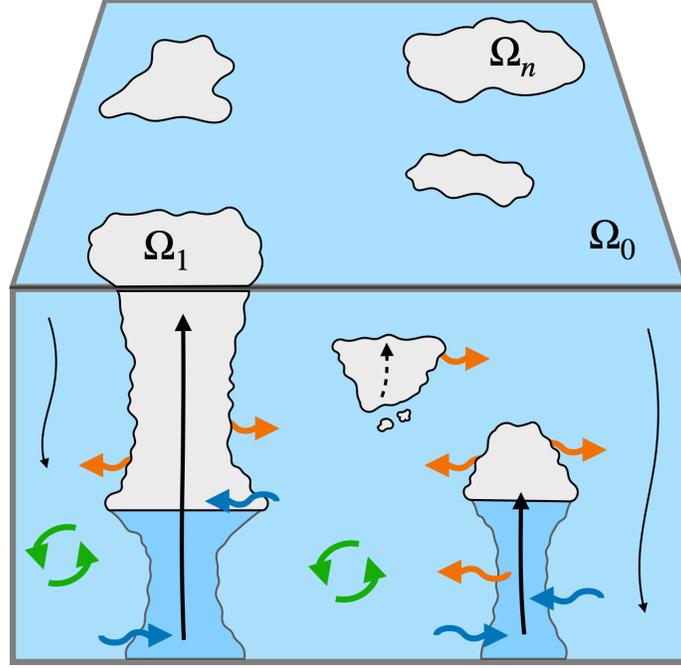


Figure 2.3: Schematic of the extended EDMF scheme decomposition within an atmospheric column, from the surface to some height below the highest cloud top. Arrows as in Figure 2.2. The thermodynamic state of the updrafts at the surface, along with entrainment and detrainment processes, determine their area fraction, mass flux, and termination height. Transient terms allow mature convective clouds to dissipate over a finite amount of time.

definition of buoyancy, the subdomain-averaged vertical momentum equation reads

$$\begin{aligned} \frac{\partial(\rho a_i \bar{w}_i)}{\partial t} + \nabla_h \cdot (\rho a_i \langle \mathbf{u}_h \rangle \bar{w}_i) + \frac{\partial(\rho a_i \bar{w}_i \bar{w}_i)}{\partial z} = -\nabla \cdot (\rho a_i \overline{\mathbf{u}'_i w'_i}) \\ + \sum_{j \neq i} \left[ (E_{ij} + \hat{E}_{ij}) \bar{w}_j - (\Delta_{ij} + \hat{\Delta}_{ij}) \bar{w}_i \right] + \rho a_i \bar{b}_i - \rho a_i \frac{\partial}{\partial z} \left( \frac{\bar{p}_i^\dagger}{\rho} \right) + \rho a_i \bar{S}_{w,i}. \end{aligned} \quad (2.45)$$

This equation is the extended EDMF version of the equilibrium equation (2.9) used in steady EDMF schemes. Here, large-scale horizontal advection, pressure gradients, and turbulent fluxes within the subdomain may modify the subdomain vertical velocity evolution. Importantly, no empirical parameter multiplies the buoyancy and entrainment terms. This highlights that parameters  $c_1$  and  $c_2$  in equation (2.9) represent a closure of the pressure gradient term; explicit perturbation pressure gradient closures typically follow a similar functional form (J. He et al., 2021).

### Closures within the extended EDMF scheme

Equations (2.40)–(2.45) define the subdomain dynamics given the grid-mean equations and parameterizations for the following unresolved processes:

- All within-subdomain turbulent fluxes of the form  $\overline{\mathbf{u}'_i \phi'_i}$ . These include horizontal and vertical turbulent fluxes;
- Dynamical entrainment  $E_{ij}$ , detrainment  $\Delta_{ij}$ , and turbulent entrainment  $\hat{E}_{ij}$ ;
- Subdomain perturbation pressure gradients  $\partial(\bar{p}_i^\dagger/\rho)/\partial z$  in the vertical momentum equations;
- Heating (or cooling) and moistening (or drying) due to cloud microphysical processes, which partially define the terms  $\bar{S}_{q_{t,i}}$  and  $\bar{S}_{\theta_{l,i}}$  in equations (2.40) and (2.41) and can have a strong influence in moist atmospheric dynamics.

Parameterizations of these processes are necessarily imperfect approximations of their real effect on the dynamics of the atmosphere, be they entirely data-driven or based on simplified physical balances (Schneider et al., 2017b; Lopez-Gomez et al., 2022b). As such, they introduce important representation or model errors in the dynamics (Brynjarsdóttir and O’Hagan, 2014; van Leeuwen, 2015). These errors have been shown to be major contributors to the biases and uncertainty in state-of-the-art climate projections, which is why improving parameterizations is one of the core efforts in atmospheric dynamics research today (Sherwood et al., 2014; Zelinka et al., 2020). Examples of closures designed for extended EDMF schemes are the diffusive turbulent flux closure proposed in Lopez-Gomez et al. (2020), the entrainment and detrainment closures found in Cohen et al. (2020), and the perturbation pressure gradient closure introduced by J. He et al. (2021); the first of these closures is the subject of Chapter 3.

### **The subdomain covariance equations**

Many proposed parameterizations for unresolved processes do not have a closed diagnostic form. Instead, they rely on additional prognostic equations that carry information about subgrid-scale variability. By computing spatially filtered versions of products of the equations (2.14)–(2.18), we may obtain prognostic equations for higher order statistical moments of the state  $\{\mathbf{u}, \theta_l, q_t\}$ . A hierarchy of systems of equations can then be obtained by selecting the highest order of the statistical moments for which prognostic equations are retained. Mellor and Yamada (1974) provide a description of this hierarchy when filtering at the scale of the grid-volume  $\Omega$ ; a similar hierarchy applied to the subdomain equations is given by Thuburn et al. (2022a).

The extended EDMF scheme proposed by Tan et al. (2018) and Cohen et al. (2020) introduce additional equations for the environmental turbulence kinetic energy, and for the environmental covariance of the moist adiabatic variables  $q_t$  and  $\theta_l$ . The environmental turbulence kinetic energy is used as an ingredient for all within-subdomain turbulent flux closures, and the thermodynamic covariances are used to estimate the average effect of processes that depend nonlinearly on the thermodynamic state of the fluid due to phase transitions (Lopez-Gomez et al., 2020). No covariance equations are carried for the updrafts, wherein correlations between turbulent fluctuations of all fields are assumed negligible; updraft variability can be captured by increasing the number of subdomains. In the Mellor and Yamada (1974) hierarchy, the extended EDMF scheme combines an environmental level 3 model with an updraft "level 0" model<sup>3</sup>. Common hierarchies for EDMF schemes involve the use of an environmental level 2.5 model, which retain a single second-order moment equation for the turbulence kinetic energy (Mellor and Yamada, 1982; Kurowski and Teixeira, 2018).

The derivation of the subdomain covariance equation for two thermodynamic fields, for instance  $q_t$  and  $\theta_l$ , follows a similar structure to the derivation of the subdomain mean equation outlined in previous pages. Consider equation (2.35) with the change of variable  $q_t \rightarrow q_t\theta_l$ ,

$$\begin{aligned} \frac{\partial(\rho a_i \overline{q_{t,i}\theta_{l,i}})}{\partial t} = & -\nabla \cdot [\rho a_i (\overline{q_{t,i}\theta_{l,i}} \bar{\mathbf{u}}_i + \overline{(q_{t,i}\theta_{l,i})' \mathbf{u}'_i})] \\ & - \frac{1}{V_T} \int_{\partial\Omega_i^{sg}(t)} \rho q_t \theta_l \mathbf{u}_r \cdot \mathbf{n} dS + \rho a_i \bar{S}_{\theta_{l,i}, q_{t,i}}. \end{aligned} \quad (2.46)$$

Using the entrainment decomposition (2.36),

$$\begin{aligned} \frac{\partial(\rho a_i \overline{q_{t,i}\theta_{l,i}})}{\partial t} = & -\nabla \cdot [\rho a_i (\overline{q_{t,i}\theta_{l,i}} \bar{\mathbf{u}}_i + \overline{(q_{t,i}\theta_{l,i})' \mathbf{u}'_i})] \\ & - \frac{A_i^{sg}}{V_T} \left( \overline{\rho q_t \theta_l \widehat{u}_{r,n}} + \overline{\rho (q_t \theta_l)' \widehat{u}'_{r,n}} \right)_i + \rho a_i \bar{S}_{\theta_{l,i}, q_{t,i}}. \end{aligned} \quad (2.47)$$

The equation for covariance  $\overline{q'_{t,i}\theta'_{l,i}}$  can be obtained from equation (2.47), the subdomain-mean equations (2.40) and (2.41), and the subdomain continuity equa-

<sup>3</sup>Level 1, the lowest level in the Mellor-Yamada hierarchy, provides diagnostic equations for turbulent covariances. Here, we refer to a lower level in which all covariances are neglected.

tion (2.42),

$$\begin{aligned} \frac{\partial(\rho a_i \overline{q'_{t,i} \theta'_{l,i}})}{\partial t} &= \frac{\partial(\rho a_i \overline{q_{t,i} \theta_{l,i}})}{\partial t} \\ &\quad - \bar{\theta}_{l,i} \frac{\partial(\rho a_i \bar{q}_{t,i})}{\partial t} - \bar{q}_{t,i} \frac{\partial(\rho a_i \bar{\theta}_{l,i})}{\partial t} + \bar{q}_{t,i} \bar{\theta}_{l,i} \frac{\partial(\rho a_i)}{\partial t}, \end{aligned} \quad (2.48)$$

which after substitution reads

$$\begin{aligned} \frac{\partial(\rho a_i \overline{q'_{t,i} \theta'_{l,i}})}{\partial t} + \nabla_h \cdot (\rho a_i \langle \mathbf{u}_h \rangle \overline{q'_{t,i} \theta'_{l,i}}) + \frac{\partial(\rho a_i \overline{w_i q'_{t,i} \theta'_{l,i}})}{\partial z} &= \\ - \nabla \cdot (\rho a_i \overline{\mathbf{u}'_i q'_{t,i} \theta'_{l,i}}) - \rho a_i \overline{w'_i q'_{t,i}} \frac{\partial \bar{\theta}_{l,i}}{\partial z} - \rho a_i \overline{w'_i \theta'_{l,i}} \frac{\partial \bar{q}_{t,i}}{\partial z} & \\ - \rho \frac{A_i^{sg}}{V_T} \left( \overline{q'_t \theta'_l u'_{r,n}} - (\bar{\theta}_{l,i} - \hat{\theta}_l) \overline{u'_{r,n} q'_t} - (\bar{q}_{t,i} - \hat{q}_t) \overline{u'_{r,n} \theta'_l} \right)_i & \\ - \rho \frac{A_i^{sg}}{V_T} \left( \hat{u}_{r,n} (\hat{q}_t - \bar{q}_{t,i}) (\hat{\theta}_l - \bar{\theta}_{l,i}) + \hat{u}_{r,n} \overline{q'_t \theta'_l} \right)_i & \\ - \rho a_i \overline{D_{q'_{t,i} \theta'_{l,i}}} + \rho a_i (\overline{S'_{q_{t,i} \theta'_{l,i}}} + \overline{S'_{\theta_{l,i} q'_{t,i}}}), & \end{aligned} \quad (2.49)$$

where  $\overline{D_{q'_{t,i} \theta'_{l,i}}}$  is the turbulent covariance dissipation. Here, differences between subdomain-means and the interface value,  $\bar{\phi}_i - \hat{\phi}$ , must be modeled as  $\bar{\phi}_i^* = \bar{\phi}_i - \langle \phi \rangle$  to ensure conservation of the grid-mean covariance. The last term in equation (2.49) follows from the decomposition (2.48), given that

$$\overline{S_{q_{t,i} \theta_{l,i}}} = \overline{q_{t,i} S_{\theta_{l,i}}} + \overline{\theta_{l,i} S_{q_{t,i}}}. \quad (2.50)$$

Finally, substituting the entrainment fluxes (2.37) and (2.39) in equation (2.49), we obtain

$$\begin{aligned} \frac{\partial(\rho a_i \overline{q'_{t,i} \theta'_{l,i}})}{\partial t} + \nabla_h \cdot (\rho a_i \langle \mathbf{u}_h \rangle \overline{q'_{t,i} \theta'_{l,i}}) + \frac{\partial(\rho a_i \overline{w_i q'_{t,i} \theta'_{l,i}})}{\partial z} &= \\ - \underbrace{\nabla \cdot (\rho a_i \overline{\mathbf{u}'_i q'_{t,i} \theta'_{l,i}})}_{\text{turb. transport}} - \underbrace{\rho a_i \overline{w'_i q'_{t,i}} \frac{\partial \bar{\theta}_{l,i}}{\partial z}}_{\text{turb. production}} - \underbrace{\rho a_i \overline{w'_i \theta'_{l,i}} \frac{\partial \bar{q}_{t,i}}{\partial z}}_{\text{turb. production}} & \\ + \sum_{j \neq i} \left( \underbrace{\hat{E}_{ij} (\overline{q'_{t,j} \theta'_{l,j}} - \overline{q'_{t,i} \theta'_{l,i}})}_{\text{turb. entrainment}} + \underbrace{\bar{\theta}_{l,i}^* \hat{E}_{ij} (\bar{q}_{t,i} - \bar{q}_{t,j}) + \bar{q}_{t,i}^* \hat{E}_{ij} (\bar{\theta}_{l,i} - \bar{\theta}_{l,j})}_{\text{turb. entrainment production}} \right) & \\ + \sum_{j \neq i} \left( \underbrace{E_{ij} \overline{q'_{t,j} \theta'_{l,j}}}_{\text{dyn. entrainment}} - \underbrace{\Delta_{ij} \overline{q'_{t,i} \theta'_{l,i}}}_{\text{dyn. detrainment}} + \underbrace{E_{ij} (\bar{q}_{t,i} - \bar{q}_{t,j}) (\bar{\theta}_{l,i} - \bar{\theta}_{l,j})}_{\text{dyn. entrainment flux}} \right) & \\ - \underbrace{\rho a_i \overline{D_{q'_{t,i} \theta'_{l,i}}}}_{\text{turb. dissipation}} + \underbrace{\rho a_i (\overline{S'_{q_{t,i} \theta'_{l,i}}} + \overline{S'_{\theta_{l,i} q'_{t,i}}})}_{\text{turb. source correlations}}. & \end{aligned} \quad (2.51)$$

The subdomain covariance equation (2.51) is very general, and current implementations of the extended EDMF scheme simplify it in some ways. The main simplification comes from the assumption that only one of the subdomains, the environment, has significant subgrid-scale variability. In all other subdomains, turbulent correlations are neglected. Then it follows that equation (2.51) is only retained for the environment, and that all fluxes  $\overline{q'_{t,j}\theta'_{l,j}}$  for  $j > 0$  are neglected, yielding the simplified environmental covariance equation

$$\begin{aligned}
& \frac{\partial(\rho a_0 \overline{q'_{t,0}\theta'_{l,0}})}{\partial t} + \nabla_h \cdot (\rho a_0 \langle \mathbf{u}_h \rangle \overline{q'_{t,0}\theta'_{l,0}}) + \frac{\partial(\rho a_0 \overline{w_0 q'_{t,0}\theta'_{l,0}})}{\partial z} = \\
& \quad - \nabla \cdot (\rho a_0 \overline{\mathbf{u}'_0 q'_{t,0}\theta'_{l,0}}) - \rho a_0 \overline{w'_0 q'_{t,0}} \frac{\partial \bar{\theta}_{l,0}}{\partial z} - \rho a_0 \overline{w'_0 \theta'_{l,0}} \frac{\partial \bar{q}_{t,0}}{\partial z} \\
& \quad + \sum_{j>0} \left( - \hat{E}_{0j} \overline{q'_{t,0}\theta'_{l,0}} + \bar{\theta}'_{l,0} \hat{E}_{0j} (\bar{q}_{t,0} - \bar{q}_{t,j}) + \bar{q}'_{t,0} \hat{E}_{0j} (\bar{\theta}_{l,0} - \bar{\theta}_{l,j}) \right) \\
& \quad + \sum_{j>0} \left( - \Delta_{0j} \overline{q'_{t,0}\theta'_{l,0}} + E_{0j} (\bar{q}_{t,0} - \bar{q}_{t,j}) (\bar{\theta}_{l,0} - \bar{\theta}_{l,j}) \right) \\
& \quad - \rho a_0 \overline{D_{q'_{t,0}\theta'_{l,0}}} + \rho a_0 (\overline{S'_{q_{t,0}\theta'_{l,0}}} + \overline{S'_{\theta_{l,0}q'_{t,0}}}). \quad (2.52)
\end{aligned}$$

Here, the turbulent transport and shear production terms in the second line require additional closures. These are typically modeled as diffusive, using expression (2.11), where  $\phi \rightarrow \{q_t \theta_l, q_t, \theta_l\}$  in each of the unresolved terms in equation (2.52).

A distinctly important covariance is the environmental turbulence kinetic energy (TKE)  $\bar{e}_0$ , since it is used in level 2.5 and level 3 schemes of the Mellor-Yamada hierarchy,

$$\bar{e}_0 = \frac{\overline{\mathbf{u}'_0 \cdot \mathbf{u}'_0}}{2} = \frac{1}{2} (\overline{u'_0 u'_0} + \overline{v'_0 v'_0} + \overline{w'_0 w'_0}). \quad (2.53)$$

It is generally used to construct diagnostic models of within-subdomain turbulent

fluxes. Its prognostic equation in flux form reads

$$\begin{aligned}
& \frac{\partial(\rho a_0 \bar{e}_0)}{\partial t} + \nabla_h \cdot (\rho a_0 \langle \mathbf{u}_h \rangle \bar{e}_0) + \frac{\partial(\rho a_0 \bar{w}_0 \bar{e}_0)}{\partial z} = \\
& \underbrace{- \frac{\partial(\rho a_0 \overline{w'_0 e'_0})}{\partial z}}_{\text{vert. turb. transport}} - \underbrace{\nabla_h \cdot (\rho a_0 \overline{\mathbf{u}'_{h,0} e'_0})}_{\text{horiz. turb. transport}} + \underbrace{\rho a_0 \overline{w'_0 b'_0}}_{\text{buoy. production}} + \sum_{i>0} \underbrace{\rho a_i (\bar{w}_i^* - \bar{w}_0^*) \frac{\partial \bar{\Psi}_i^*}{\partial z}}_{\text{pressure - velocity corr.}} \\
& \underbrace{- \rho a_0 \left[ \overline{w'_0 u'_0} \frac{\partial \langle u \rangle}{\partial z} + \overline{w'_0 v'_0} \frac{\partial \langle v \rangle}{\partial z} + \overline{w'_0 w'_0} \frac{\partial \bar{w}_0}{\partial z} \right]}_{\text{vertical shear production}} - \underbrace{\rho a_0 \mathcal{D}}_{\text{turb. dissipation}} \\
& \underbrace{- \rho a_0 (\overline{\mathbf{u}'_{h,0} u'_0} \cdot \nabla_h \langle u \rangle + \overline{\mathbf{u}'_{h,0} v'_0} \cdot \nabla_h \langle v \rangle + \overline{\mathbf{u}'_{h,0} w'_0} \cdot \nabla_h \bar{w}_0)}_{\text{horizontal shear production}} \\
& + \sum_j \left[ \underbrace{E_{0j} \frac{1}{2} (\bar{w}_j - \bar{w}_0)^2}_{\text{conv. energy entrainment}} - \underbrace{(\Delta_{0j} + \hat{E}_{0j}) \bar{e}_0}_{\text{detrainment}} - \underbrace{\hat{E}_{0j} \bar{w}_0^* (\bar{w}_j^* - \bar{w}_0^*)}_{\text{turb. entrainment production}} \right], \quad (2.54)
\end{aligned}$$

where  $\Psi = p/\rho$  is the pressure potential. The environmental TKE is only a fraction of the total turbulence kinetic energy, which also includes the convective energy due to the subdomain velocities  $\bar{w}_i$ . However, no additional equations are necessary for convective TKE, since it is explicitly resolved through the subdomain mean momentum equations. The energetic pathways between environmental TKE and convective TKE are given in Appendix B of Chapter 3.

## 2.5 The full system of equations

The full system of equations necessary to evolve the dynamics of the extended EDMF scheme is formed by:

- The grid-averaged equations (2.21)–(2.26),

$$\frac{\partial \rho}{\partial t} + \nabla_h \cdot (\rho \langle \mathbf{u}_h \rangle) + \frac{\partial(\rho \langle w \rangle)}{\partial z} = 0, \quad (2.55)$$

$$\begin{aligned}
& \frac{\partial(\rho \langle \mathbf{u}_h \rangle)}{\partial t} + \nabla_h \cdot (\rho \langle \mathbf{u}_h \rangle \otimes \langle \mathbf{u}_h \rangle) + \frac{\partial(\rho \langle w \rangle \langle \mathbf{u}_h \rangle)}{\partial z} = \\
& - \nabla_h \cdot (\rho \langle \mathbf{u}_h^* \otimes \mathbf{u}_h^* \rangle) - \frac{\partial(\rho \langle w^* \mathbf{u}_h^* \rangle)}{\partial z} - \nabla_h \langle p^\dagger \rangle_R + \rho S_{\langle \mathbf{u}_h \rangle}, \quad (2.56)
\end{aligned}$$

$$\begin{aligned} \frac{\partial(\rho\langle w \rangle)}{\partial t} + \nabla_h \cdot (\rho\langle \mathbf{u}_h \rangle \langle w \rangle) + \frac{\partial(\rho\langle w \rangle \langle w \rangle)}{\partial z} = \\ - \nabla_h \cdot (\rho\langle \mathbf{u}_h^* w^* \rangle) - \frac{\partial(\rho\langle w^* w^* \rangle)}{\partial z} + \rho\langle b \rangle - \frac{\partial\langle p^\dagger \rangle_R}{\partial z} + \rho S_{\langle w \rangle}, \end{aligned} \quad (2.57)$$

$$\begin{aligned} \frac{\partial(\rho\langle \theta_l \rangle)}{\partial t} + \nabla_h \cdot (\rho\langle \mathbf{u}_h \rangle \langle \theta_l \rangle) + \frac{\partial(\rho\langle w \rangle \langle \theta_l \rangle)}{\partial z} = \\ - \nabla_h \cdot (\rho\langle \mathbf{u}_h^* \theta_l^* \rangle) - \frac{\partial(\rho\langle w^* \theta_l^* \rangle)}{\partial z} + \rho S_{\langle \theta_l \rangle}, \end{aligned} \quad (2.58)$$

$$\begin{aligned} \frac{\partial(\rho\langle q_t \rangle)}{\partial t} + \nabla_h \cdot (\rho\langle \mathbf{u}_h \rangle \langle q_t \rangle) + \frac{\partial(\rho\langle w \rangle \langle q_t \rangle)}{\partial z} = \\ - \nabla_h \cdot (\rho\langle \mathbf{u}_h^* q_t^* \rangle) - \frac{\partial(\rho\langle w^* q_t^* \rangle)}{\partial z} + \rho S_{\langle q_t \rangle}, \end{aligned} \quad (2.59)$$

$$\langle p \rangle_R = \rho R_d \langle T_v \rangle; \quad (2.60)$$

- The subdomain-mean equations (2.40)–(2.45) for  $n$  subdomains, the updrafts,

$$\frac{\partial(\rho a_i)}{\partial t} + \nabla_h \cdot (\rho a_i \langle \mathbf{u}_h \rangle) + \frac{\partial(\rho a_i \bar{w}_i)}{\partial z} = \sum_{j \neq i} (E_{ij} - \Delta_{ij}), \quad (2.61)$$

$$\begin{aligned} \frac{\partial(\rho a_i \bar{w}_i)}{\partial t} + \nabla_h \cdot (\rho a_i \langle \mathbf{u}_h \rangle \bar{w}_i) + \frac{\partial(\rho a_i \bar{w}_i \bar{w}_i)}{\partial z} = -\nabla \cdot (\rho a_i \overline{\mathbf{u}'_i w'_i}) \\ + \sum_{j \neq i} \left[ (E_{ij} + \hat{E}_{ij}) \bar{w}_j - (\Delta_{ij} + \hat{\Delta}_{ij}) \bar{w}_i \right] + \rho a_i \bar{b}_i - \rho a_i \frac{\partial}{\partial z} \left( \frac{\bar{p}_i^\dagger}{\rho} \right) + \rho a_i \bar{S}_{w,i}, \end{aligned} \quad (2.62)$$

$$\begin{aligned} \frac{\partial(\rho a_i \bar{\theta}_{l,i})}{\partial t} + \nabla_h \cdot (\rho a_i \bar{\theta}_{l,i} \langle \mathbf{u}_h \rangle) + \frac{\partial(\rho a_i \bar{\theta}_{l,i} \bar{w}_i)}{\partial z} = -\nabla \cdot (\rho a_i \overline{\theta'_{l,i} \mathbf{u}'_i}) \\ + \sum_{j \neq i} (E_{ij} \bar{\theta}_{l,j} - \Delta_{ij} \bar{\theta}_{l,i}) + \sum_{j \neq i} \hat{E}_{ij} (\bar{\theta}_{l,j} - \bar{\theta}_{l,i}) + \rho a_i \bar{S}_{\theta_{l,i}}, \end{aligned} \quad (2.63)$$

$$\begin{aligned} \frac{\partial(\rho a_i \bar{q}_{t,i})}{\partial t} + \nabla_h \cdot (\rho a_i \bar{q}_{t,i} \langle \mathbf{u}_h \rangle) + \frac{\partial(\rho a_i \bar{q}_{t,i} \bar{w}_i)}{\partial z} = -\nabla \cdot (\rho a_i \overline{q'_{t,i} \mathbf{u}'_i}) \\ + \sum_{j \neq i} (E_{ij} \bar{q}_{t,j} - \Delta_{ij} \bar{q}_{t,i}) + \sum_{j \neq i} \hat{E}_{ij} (\bar{q}_{t,j} - \bar{q}_{t,i}) + \rho a_i \bar{S}_{q_{t,i}}, \end{aligned} \quad (2.64)$$

$$\rho_i = \frac{\langle p \rangle_R}{R_d \bar{T}_{v,i}}. \quad (2.65)$$

The decomposition (2.3) couples all subdomain equations with the grid-averaged equations, such that the equations for one subdomain are defined implicitly. Cohen et al. (2020) take the environmental mean equations to be defined implicitly;

- An equation for the environmental turbulence kinetic energy  $\bar{e}_0$

$$\begin{aligned}
& \frac{\partial(\rho a_0 \bar{e}_0)}{\partial t} + \nabla_h \cdot (\rho a_0 \langle \mathbf{u}_h \rangle \bar{e}_0) + \frac{\partial(\rho a_0 \bar{w}_0 \bar{e}_0)}{\partial z} = - \frac{\partial(\rho a_0 \overline{w'_0 e'_0})}{\partial z} + \rho a_0 \overline{w'_0 b'_0} \\
& - \rho a_0 \left[ \overline{w'_0 u'_0} \frac{\partial \langle u \rangle}{\partial z} + \overline{w'_0 v'_0} \frac{\partial \langle v \rangle}{\partial z} + \overline{w'_0 w'_0} \frac{\partial \bar{w}_0}{\partial z} \right] + \rho \sum_{i>0} a_i (\bar{w}_i^* - \bar{w}_0^*) \frac{\partial \bar{\Psi}_i^*}{\partial z} \\
& + \sum_j \left[ E_{0j} \frac{1}{2} (\bar{w}_j - \bar{w}_0)^2 - (\Delta_{0j} + \hat{E}_{0j}) \bar{e}_0 - \hat{E}_{0j} \bar{w}_0^* (\bar{w}_j^* - \bar{w}_0^*) \right] - \rho a_0 \mathcal{D} \\
& - \nabla_h \cdot (\rho a_0 \overline{\mathbf{u}'_{h,0} e'_0}) - \rho a_0 \left( \overline{\mathbf{u}'_{h,0} u'_0} \cdot \nabla_h \langle u \rangle + \overline{\mathbf{u}'_{h,0} v'_0} \cdot \nabla_h \langle v \rangle + \overline{\mathbf{u}'_{h,0} w'_0} \cdot \nabla_h \bar{w}_0 \right); \tag{2.66}
\end{aligned}$$

- Equations for the environmental covariances  $\overline{\theta'_{l,0} q'_{l,0}}$ ,  $\overline{q'^2_{l,0}}$  and  $\overline{\theta'^2_{l,0}}$ ,

$$\begin{aligned}
& \frac{\partial(\rho a_0 \overline{\theta'_{l,0} q'_{l,0}})}{\partial t} + \nabla_h \cdot (\rho a_0 \langle \mathbf{u}_h \rangle \overline{\theta'_{l,0} q'_{l,0}}) + \frac{\partial(\rho a_0 \bar{w}_0 \overline{\theta'_{l,0} q'_{l,0}})}{\partial z} = \\
& - \nabla \cdot (\rho a_0 \overline{\mathbf{u}'_0 \theta'_{l,0} q'_{l,0}}) - \rho a_0 \overline{w'_0 q'_{l,0}} \frac{\partial \bar{\theta}_{l,0}}{\partial z} - \rho a_0 \overline{w'_0 \theta'_{l,0}} \frac{\partial \bar{q}_{l,0}}{\partial z} \\
& + \sum_{j>0} \left( -\hat{E}_{0j} \overline{\theta'_{l,0} q'_{l,0}} + \bar{\theta}_{l,0}^* \hat{E}_{0j} (\bar{q}_{l,0} - \bar{q}_{l,j}) + \bar{q}_{l,0}^* \hat{E}_{0j} (\bar{\theta}_{l,0} - \bar{\theta}_{l,j}) \right) \\
& + \sum_{j>0} \left( -\Delta_{0j} \overline{\theta'_{l,0} q'_{l,0}} + E_{0j} (\bar{q}_{l,0} - \bar{q}_{l,j}) (\bar{\theta}_{l,0} - \bar{\theta}_{l,j}) \right) \\
& - \rho a_0 \overline{D_{\theta'_{l,0} q'_{l,0}}} + \rho a_0 (\overline{S'_{\theta_{l,0} q'_{l,0}}} + \overline{S'_{\theta_{l,0} q'_{l,0}}}), \tag{2.67}
\end{aligned}$$

$$\begin{aligned}
& \frac{\partial(\rho a_0 \overline{q'^2_{l,0}})}{\partial t} + \nabla_h \cdot (\rho a_0 \langle \mathbf{u}_h \rangle \overline{q'^2_{l,0}}) + \frac{\partial(\rho a_0 \bar{w}_0 \overline{q'^2_{l,0}})}{\partial z} = - \nabla \cdot (\rho a_0 \overline{\mathbf{u}'_0 q'^2_{l,0}}) \\
& - 2 \rho a_0 \overline{w'_0 q'^2_{l,0}} \frac{\partial \bar{q}_{l,0}}{\partial z} + \sum_{j>0} \left( -\hat{E}_{0j} \overline{q'^2_{l,0}} + 2 \bar{q}_{l,0}^* \hat{E}_{0j} (\bar{q}_{l,0} - \bar{q}_{l,j}) \right) \\
& + \sum_{j>0} \left( -\Delta_{0j} \overline{q'^2_{l,0}} + E_{0j} (\bar{q}_{l,0} - \bar{q}_{l,j})^2 \right) - \rho a_0 \overline{D_{q'^2_{l,0}}} + 2 \rho a_0 \overline{S'_{q'_{l,0} q'_{l,0}}}, \tag{2.68}
\end{aligned}$$

$$\begin{aligned}
& \frac{\partial(\rho a_0 \overline{\theta'_{l,0}{}^2})}{\partial t} + \nabla_h \cdot (\rho a_0 \langle \mathbf{u}_h \rangle \overline{\theta'_{l,0}{}^2}) + \frac{\partial(\rho a_0 \overline{w_0 \theta'_{l,0}{}^2})}{\partial z} = -\nabla \cdot (\rho a_0 \overline{\mathbf{u}'_0 \theta'_{l,0}{}^2}) \\
& - 2\rho a_0 \overline{w'_0 \theta'_{l,0}} \frac{\partial \bar{\theta}_{l,0}}{\partial z} + \sum_{j>0} \left( -\hat{E}_{0j} \overline{\theta'_{l,0}{}^2} + 2\bar{\theta}_{l,0}^* \hat{E}_{0j} (\bar{\theta}_{l,0} - \bar{\theta}_{l,j}) \right) \\
& + \sum_{j>0} \left( -\Delta_{0j} \overline{\theta'_{l,0}{}^2} + E_{0j} (\bar{\theta}_{l,0} - \bar{\theta}_{l,j})^2 \right) - \rho a_0 \overline{D_{\theta'_{l,0}{}^2}} + 2\rho a_0 \overline{S'_{\theta_{l,0}} \theta'_{l,0}}. \quad (2.69)
\end{aligned}$$

In total, evolving the extended EDMF scheme in time requires integrating  $10 + 4n$  prognostic equations. This number can be reduced to  $7 + 4n$  if an environmental model of level 2.5 is selected instead, by neglecting transience in the thermodynamic covariance equations.

## 2.6 Similarities between the extended EDMF scheme and other approaches

The extended EDMF scheme is one of many candidate frameworks developed to unify parameterizations of low tropospheric mixing in Earth System Models. Here we briefly describe its similarities and differences with other recently developed schemes in the literature.

### The multi-fluid scheme

Parallel to the development of the extended EDMF scheme (Tan et al., 2018; Cohen et al., 2020; Lopez-Gomez et al., 2020), a similar framework known as the multi-fluid scheme has been developed (Thuburn et al., 2018; Weller and McIntyre, 2019; Shipley et al., 2022; Thuburn et al., 2022a). Both frameworks are derived from spatial averaging of the governing equations over free moving subdomains; a process that can also be described as conditional filtering (Thuburn et al., 2022a). By considering subdomains with moving boundaries, they generalize traditional EDMF schemes by allowing updrafts to expand and contract as a function of the forcing. Retaining transience in the subdomain volume equations endows the extended EDMF and multi-fluid schemes with convective memory.

One important difference between the extended EDMF scheme and the multi-fluid scheme is that the latter remains fully compressible; the subdomain density is retained in the subdomain continuity equations. In this case, the continuity equation for area fraction (2.61) is replaced by an equation for the subdomain mass fraction  $m_i = \rho_i a_i$ , which is then used in the rest of subdomain equations (Thuburn et al.,

2022b). The area fraction becomes a diagnostic quantity  $a_i = m_i/\rho_i$ , given the subdomain density  $\rho_i = \rho_i(\bar{\theta}_{l,i}, \bar{q}_{t,i})$ .

Another generalization of the multi-fluid scheme is its retention of the subdomain covariances within the updrafts. This is also done in the derivation of the extended EDMF scheme in Section 2.4, but practical implementations do neglect them (Cohen et al., 2020; Lopez-Gomez et al., 2020). Retaining second-order moments within the updrafts increases the number of equations of the scheme to  $10 + 8n$ . Due to the increased computational cost, (Thuburn et al., 2022b) truncate the scheme at level 2.5 in the Mellor-Yamada hierarchy; neglecting transience, advection and third-order terms in all thermodynamic covariances. This reduces the number of prognostic equations to  $7 + 5n$ . Whether retaining diagnostic second-order moments for all subdomains or prognostic ones for the environment leads to improved performance remains to be investigated.

All other differences relate to the closure of the entrainment fluxes (2.36), and the treatment of pressure-velocity correlations and third-order terms appearing in the covariance equations (2.66)–(2.69). Regarding dynamical entrainment, Thuburn et al. (2022b) propose

$$-\frac{A_i^{sg}}{V_T}(\rho\widehat{\phi u_{r,n}})_i = \sum_{j \neq i} (E_{ij}\hat{\phi}_{ji} - \Delta_{ij}\hat{\phi}_{ij}), \quad (2.70)$$

where the entrained quantity is a linear combination of the averages over the subdomains involved in the mass exchange,

$$\hat{\phi}_{ji} = \alpha\bar{\phi}_j + (1 - \alpha)\bar{\phi}_i. \quad (2.71)$$

This term reduces to the extended EDMF closure (2.37) for  $\alpha = 1$ . The mixing ratio  $\alpha$  is treated as a learnable parameter and diagnosed from large-eddy simulations in McIntyre et al. (2022). However, the authors note that adjustments to the diagnosed values of  $\alpha$  are necessary when the multi-fluid system is used prognostically, for velocities in the boundary layer to remain realistic.

The mass entrainment and detrainment fluxes  $E_{ij}$  and  $\Delta_{ij}$  are parameterized as the sum of contributions from 4 processes: entrainment by static instability, due to relaxation to a pre-defined reference, due to turbulent mixing, and finally forced detrainment of negatively buoyant air at the boundary layer top.

In addition, Thuburn et al. (2022a) introduce the turbulent entrainment closure

$$-\frac{A_i^{sg}}{V_T}(\rho\widehat{\phi' u'_{r,n}})_i = \frac{1}{\rho} \left[ \sum_{j \neq i} m_j \nabla \cdot (m_i \overline{\mathbf{u}'_i \phi'_i}) - m_i \sum_{j \neq i} \nabla \cdot (m_j \overline{\mathbf{u}'_j \phi'_j}) \right]. \quad (2.72)$$

This turbulent entrainment closure homogenizes contributions from SGS fluxes due to area fraction gradients across subdomains. The authors show this term to be necessary if one wants to enforce that subdomains with identical thermodynamic states and different area fractions experience similar turbulent tendencies. In the extended EDMF scheme, where updraft turbulent fluxes are neglected, this term would reduce to

$$-\frac{A_i^{sg}}{V_T}(\rho\widehat{\phi'u'_{r,n}})_i = -a_i\nabla \cdot (\rho a_0\overline{\mathbf{u}'_0\phi'_0}), \quad i > 0, \quad (2.73)$$

and in the environment,

$$-\frac{A_0^{sg}}{V_T}(\rho\widehat{\phi'u'_{r,n}})_0 = \sum_{j=0} a_j\nabla \cdot (\rho a_0\overline{\mathbf{u}'_0\phi'_0}). \quad (2.74)$$

This turbulent entrainment closure could be combined with the diffusive closure (2.39). Similar turbulent entrainment closures may be found for the covariance equations, by homogenizing the area fraction contribution from second and third order terms. Their expressions are derived in Section 6 of Thuburn et al. (2022a).

The implemented versions of the extended EDMF scheme and the multi-fluid framework use different closures for all unresolved terms, and for sources related to microphysical sources. However, these are easily interchangeable when the differences discussed above are taken into account. The multi-fluid model has been successfully used to simulate a range of canonical dynamical regimes, including dry convection (Thuburn et al., 2019), Rayleigh–Bénard convection (Shiple et al., 2022), and shallow convection (McIntyre et al., 2022). Results of the scheme under other dynamical regimes such as stratocumulus-topped boundary layers, stable boundary layers, or deep convection, have not been published as of this writing.

### **The steady-state multiplume EDMF scheme: JPL-EDMF and ED(MF)<sup>n</sup>**

The extended EDMF and multi-fluid schemes leverage a time and vertically varying partition of the grid-mean flow to regulate the intensity of convective overturning and retain convective memory. Although these schemes allow for an arbitrary number of updrafts, using more than one updraft in the extended EDMF scheme does not lead to significant performance improvements in practice – rendering the linear increase in number of equations unjustifiable (Cohen et al., 2020; Lopez-Gomez et al., 2020). This also seems to be the case for the multi-fluid approach, since numerical implementations in the literature use a single updraft (Thuburn et al., 2022b; McIntyre et al., 2022).

In contrast, research efforts focused on the steady-state EDMF scheme discussed in Section 2.3 have evolved to include a significantly higher number of updrafts. Here, we focus our attention on two models: the JPL-EDMF (Suselj et al., 2013; Suselj et al., 2019a) and the ED(MF)<sup>n</sup> scheme (Neggers, 2015). The computational cost of evaluating these models scales more favorably with the number of updrafts, since the subdomain equations (2.61)–(2.64) are substituted by one-dimensional ordinary differential equations for the vertical velocity (Eq. 2.9), and for all thermodynamic fields,

$$\frac{\partial \bar{\phi}_i}{\partial z} = \epsilon_{i0}(\bar{\phi}_0 - \bar{\phi}_i) + \frac{S_{\bar{\phi}_i}}{\bar{w}_i}. \quad (2.75)$$

In equation (2.75),  $\epsilon_{i0} = E_{i0}/(\rho a_i \bar{w}_i)$  is the fractional dynamical entrainment rate of environmental air. The mass flux (2.6) is then defined by an area fraction profile that is fixed a priori. Scalability is improved even further in some implementations, where a single updraft initialized at the surface is split into multiple updrafts above its lifting condensation level (Suselj et al., 2013). This splitting represents a trigger function much like the ones used in early mass flux schemes (Section 2.2), so most state-of-the-art EDMF schemes now initialize all updrafts at the surface (Suselj et al., 2019b).

The excellent performance of these schemes in standard benchmarks (Suselj et al., 2019a; Suselj et al., 2022; Neggers and Griewank, 2022) suggests that increasing the number of updrafts can partially compensate for structural errors due to the simplified description of their dynamics given by equations (2.9) and (2.75). To look for compensation mechanisms in multiplume schemes, it is useful to first distill the approximations in their dynamics that are absent in the extended EDMF scheme.

One significant simplification is the approximation of the updraft area fraction as a step function,

$$a_i = \begin{cases} a & \text{if } \bar{w}_i > 0, \\ 0 & \text{otherwise,} \end{cases} \quad i = 1, \dots, n, \quad (2.76)$$

where  $a$  is a user-defined constant (Neggers, 2015; Suselj et al., 2019a). Apart from forced detrainment of the entire area fraction when updraft velocities become negative, mass exchange processes do not modify the updraft area fraction. This constraint makes it impossible for steady-state EDMF schemes with a single updraft to reproduce the vertical structure of convective clouds, characterized by a varying cloud fraction with several maxima and minima (Singer et al., 2021). The use of multiple updrafts with different lifting condensation levels and termination heights

can compensate for this (Neggers, 2015; Neggers and Griewank, 2022), since the total updraft area fraction then becomes

$$a_u = \sum_{i>0} a_i H(-\bar{w}_i), \quad (2.77)$$

where  $H(\cdot)$  is the Heaviside step function. This can be seen in Figure 2.2. The vertical smoothness of the area fraction (2.77) is a function of the number of updrafts; a typical value found in the literature is  $n \sim 10$ . Although multiplume schemes with fixed area fraction allow for vertical variations in cloud fraction, below cloud base the total updraft area fraction is strictly non-increasing, which can be detrimental to model performance. Neggers and Griewank (2022) address this issue by introducing an empirical model of  $a_i$  within the mixed-layer that increases linearly with height and depends on updraft size. They show that this model reduces the misfit of their scheme with respect to resolved simulations of shallow cumulus convection.

Fixing the updraft area fraction also defines the fractional detrainment rate  $\delta_{i0} = D_{i0}/(\rho a_i \bar{w}_i)$ , given a closure for entrainment. Comparing the implied net entrainment rate to closures used in unsteady schemes can provide insight into potential models for area fraction in multiplume steady EDMF schemes. From the steady-state mass flux equation (2.7),

$$\epsilon_{i0} - \delta_{i0} = \frac{\partial}{\partial z} [\ln(\rho \bar{w}_i)], \quad \bar{w}_i > 0. \quad (2.78)$$

There is net entrainment in regions of accelerating ascent, and net detrainment in regions of updraft deceleration. For a given stratification, the net entrainment/detrainment can be written in terms of the wavenumber  $\bar{b}_i/\bar{w}_i^2$  if equation (2.9) is used for the vertical velocity,

$$\epsilon_{i0} - \delta_{i0} = \frac{\partial}{\partial z} [\ln(\rho)] + c_1 \frac{\bar{b}_i}{\bar{w}_i^2} - c_2 \epsilon_{i0}, \quad \bar{w}_i > 0. \quad (2.79)$$

Cohen et al. (2020) show improved performance of the extended EDMF scheme when the difference  $\epsilon_{i0} - \delta_{i0}$  is a function of the relative humidity difference between updrafts and the environment, as previously suggested by the idealized theoretical analysis of Savre and Herzog (2019). This dependence could be introduced in parameterizations of the area fraction in steady EDMF schemes.

Finally, two additional simplifications separate steady multiplume models from the extended EDMF scheme: the absence of horizontal advection and transient terms in

the subdomain equations. Retaining transient terms represents a form of convective memory (Cohen et al., 2020). In the context of the ED(MF)<sup>n</sup> scheme, Neggers and Griewank (2022) present an alternative approach to retain history dependence. They couple their multiplume EDMF scheme to a thermal population model, based on two-dimensional cellular automata (Wolfram, 1983), that regulates the area fraction occupied by different populations with time. A benefit of this model is that it provides a scale-aware mechanism to parameterize convective aggregation, a phenomenon that is difficult to encode using a single-plume extended EDMF (or multi-fluid) scheme (Gentine et al., 2018).

Whether transient schemes such as the extended EDMF or steady-state multiplume schemes like ED(MF)<sup>n</sup> are the best choice for implementation in global models remains to be seen. Transient schemes are by construction more complete models, but the outcome of a comparison is far from evident when the implementation is conditioned by a finite computational budget. A systematic comparison of the hierarchy of schemes, given pre-defined spatial and time resolutions, would provide insight into this question.

## 2.7 Discussion

Aggregate models of atmospheric turbulence and convection have undergone substantial change since the early work of Manabe in the 1960s. The mass flux schemes that had their roots in insight gained from observational campaigns (Yanai et al., 1973; Arakawa and Schubert, 1974) have been integrated into frameworks that unify the description of turbulence, dry, shallow, and deep convection (Siebesma and Teixeira, 2000; Siebesma et al., 2007; Soares et al., 2004; Suselj et al., 2019a). Transient versions of these schemes consistent with a systematic filtering of the governing equations of motion have recently been derived (Tan et al., 2018; Cohen et al., 2020; Thuburn et al., 2022a), providing even more flexible and seamless frameworks to parameterize unresolved processes that remain consistent across scales.

Most of these advances, however, have not been implemented in operational models. Climate models in CMIP6 still largely employ separate schemes for boundary layer turbulence, shallow convection, and deep convection (Stevens et al., 2013; T. Wu et al., 2019; Walters et al., 2019; Madeleine et al., 2020). In the best of cases, the parameterizations of turbulence and shallow convection have been unified, but remain independent from deep convection closures (Danabasoglu et al., 2020). The added cost of implementing new parameterizations and making them efficient and

stable, the laborious process of model calibration (Mauritsen et al., 2012; Hourdin et al., 2017), the requirement to produce an ever-increasing number of climate simulations for intercomparison projects (Eyring et al., 2016; Haarsma et al., 2016), and the limiting timescale of the CMIP cycle, are all reasons that may explain why this is the case. The gap between recently developed turbulence and convection schemes and those implemented operationally is also apparent in numerical weather prediction (Cottrill et al., 2013; Han et al., 2017; Takaya et al., 2018).

This gap often leads to mistrust in the turbulence and convection schemes of today, based on the biases introduced by older, now operational schemes (Gentine et al., 2018). Although advances in data-driven parameterization schemes in recent years are promising (Brenowitz and Bretherton, 2018; Brenowitz et al., 2020; Yuval and O’Gorman, 2020; Yuval et al., 2021; X. Wang et al., 2022), biases will remain if the gap between operational and state-of-the-art models is not reduced. A concerted effort toward the adoption of unified schemes that have extensively shown improved skill over a wide range of atmospheric regimes in single-column settings (Suselj et al., 2019a; Cohen et al., 2020; Lopez-Gomez et al., 2020) and global simulations (Hogan et al., 2014; Suselj et al., 2014; Suselj et al., 2021) would represent a worthy step forward in climate modeling.

The derivation of the extended EDMF scheme, its historical context, and its similarities and differences with closely related turbulence and convection schemes have been discussed in this chapter. As stated in Section 2.4, none of these schemes provide a closed system of equations. They constitute sets of prognostic equations that enable a description of complex atmospheric dynamics in terms of more elementary processes that still need to be modeled; with the added benefit that conservation laws are engrained in the schemes. Examples of remaining unresolved processes are turbulent mixing and entrainment. For some of these processes, closures may be found by making additional approximations about the leading order dynamics governing them. One such example is the closure of turbulent diffusion within the quasi-isotropic environment surrounding convective updrafts; this closure is the topic of the next chapter.

*Chapter 3***A GENERALIZED MIXING LENGTH CLOSURE FOR  
EDDY-DIFFUSIVITY MASS-FLUX SCHEMES OF  
TURBULENCE AND CONVECTION**

Lopez-Gomez, Ignacio, Yair Cohen, Jia He, Anna Jaruga, and Tapio Schneider (2020). “A generalized mixing length closure for eddy-diffusivity mass-flux schemes of turbulence and convection”. In: *Journal of Advances in Modeling Earth Systems* 12, e2020MS002161. DOI: 10.1029/2020MS002161.

**3.1 Abstract**

Because of their limited spatial resolution, numerical weather prediction and climate models have to rely on parameterizations to represent atmospheric turbulence and convection. Historically, largely independent approaches have been used to represent boundary layer turbulence and convection, neglecting important interactions at the subgrid scale. Here we build on an eddy-diffusivity mass-flux (EDMF) scheme that represents all subgrid-scale mixing in a unified manner, partitioning subgrid-scale fluctuations into contributions from local diffusive mixing and coherent advective structures and allowing them to interact within a single framework. The EDMF scheme requires closures for the interaction between the turbulent environment and the plumes and for local mixing. A second-order equation for turbulence kinetic energy (TKE) provides one ingredient for the diffusive local mixing closure, leaving a mixing length to be parameterized. Here, we propose a new mixing length formulation, based on constraints derived from the TKE balance. It expresses local mixing in terms of the same physical processes in all regimes of boundary layer flow. The formulation is tested at a range of resolutions and across a wide range of boundary layer regimes, including a stably stratified boundary layer, a stratocumulus-topped marine boundary layer, and dry convection. Comparison with large eddy simulations (LES) shows that the EDMF scheme with this diffusive mixing parameterization accurately captures the structure of the boundary layer and clouds in all cases considered.

### 3.2 Introduction

Turbulence is ubiquitous in the planetary boundary layer. Small-scale chaotic air motions enhance mixing of energy and moisture in the lower troposphere. Under statically unstable conditions, convective updrafts and downdrafts further increase the vertical transport of energy and moisture between the surface and the air aloft. Together, turbulence and convection shape the vertical distribution of temperature and water vapor that sustains clouds. However, these processes act on scales far too small to be resolved in global climate models (GCMs), with resolutions constrained by current computational power (Schneider et al., 2017a). Although the unabated increase in processing power will make globally resolving deep convective processes routine in the coming years (Kajikawa et al., 2016), resolving turbulent mixing and shallow convection will remain an intractable problem for decades. Instead, parameterizations have to be used to approximate the average effect of these subgrid-scale processes on the grid scale.

Conventional parameterizations consider atmospheric turbulence and convection as independent processes, neglecting interactions that alter their combined effect on the large scale. These parameterizations are often regime-dependent, leading to models that artificially split the spectrum of atmospheric conditions into a discrete number of cases. Examples of such case-dependent approaches include parameterizations of cumulus (Arakawa, 2004) and stratocumulus clouds (Lilly, 1968; Schubert, 1976). However accurate, the use of disparate schemes for different conditions complicates a seamless representation of subgrid-scale processes in the lower troposphere.

Several approaches to obtain a unified model of turbulence and convection have been proposed (Lappen and Randall, 2001; Park, 2014; Thuburn et al., 2018). Here we focus on the extended formulation of an eddy-diffusivity mass-flux (EDMF) scheme developed in Tan et al. (2018), which in turn built on work by Siebesma and Teixeira (2000), Soares et al. (2004), and Siebesma et al. (2007) and Angevine et al. (2010), among others. In the EDMF framework, the flow within each grid cell is decomposed into several distinct subdomains, representing coherent convective structures and their relatively isotropic turbulent environment. Convective transport is captured by mass flux terms that depend on differences between subdomain-mean properties; more isotropic turbulent transport, associated with small-scale fluctuations within each subdomain, is captured by eddy diffusion closures.

The extended EDMF framework uses additional prognostic equations for subdomain variables, such as the environmental turbulence kinetic energy, and it requires clo-

asures for local turbulent fluxes and for the mass exchange between subdomains (Tan et al., 2018). Even though the EDMF framework arises from the need for a unified model of turbulence and convection, the parameterizations used for entrainment and turbulent mixing are usually defined differently for each regime (Suselj et al., 2013; Witek et al., 2011b). The development of regime-independent parameterizations for the required closures is the last step in the construction of a unified model of atmospheric turbulence and convection.

Here, a regime-independent closure for turbulent mixing within the EDMF framework is proposed. Section 2 reviews the decomposition of subgrid-scale fluxes in the extended EDMF scheme. Section 3 introduces the formulation of the closure. Section 4 illustrates the performance of the EDMF scheme with the turbulent mixing closure in boundary layer regimes where vertical transport is strongly dependent on the turbulence closure used: the stable boundary layer (SBL), the stratocumulus-topped boundary layer (STBL), and dry convection. The performance of the extended EDMF scheme with this closure in moist-convective cases is demonstrated in a companion paper (Cohen et al., 2020). Finally, Section 5 summarizes the results and conclusions.

### 3.3 EDMF framework

In the EDMF framework, each grid-cell volume is decomposed into  $n$  updrafts or downdrafts (labeled by index  $i = 1, \dots, n$ ) and an environment (labeled by index  $i = 0$ ) in which they are embedded. Following this decomposition, the grid-mean value of variable  $\psi$  may be written as

$$\langle \psi \rangle = \sum_{i \geq 0}^n a_i \bar{\psi}_i. \quad (3.1)$$

Here, angle brackets  $\langle \cdot \rangle$  denote the grid mean,  $\bar{\psi}_i$  denotes the Favre average of  $\psi$  over subdomain  $i$ , and  $a_i$  is the mean horizontal cross-sectional area covered by subdomain  $i$  within the grid cell. This partition is motivated by the anisotropy of turbulent convective flows, in which isotropic turbulent eddies coexist with coherent columnar structures that induce a strong vertical transport (Bjerknes, 1938). The subdomain decomposition is simplified for the horizontal velocity vector  $\mathbf{u}_h$ , which is taken to have the same mean value for all subdomains,  $\bar{\mathbf{u}}_{h,i} = \langle \mathbf{u}_h \rangle$ . Applying the subdomain decomposition to higher-order moments introduces additional terms associated with the difference between grid and subdomain means. For the vertical

subgrid-scale flux of  $\psi$ , this leads to

$$\langle w^* \psi^* \rangle = \sum_{i \geq 0}^n a_i \left( \overline{w'_i \psi'_i} + \bar{w}_i^* \bar{\psi}_i^* \right). \quad (3.2)$$

Here,  $w$  is the vertical velocity,  $\psi^* = \psi - \langle \psi \rangle$ ,  $\psi'_i = \psi - \bar{\psi}_i$ , and  $\bar{\psi}_i^* = \bar{\psi}_i - \langle \psi \rangle$ . The decomposition (3.2) partitions the subgrid-scale flux into contributions from small-scale fluctuations, associated with turbulence, and subdomain-mean terms, representative of convection. In the following, we will refer to these contributions as turbulent and convective fluxes, respectively.

The subdomain-mean terms can be explicitly solved for by introducing  $n$  prognostic subdomain equations for each variable and an additional equation for each plume area fraction  $a_i$ , which may be diagnostic or prognostic. Cohen et al. (2020) derive the subdomain equations used in the EDMF framework, starting from the Navier-Stokes equations. The use of prognostic subdomain equations means that convective fluxes such as  $\bar{w}_i^* \bar{\psi}_i^*$  in (3.2) are explicitly solved for, while turbulent fluxes like  $\overline{w'_i \psi'_i}$  must be modeled. Turbulent fluxes within each subdomain are modeled as downgradient and proportional to an eddy diffusivity  $K_{\psi,i}$ , where  $\psi$  is the property being transported. For the vertical turbulent flux in (3.2), this gives

$$\overline{w'_i \psi'_i} = -K_{\psi,i} \frac{\partial \bar{\psi}_i}{\partial z}. \quad (3.3)$$

The eddy diffusivity  $K_{\psi,i}$  is proportional to a characteristic velocity scale and the length scale of the eddies driving the transport, both of which must be parameterized.

Proposed closures for the eddy diffusivity vary from simple diagnostic expressions to second-order models that introduce prognostic equations for both scales (Umlauf and Burchard, 2003). The 1.5-order turbulence kinetic energy (TKE) model is a particularly popular choice due to its balance between accuracy and computational efficiency (Mellor and Yamada, 1982). The 1.5-order model, also referred to as the Level 2.5 model in the Mellor-Yamada hierarchy, makes use of a prognostic equation for TKE and a diagnostic expression for the mixing length. In the EDMF framework, the grid-mean TKE  $\langle e \rangle$  can be decomposed following expression (3.2) for second-order moments as

$$\langle e \rangle = \sum_{i \geq 0}^n a_i \left( \bar{e}_i + \frac{\bar{w}_i^* \bar{w}_i^*}{2} \right), \quad (3.4)$$

where  $\bar{e}_i$  is the TKE of subdomain  $i$ , and the second term represents the corresponding convective kinetic energy. This expression can be simplified by assuming that

for the updrafts and downdrafts ( $i > 0$ ), the contribution to the grid-mean TKE from small-scale turbulence is negligible compared to the convective term, an assumption commonly made in EDMF schemes:

$$\langle e \rangle = a_0 \bar{e}_0 + \sum_{i \geq 0}^n a_i \frac{\bar{w}_i^* \bar{w}_i^*}{2}. \quad (3.5)$$

Thus, grid-mean TKE is given by the sum of environmental TKE and convective TKE. The TKE decomposition (3.5) can also be obtained by assuming a small updraft and downdraft area fraction and similar turbulence intensity in all subdomains (Siebesma et al., 2007). However, the equations derived for the subdomain second-order moments with these two approaches differ in the source terms that appear due to entrainment processes between subdomains. The former approximation is favored here to allow for the use of this framework in high-resolution models, where the assumption of slender updrafts may become inadequate (Randall, 2013).

Given an updraft area fraction  $a_i$ , which may be diagnostic or prognostic (Tan et al., 2018), the grid-mean TKE is determined by the environmental TKE  $\bar{e}_0$  and the subdomain-mean vertical velocities  $\bar{w}_i$ . The subdomain-mean vertical velocity equation for subdomain  $i$  is

$$\begin{aligned} \frac{\partial(\rho a_i \bar{w}_i)}{\partial t} + \frac{\partial(\rho a_i \bar{w}_i^2)}{\partial z} + \nabla_h \cdot (\rho a_i \bar{\mathbf{u}}_{h,i} \bar{w}_i) = & -\frac{\partial(\rho a_i \overline{w'_i w'_i})}{\partial z} - \nabla_h \cdot (\rho a_i \overline{\mathbf{u}'_{h,i} w'_i}) \\ & + \sum_{j \neq i} [E_{ij} \bar{w}_j - \Delta_{ij} \bar{w}_i + \hat{E}_{ij} (\bar{w}_j - \bar{w}_i)] + \rho a_i \bar{b}_i - \rho a_i \frac{\partial \bar{\Psi}_i^\dagger}{\partial z}, \end{aligned} \quad (3.6)$$

where  $\nabla_h$  is the horizontal gradient operator,  $\Psi = p/\rho$  is the pressure potential and the turbulent transport terms on the right-hand side are negligible for all subdomains except the environment ( $i = 0$ ). Subgrid density changes are only considered in the buoyancy term, such that  $\rho = \langle \rho \rangle$  in the previous equation, in order to avoid creation of spurious acoustic modes through the subdomain decomposition (Cohen et al., 2020). The buoyancy  $\bar{b}_i$  and the pressure potential anomaly  $\bar{\Psi}_i^\dagger$  are defined with respect to a reference hydrostatic pressure profile  $p_h(z)$  and density  $\rho_h(z)$ , related by  $\partial_z p_h = -\rho_h g$ :

$$\bar{b}_i = -g \frac{\bar{p}_i - \rho_h}{\rho}, \quad \frac{\partial \bar{\Psi}_i^\dagger}{\partial z} = \frac{\partial}{\partial z} \left( \frac{\bar{p}_i}{\rho} \right) + g \frac{\rho_h}{\rho}. \quad (3.7)$$

Here,  $\bar{p}_i$  is the subdomain-mean pressure. Density appears inside the pressure gradients in (3.6) and (3.7) to ensure thermodynamic consistency of the subgrid-scale anelastic approximation (Cohen et al., 2020). Interactions between subdomains

are captured by entrainment and detrainment fluxes. In the vertical velocity equation (3.6),  $\Delta_{ij}$  is the dynamical detrainment of air mass from subdomain  $i$  into subdomain  $j$ , and  $E_{ij}$  and  $\hat{E}_{ij}$  are the dynamical and turbulent entrainment from subdomain  $j$  into subdomain  $i$ , respectively. It is assumed that entrainment events occur over timescales much shorter than the eddy turnover rate  $K_{\psi,i}/\bar{e}_i$ , so that entrained air carries the properties of the subdomain it detrains from. In addition, for now we assume entrainment occurs only between convective plumes and the environment, not among plumes.

The prognostic equation for environmental TKE can be written in non-conservative form as (Cohen et al., 2020)

$$\begin{aligned} \frac{\partial \bar{e}_0}{\partial t} + \bar{w}_0 \frac{\partial \bar{e}_0}{\partial z} + \langle \mathbf{u}_h \rangle \cdot \nabla_h \bar{e}_0 = & -\overline{w'_0 u'_0} \frac{\partial \langle u \rangle}{\partial z} - \overline{w'_0 v'_0} \frac{\partial \langle v \rangle}{\partial z} - \overline{w_0'^2} \frac{\partial \bar{w}_0}{\partial z} + \overline{w'_0 b'_0} - \mathcal{P} \\ - \frac{1}{\rho a_0} \frac{\partial}{\partial z} \left( \rho a_0 \overline{w'_0 e'_0} \right) + \sum_{i>0} \left[ \frac{\Delta_{i0}}{\rho a_0} \left( \frac{(\bar{w}_i - \bar{w}_0)^2}{2} - \bar{e}_0 \right) - \frac{\hat{E}_{i0}}{\rho a_0} (\bar{w}_0^* (\bar{w}_i - \bar{w}_0) + \bar{e}_0) \right] - \mathcal{D} \\ - \frac{1}{\rho a_0} \nabla_h \cdot \left( \rho a_0 \overline{\mathbf{u}'_{h,0} e'_0} \right) - \overline{\mathbf{u}'_{h,0} u'_0} \cdot \nabla_h \langle u \rangle - \overline{\mathbf{u}'_{h,0} v'_0} \cdot \nabla_h \langle v \rangle - \overline{\mathbf{u}'_{h,0} w'_0} \cdot \nabla_h \bar{w}_0. \end{aligned} \quad (3.8)$$

Here,  $\langle u \rangle$  and  $\langle v \rangle$  are the components of  $\langle \mathbf{u}_h \rangle$ ,  $\mathcal{P}$  is the velocity-pressure gradient correlation, and  $\mathcal{D}$  is the turbulent dissipation. All sources and sinks of  $\bar{e}_0$  account for unresolved processes on the grid scale, so they must be parameterized. Subdomain covariances in (3.8) are modeled diffusively, with the environmental eddy diffusivity  $K_\psi$  defined as

$$K_\psi = c_\psi l \bar{e}_0^{1/2}, \quad (3.9)$$

where  $l$  is the mixing length, and  $c_\psi$  is a fitting parameter. The subscript 0 in the eddy diffusivity is dropped to simplify notation. The coefficient  $c_\psi$  is taken to be equal to  $c_h$  for the diffusion of all fields except for momentum, for which  $c_\psi = c_m$ . The eddy viscosity  $K_m$  is related to  $K_h$  through the turbulent Prandtl number  $\text{Pr}_t$ , such that  $K_m = \text{Pr}_t K_h$ .

Under the assumption that subgrid-scale pressure work on the grid mean is negligible,  $\mathcal{P}$  is taken as opposite to the pressure work on the plumes (Tan et al., 2018),

$$\mathcal{P} = \left[ \overline{w'_0 \left( \frac{\partial \Psi}{\partial z} \right)'_0} + \overline{u'_0 \left( \frac{\partial \Psi}{\partial x} \right)'_0} + \overline{v'_0 \left( \frac{\partial \Psi}{\partial y} \right)'_0} \right] = - \sum_{i>0} \frac{a_i}{a_0} (\bar{w}_i^* - \bar{w}_0^*) \frac{\partial \bar{\Psi}_i^*}{\partial z}. \quad (3.10)$$

The last term in (3.10) appears as a sink term in the convective TKE balance, which is derived in 3.8. Hence,  $\mathcal{P}$  acts as a return-to-isotropy term on the full grid,

transferring momentum from the strongly anisotropic coherent structures into the relatively isotropic eddies in the environment. The pressure work on the plumes is formulated in terms of contributions from a virtual mass term (Gregory, 2001), an advective term and a drag term (Romps and Charn, 2015), yielding the following expression for the velocity-pressure gradient correlation:

$$\mathcal{P} = - \sum_{i>0} \frac{a_i}{a_0} (\bar{w}_i^* - \bar{w}_0^*) \left( \alpha_b \bar{b}_i^* - \alpha_d \bar{w}_i^* \frac{\partial \bar{w}_i^*}{\partial z} + \alpha_d \frac{(\bar{w}_i^* - \bar{w}_0^*) |\bar{w}_i^* - \bar{w}_0^*|}{H_i} \right), \quad (3.11)$$

where  $\alpha_a$  and  $\alpha_d$  are constant parameters,  $H_i$  is the plume height and  $\alpha_b$  is a function of the aspect ratio of the plume. Finally, assuming statistical equilibrium at scales  $l$  (Vassilicos, 2015), turbulent dissipation can be estimated from the spectral transport relation that follows from Kolmogorov's theory of inertial turbulence, giving Taylor's dissipation surrogate

$$\mathcal{D} = c_d \frac{\bar{e}_0^{3/2}}{l}. \quad (3.12)$$

Here,  $c_d$  is an empirical coefficient and  $l$  is the dissipation length, taken to be equal to the mixing length in our model. Expressions (3.3) and (3.5)–(3.12) provide closure to a 1.5-order model of turbulence within the EDMF framework, given diagnostic expressions for the mixing length  $l$  and for entrainment and detrainment.

### 3.4 Mixing length formulation

We seek to obtain a regime-independent eddy diffusivity closure that provides an accurate representation of turbulent subgrid-scale fluxes, over a wide range of host model resolutions. Thus, the eddy diffusivity should reduce to an LES-type closure at high resolution, while being able to account for the processes that modify turbulent fluxes at larger scales. The formulation of the closure is organized following this logic.

In section 3.4, we first adapt a minimum TKE dissipation closure proposed for LES subgrid models (Abkar and Moin, 2017) to the EDMF framework. Given the diffusive closure (3.3) and the eddy diffusivity (3.9), the minimum dissipation assumption can be used to construct a mixing length closure. This mixing length closure is shown to be equivalent to other proposed closures (e.g., Grisogono (2010)) for stable stratification, but additional entrainment terms appear in the general case. Section 3.4 highlights the limitations of this closure for climate modeling and weather prediction purposes when a prognostic TKE equation is used. Section 3.4 then introduces a modified mixing length closure, which builds on the

minimum dissipation model and corrects its shortcomings by introducing additional mechanisms of net TKE dissipation.

### Minimum Dissipation of Environmental TKE

As in Versteppen (2011) and Abkar and Moin (2017), we assume that at the small scales represented by the environment in the EDMF scheme, TKE is dissipated at least at the rate at which it is produced. This condition translates into an inequality for the production and dissipation terms in the environmental TKE budget (3.8):

$$\begin{aligned} & \overline{w'_0 b'_0} - \overline{w'_0 u'_0} \frac{\partial \langle u \rangle}{\partial z} - \overline{w'_0 v'_0} \frac{\partial \langle v \rangle}{\partial z} - \overline{w_0'^2} \frac{\partial \bar{w}_0}{\partial z} - \overline{\mathbf{u}'_{h,0} u'_0} \cdot \nabla_h \langle u \rangle - \overline{\mathbf{u}'_{h,0} v'_0} \cdot \nabla_h \langle v \rangle \\ & - \overline{\mathbf{u}'_{h,0} w'_0} \cdot \nabla_h \bar{w}_0 + \sum_{i>0} \left[ \frac{\Delta_{i0}}{\rho a_0} \left( \frac{(\bar{w}_i - \bar{w}_0)^2}{2} - \bar{e}_0 \right) - \frac{\hat{E}_{i0}}{\rho a_0} (\bar{w}_0^* (\bar{w}_i - \bar{w}_0) + \bar{e}_0) \right] \leq \mathcal{D}. \end{aligned} \quad (3.13)$$

Here, the terms involving TKE injection from entrained air are also taken to be locally balanced by dissipation, consistent with the assumption that entrainment events occur over timescales much shorter than the eddy turnover time  $K_{\psi,i}/\bar{e}_i$ . Note that the net dissipation condition (3.13) does not include redistribution terms, such as the turbulent transport or the velocity-pressure gradient correlation  $\mathcal{P}$ . Moreover, the inequality (3.13) represents a local condition for the environment, and it does not preclude net subgrid-scale energy production due to processes such as convection, represented by plumes. Denoting the difference between the right-hand side and the left-hand side of (3.13) as the net environmental TKE dissipation  $\gamma_0$ , the prognostic environmental TKE equation (3.8) reduces to

$$\frac{\partial \bar{e}_0}{\partial t} + \bar{w}_0 \frac{\partial \bar{e}_0}{\partial z} + \langle \mathbf{u}_h \rangle \cdot \nabla_h \bar{e}_0 = -\frac{1}{\rho a_0} \left[ \frac{\partial}{\partial z} \left( \rho a_0 \overline{w'_0 e'_0} \right) + \nabla_h \cdot \left( \rho a_0 \overline{\mathbf{u}'_{h,0} e'_0} \right) \right] - \mathcal{P} - \gamma_0. \quad (3.14)$$

Here,  $\mathcal{P}$  captures the effect of plumes on the environmental TKE. The evolution of the grid-mean TKE that follows from decomposition (3.5) and the simplified prognostic equation (3.14) is

$$\begin{aligned} & \frac{\partial \langle e \rangle}{\partial t} + \langle \mathbf{u}_h \rangle \cdot \nabla_h \langle e \rangle + \langle w \rangle \frac{\partial \langle e \rangle}{\partial z} + \frac{1}{\rho} \left[ \nabla_h \cdot \left( \rho \langle \mathbf{u}_h^* e^* \rangle \right) + \frac{\partial (\rho \langle w^* e^* \rangle)}{\partial z} \right] = \sum_i a_i \left( \bar{w}_i^* \bar{b}_i^* - \bar{w}_i^{*2} \frac{\partial \langle w \rangle}{\partial z} \right) \\ & - a_0 \gamma_0 + a_0 \left( \overline{w_0'^2} \frac{\partial \bar{w}_0^*}{\partial z} + \overline{\mathbf{u}'_{h,0} w'_0} \cdot \nabla_h \bar{w}_0^* \right) - \sum_{i>0} \left[ \frac{\Delta_{i0}}{\rho} \frac{(\bar{w}_i - \bar{w}_0)^2}{2} - \frac{\hat{E}_{i0}}{\rho} \bar{w}_0^* (\bar{w}_i - \bar{w}_0) \right]. \end{aligned} \quad (3.15)$$

A detailed derivation of this equation and the subgrid-scale kinetic energy pathways in the extended EDMF scheme is described in 3.8. Under the net dissipation closure

(3.13), grid-mean TKE production occurs through the first two terms on the right-hand side of (3.15): the convective buoyancy flux and the subdomain-scale shear production.

The net dissipation condition (3.13) can be written in terms of the mixing length by introducing the closures described in section 3.3. Using Taylor's dissipation surrogate (3.12) and downgradient closures for the shear and buoyancy terms of the form

$$\sum_{j=1}^3 \sum_{k=1}^3 -u'_{k,0} u'_{j,0} \frac{\partial \bar{u}_{j,0}}{\partial x_k} = \sum_{j=1}^3 \sum_{k=1}^3 K_m \left( \frac{\partial \bar{u}_{j,0}}{\partial x_k} \right)^2, \quad \overline{w'_0 b'_0} = -K_h \frac{\partial \bar{b}_0}{\partial z}, \quad (3.16)$$

the inequality (3.13) leads to a condition for the maximum value of the mixing length  $l$  at which the net dissipation  $\gamma_0$  is still positive semidefinite:

$$\left\{ \sum_{k=1}^3 \left[ \left( \frac{\partial \langle u \rangle}{\partial x_k} \right)^2 + \left( \frac{\partial \langle v \rangle}{\partial x_k} \right)^2 + \left( \frac{\partial \bar{w}_0}{\partial x_k} \right)^2 \right] - \frac{1}{\text{Pr}_t} \frac{\partial \bar{b}_0}{\partial z} \right\} l^2 + \sum_{i>0} \left[ \frac{\Delta_{i0}}{\rho a_0} \left( \frac{(\bar{w}_i - \bar{w}_0)^2}{2} - \bar{e}_0 \right) - \frac{\hat{E}_{i0}}{\rho a_0} (\bar{w}_0^* (\bar{w}_i - \bar{w}_0) + \bar{e}_0) \right] l \leq \frac{c_d}{c_m} \bar{e}_0. \quad (3.17)$$

Here, the environmental buoyancy gradient is computed following Tan et al. (2018), taking into account possible phase change effects. In (3.16) and (3.17),  $x_k$  and  $u_{k,0}$  represent the  $k$ -th coordinate and  $k$ -th velocity component in the environment, respectively. From the inequality (3.17), an expression for the mixing length that minimizes turbulent dissipation can be obtained by solving for  $l$ . This is equivalent to setting  $\gamma_0 = 0$  in (3.14) and (3.15). For the resulting value of the mixing length, production and dissipation of TKE are locally balanced:

$$l_{\text{tke}} = \frac{\sqrt{\Delta} - \mathcal{I}}{2(\mathcal{S}_l + \mathcal{B}_l)} = -\frac{\mathcal{I}}{2(\mathcal{S}_l + \mathcal{B}_l)} + \frac{\sqrt{\mathcal{I}^2 + 4(\mathcal{S} + \mathcal{B})\mathcal{D}}}{2(\mathcal{S}_l + \mathcal{B}_l)}. \quad (3.18)$$

Here,  $\Delta$  is the discriminant and the different terms are given by

$$\mathcal{S}_l + \mathcal{B}_l = c_m \bar{e}_0^{1/2} \left\{ \sum_{k=1}^3 \left[ \left( \frac{\partial \langle u \rangle}{\partial x_k} \right)^2 + \left( \frac{\partial \langle v \rangle}{\partial x_k} \right)^2 + \left( \frac{\partial \bar{w}_0}{\partial x_k} \right)^2 \right] - \frac{1}{\text{Pr}_t} \frac{\partial \bar{b}_0}{\partial z} \right\},$$

$$\mathcal{I} = \sum_{i>0} \left[ \frac{\Delta_{i0}}{\rho a_0} \left( \frac{(\bar{w}_i - \bar{w}_0)^2}{2} - \bar{e}_0 \right) - \frac{\hat{E}_{i0}}{\rho a_0} (\bar{w}_0^* (\bar{w}_i - \bar{w}_0) + \bar{e}_0) \right], \quad (3.19)$$

$$\mathcal{S} + \mathcal{B} = (\mathcal{S}_l + \mathcal{B}_l)l.$$

In (3.18), the product  $(\mathcal{S} + \mathcal{B})\mathcal{D}$  is independent of the mixing length, so  $l_{\text{tke}}$  can be readily evaluated. Although the term  $(\mathcal{S} + \mathcal{B})$  is sign-indefinite, the discriminant

$$\Delta = \mathcal{I}^2 + 4(\mathcal{S} + \mathcal{B})\mathcal{D}$$

in (3.18) can be shown to remain positive semidefinite even when the shear and buoyancy terms result in TKE destruction, provided that the inequality (3.13) holds. This is because the minimum dissipation balance requires

$$\mathcal{I} = \mathcal{D} - (\mathcal{S} + \mathcal{B}), \quad (3.20)$$

so that the expression for the discriminant  $\Delta$  is of the form

$$\Delta = [\mathcal{D} - (\mathcal{S} + \mathcal{B})]^2 + 4(\mathcal{S} + \mathcal{B})\mathcal{D} = [\mathcal{D} + (\mathcal{S} + \mathcal{B})]^2 \geq 0. \quad (3.21)$$

The mixing length  $l_{\text{tke}}$  depends on local characteristics of the environment and on the vertical velocity difference between subdomains, which enters the injection term  $\mathcal{I}$  in (3.19). Hence, convection modifies the environmental diffusive transport directly through entrainment processes. In addition, convection also regulates the time evolution of turbulent fluxes through its effect on the prognostic environmental TKE equation (3.14), captured by  $\mathcal{P}$ .

This approach can also be applied to turbulence models that retain covariance terms  $\overline{w'_i \psi'_i}$  for other subdomains, and not only for the environment. In this case, the minimum dissipation condition may be used to obtain a characteristic mixing length  $l_{\text{tke},i}$  for each subdomain. However, variance within plumes can also be accounted for by variance among plumes when the number of subdomains is increased.

In stably stratified boundary layers, where convection is inhibited, pressure work and entrainment fluxes in (3.6) act to homogenize the different subdomains, such that  $\bar{\psi}_i^* \rightarrow 0$  for any variable  $\psi$  and  $a_0 \rightarrow 1$  (i.e., there are no convective plumes). Under these conditions, the minimum dissipation mixing length (3.18) reduces to the expression proposed by Grisogono (2010) for steady-state stable boundary layer (SBL) flow:

$$\begin{aligned} l_{\text{tke}} &= \frac{\sqrt{(\mathcal{S} + \mathcal{B})\mathcal{D}}}{(\mathcal{S}_l + \mathcal{B}_l)} \\ &= \sqrt{\frac{c_d}{c_m} \langle e \rangle} \left\{ \sum_{k=1}^3 \left[ \left( \frac{\partial \langle u \rangle}{\partial x_k} \right)^2 + \left( \frac{\partial \langle v \rangle}{\partial x_k} \right)^2 + \left( \frac{\partial \langle w \rangle}{\partial x_k} \right)^2 \right] - \frac{1}{\text{Pr}_t} \frac{\partial \langle b \rangle}{\partial z} \right\}^{-1/2}. \quad (3.22) \end{aligned}$$

The balance between shear production, destruction due to stratification, and dissipation, which arises when using this mixing length, is a well-known leading-order state in neutral (Spalart, 1988) and moderately stable boundary layer flows (D. Li et al., 2016).

### Limitations of the Minimum-Dissipation Closure

Expression (3.18) for the mixing length  $l_{\text{tke}}$  captures the leading-order balance in the environmental TKE budget at small scales. However, a model with a diffusive closure based on  $l_{\text{tke}}$  cannot fully describe the dynamics of the boundary layer at the coarse resolutions typical of GCMs, on the order of  $10^4$  m in the horizontal and 10–100 m in the vertical. At these scales, the resolved horizontal gradients are weak, and the environmental TKE equation (3.14) that results from using  $l_{\text{tke}}$  can be simplified using the boundary layer approximation (neglecting horizontal relative to vertical derivatives):

$$\frac{\partial \bar{e}_0}{\partial t} + \bar{w}_0 \frac{\partial \bar{e}_0}{\partial z} = -\frac{1}{\rho a_0} \frac{\partial}{\partial z} \left( \rho a_0 \overline{w'_0 e'_0} \right) - \mathcal{P}. \quad (3.23)$$

Note that we set  $\gamma_0 = 0$  to obtain (3.23), since we are considering the case where  $l = l_{\text{tke}}$  and production locally balances dissipation. In stable conditions ( $\mathcal{P} = 0$ ), integrating the conservative form of (3.23) from the surface layer ( $z_s$ ) to the free troposphere above ( $z_i$ ) yields the evolution equation for the vertically integrated environmental TKE:

$$\int_{z_s}^{z_i} \frac{\partial(\rho a_0 \bar{e}_0)}{\partial t} dz = - [\rho a_0 \overline{w_0 e_0}]_{z_s}^{z_i} \approx - \rho a_0 K_m \left. \frac{\partial \bar{e}_0}{\partial z} \right|_{z_s}. \quad (3.24)$$

In stable conditions,  $a_0 \approx 1$  and  $\bar{\psi}_i^* \approx 0$  for any variable  $\psi$ . In addition, the absence of plumes implies that detrainment and entrainment processes are negligible. From (3.24), it follows that the evolution of the vertically integrated TKE under the minimum dissipation closure only depends on the flux from the unresolved surface layer in stable conditions. But unbalanced TKE dissipation has been observed to become increasingly important as stratification develops in field studies of the atmospheric boundary layer (D. Li et al., 2016), and it can be expected to play a role in conditions of strong surface cooling. The budget (3.24) cannot capture unbalanced TKE destruction within the boundary layer due to stratification. Furthermore, the minimum dissipation mixing length  $l_{\text{tke}}$  leads to enhanced eddy diffusion with increasing stratification, as deduced from (3.22). This is contrary to the evidence of turbulent mixing being inhibited in strong stratification, such as near strong inversions.

The limitations of a minimum dissipation model also become apparent in convectively unstable boundary layers. Integrating the TKE equation (3.23) in the vertical, the evolution of the vertically integrated environmental TKE in convective

conditions reads

$$\int_{z_s}^{z_i} \frac{\partial(\rho a_0 \bar{e}_0)}{\partial t} dz = \rho a_0 \overline{w_0 e_0} |_{z_s} + \int_{z_s}^{z_i} \rho \sum_{i>0} a_i (\bar{w}_i^* - \bar{w}_0^*) \frac{\partial \bar{\Psi}_i^*}{\partial z} dz + \int_{z_s}^{z_i} \sum_{i>0} (\Delta_{i0} - E_{i0}) \bar{e}_0 dz. \quad (3.25)$$

Here, the last term only accounts for changes in environmental area fraction and does not result in a source or sink of  $\bar{e}_0$  (Tan et al., 2018). A major difference between the SBL budget (3.24) and the convective budget (3.25) is the contribution of the velocity-pressure gradient correlation. From the velocity-pressure gradient relation (3.11), pressure work captures the important energization of turbulence in the environment owing to ascending or descending plumes (Schumann and Moeng, 1991). At the grid scale, the source of this subgrid-scale energy term is the convective buoyancy flux in (3.15), which accelerates the plumes in convective conditions.

The TKE balance (3.25) shows that, in convective conditions, the source of environmental TKE from updrafts or downdrafts can only be compensated by the flux from the unresolved surface layer. This is often a source term rather than a sink term, because shear production is surface intensified. Thus, the TKE balance (3.25) suggests an unbalanced growth of TKE in convective boundary layers. This continuous TKE increase in convective conditions is inconsistent with LES results showing quasi-stationary TKE levels in convective boundary layers (Nieuwstadt et al., 1993).

The TKE balances (3.24) and (3.25) highlight the shortcomings of the minimum dissipation balance (3.18) as a general closure for diffusive mixing in the boundary layer in stable and convective conditions. The lack of net dissipation mechanisms in the vertically integrated TKE balance hinders the correct representation of important processes, such as the shallowing of the boundary layer in the late afternoon or the sharp mixing inhibition near inversions. Moreover, it precludes reaching a quasi-stationary state in statically unstable boundary layers. Nevertheless, the limitations of the minimum dissipation model can be used to inform the construction of a generalized master length scale based on the TKE production–dissipation inequality (3.13).

The limitations of the minimum dissipation balance showcased in this section are not necessarily applicable to other turbulence models. For example, Y. He et al. (2019) use the production–dissipation condition to diagnose TKE and eddy diffusivity from a mixing length  $l$ . This allows an instantaneous adjustment of TKE to a new balanced

state, at the cost of representing convection with an empirical parameterization that has no subgrid interaction with turbulent diffusion.

### Constrained Minimization of TKE Dissipation

A master length scale that corrects the limitations of the minimum-dissipation model can be constructed by taking dissipation to be higher than production under certain circumstances. Using closures of the form (3.9) and (3.16) for the turbulent fluxes and (3.12) for the dissipation, it follows from the production–dissipation inequality (3.13) that excess dissipation occurs for  $l < l_{\text{tke}}$ . Hence, unbalanced TKE dissipation arises naturally in regions of the boundary layer where the characteristic size of environmental eddies is constrained to be smaller than  $l_{\text{tke}}$ . A general mixing length capturing this condition can be written as

$$l = s_{\min}(l_{\text{tke}}, l_1, l_2, \dots), \quad (3.26)$$

where  $l_j$  ( $j = 1, 2, \dots, N$ ) are candidate mixing lengths based on flow constraints, and  $s_{\min}(\mathbf{x})$  is a smooth minimum function defined in 3.7. The TKE production–dissipation inequality (3.17) with the closures substituted implies that the minimum length scale (3.26) provides maximum TKE dissipation among the candidate length scales. Thus, the use of the minimum length scale (3.26) is equivalent to the minimization of TKE dissipation in (3.13) subject to the constraint that dissipation exceeds the candidate dissipation rates,

$$\mathcal{D} \geq \mathcal{D}|_{l=l_j} \quad \forall j, \quad (3.27)$$

where  $\mathcal{D}|_{l=l_j}$  is the candidate dissipation rate evaluated at length scale  $l_j$ .

Our suggestion for choosing a general mixing length as a smooth minimum of various candidates contrasts with the common practice (e.g., Y. He et al., 2019; Han and Bretherton, 2019) to use the expression suggested by Blackadar (1962),

$$l_h = \left( \frac{1}{l_1} + \frac{1}{l_2} \right)^{-1}, \quad (3.28)$$

for a master length scale  $l_h$ . This length scale  $l_h$ , proportional to the harmonic mean of the candidates  $l_1$  and  $l_2$ , is smaller than both  $l_1$  and  $l_2$ . If closures similar to (3.9) and (3.12) are used in a prognostic equation for TKE, the mixing length (3.28) results in an unrealistic intensification of TKE dissipation in regions where the candidate length scales  $l_1$  and  $l_2$  are similar. This undesirable characteristic is avoided by using the smooth minimum (3.26).

We consider two limiting factors for the characteristic length scale of turbulent motion in boundary layer flows: stable stratification and the distance to solid boundaries.

### Stratification Constraints

Environmental stratification constrains the size of turbulent eddies by inhibiting the vertical displacement of air masses. Stably stratified turbulence is known to show high vertical variability and reorganization into layered structures, with most mixing occurring within the layers (Waite, 2011). The thickness of these layers is determined by the vertical scale at which the governing dynamic equations become self-similar (Billant and Chomaz, 2001; Augier et al., 2012), known as the buoyancy scale  $l_b$ . For a flow with an imposed stratification given by the Brunt-Väisälä frequency  $N_e$ , this length scale is

$$l_b = c_b \frac{(\bar{\epsilon}_0)^{1/2}}{N_e}, \quad (3.29)$$

where  $c_b$  is an empirical coefficient. It is important to note that imposing  $l_b$  as an upper bound for the size of eddies is similar to doing so by the Ozmidov scale  $l_o \sim \sqrt{\mathcal{D}/N_e^3}$  only if turbulent motions at the scale in question are assumed to be in the inertial subrange, such that (3.12) holds. In this case, an expression equivalent to (3.29) for the Ozmidov scale is

$$l_o = \left( \frac{c_b^3 \mathcal{D}}{c_d N_e^3} \right)^{1/2}. \quad (3.30)$$

However, recent experimental studies suggest that under strong stratification, turbulence may not display an inertial subrange (Grachev et al., 2013). In that case, expression (3.12) and the Ozmidov scale (3.30) may not be applicable (D. Li et al., 2016), whereas the buoyancy scale (3.29) still holds.

The buoyancy frequency of moist air depends on the latent heat release and evaporative cooling associated with the vertical displacement of air parcels. In general, the effective static stability  $N_e$  lies between the dry and the moist adiabatic limits. In the same spirit as O’Gorman (2011), we use an effective static stability of the form

$$N_e^2 = \frac{g}{\bar{\theta}_{v,0}} \left( \frac{\partial \bar{\theta}_{v,0}}{\partial z} - \lambda \frac{\partial \bar{\theta}_{v,0}}{\partial z} \Big|_{\bar{\theta}_{vl,0}} \right) = \frac{g}{\bar{\theta}_{v,0}} \left[ (1 - \lambda) \frac{\partial \bar{\theta}_{v,0}}{\partial z} + \lambda \frac{\partial \bar{\theta}_{v,0}}{\partial \bar{\theta}_{vl,0}} \frac{\partial \bar{\theta}_{vl,0}}{\partial z} \right], \quad (3.31)$$

where  $\theta_v$  is the virtual potential temperature,  $\lambda$  represents the area fraction of environmental air undergoing phase change, and  $\theta_{vl}$  is the liquid-water virtual potential temperature, defined below. In expression (3.31), we have used the definition

$$\left. \frac{\partial \bar{\theta}_{v,0}}{\partial z} \right|_{\bar{\theta}_{vl,0}} = \frac{\partial \bar{\theta}_{v,0}}{\partial z} - \frac{\partial \bar{\theta}_{v,0}}{\partial \bar{\theta}_{vl,0}} \frac{\partial \bar{\theta}_{vl,0}}{\partial z}. \quad (3.32)$$

Expression (3.31) differs from that presented in O’Gorman (2011) in the diagnosis of  $\lambda$  and the use of  $\bar{\theta}_{vl,0}$  instead of the saturated equivalent potential temperature. These differences arise from the much smaller vertical scale considered here. At scales of 10–100 m, the occurrence of phase changes is not necessarily correlated with the sign of the vertical velocity of air parcels (O’Gorman and Schneider, 2006; O’Gorman, 2011). Thus,  $\lambda$  cannot be diagnosed from vertical velocity statistics. In the non-precipitating cases considered here,  $\lambda$  is given by the environmental cloud fraction  $f_{c,0}$ . Cloud fraction diagnosis is cloud-type dependent in many current GCMs (Collins et al., 2004). In our EDMF scheme, we use a regime-independent probabilistic cloud scheme (see 3.9).

The liquid-water virtual potential temperature  $\theta_{vl}$  appearing in the effective static stability measures the buoyancy of cloudy air parcels when moist-adiabatically returned to clear conditions (Grenier and Bretherton, 2001; Marquet, 2011),

$$\bar{\theta}_{vl} \approx (1 + \eta \bar{q}_t) \bar{\theta}_l \approx \bar{\theta}_v \exp\left(-\frac{L_v \bar{q}_l}{c_p \bar{T}}\right). \quad (3.33)$$

We use  $\theta_{vl}$  instead of the saturated equivalent potential temperature because  $\theta_{vl}$  converges to  $\theta_v$  in the dry limit, while also including the effects of latent heat release. Here,  $\eta = R_v/R_d - 1$ ,  $L_v$  is the latent heat of vaporization,  $c_p$  is the specific heat of air,  $q_t$  and  $q_l$  are the total and liquid water specific humidities,  $\theta_l$  is the liquid water potential temperature,  $T$  is the temperature and  $R_v, R_d$  are the gas constants for water vapor and dry air, respectively. Expression (3.33) can be used to evaluate  $\partial \bar{\theta}_{v,0}/\partial \bar{\theta}_{vl,0}$  in (3.31) and (3.32). Note that the effective static stability (3.31) converges to the dry limit when  $q_l \rightarrow 0$  for all values of  $\lambda$ ; it reduces to  $N_e^2 = (1 - \lambda)N^2$ , with dry buoyancy frequency  $N$ , in conditions that are well mixed in  $\theta_l$  and  $q_t$ .

### Wall Constraints

The presence of boundaries also imposes an upper limit on the size of eddies near them. Following Monin and Obukhov (1954), the eddy diffusivity in the surface

layer has the form

$$K_{\psi,w} = \frac{u_* \kappa z}{\phi_\psi(\xi)} \quad (3.34)$$

where  $\xi = z/L$ ,  $\phi_\psi(\xi)$  is an empirical stability function,  $\kappa$  is the von Kármán constant,  $L$  is the Obukhov length, and  $u_*$  is the friction velocity. The upper bound for the mixing length near the surface is obtained by matching this eddy diffusivity with the expression (3.9) for the eddy diffusivity:

$$l_w = \frac{\kappa}{c_\psi \kappa_* \phi_\psi(\xi)} z. \quad (3.35)$$

Here,  $\kappa_* = \bar{e}_0^{1/2}/u_*$  is the ratio of rms turbulent velocity to the friction velocity in the surface layer. The friction velocity in our model is diagnosed using the flux-profile relationships of Byun (1990), except in free convective conditions. When the conditions for free convection are satisfied, the diagnostic of  $u_*$ , which is a function of the horizontal wind speed at the lowest model level, is modified following Beljaars (1995).

The choice of a common master length for momentum and tracer diffusion implies  $c_h \phi_h = c_m \phi_m$ , such that  $\phi_h = \text{Pr}_t \phi_m$ . In our formulation, the turbulent Prandtl number is taken to be a function of the gradient Richardson number  $\text{Ri}$ , based on a simplified cospectral budget of momentum and heat transport (Katul et al., 2013; D. Li, 2019):

$$\text{Pr}_t = \frac{2\text{Ri}}{1 + \omega_2 \text{Ri} - \sqrt{-4\text{Ri} + (1 + \omega_2 \text{Ri})^2}} \text{Pr}_{t,0}. \quad (3.36)$$

Here,  $\omega_2 = 40/13$  is a phenomenological constant, and  $\text{Pr}_{t,0}$  is the turbulent Prandtl number in neutral conditions. The stability function  $\phi_m$  is often written in the form (Businger et al., 1971; Nakanishi, 2001)

$$\phi_m = [1 + a_1(\xi)\xi]^{a_2(\xi)}, \quad a_i = a_i^- + (a_i^+ - a_i^-)H(\xi), \quad (3.37)$$

where  $H(\cdot)$  is the Heaviside function and  $a_i^-, a_i^+$  are empirical functions. The values of  $a_i^-$  are taken as negative definite to reflect the convective elongation of eddies in unstable conditions. In stable conditions, self-similarity of the flow requires that  $a_2^+ = 1$  and  $a_1^+ > 0$ , such that under strong stratification, the mixing length (3.35) becomes independent of  $\xi$ . As shown by Monin and Obukhov (1954), the asymptotic limit of  $\phi_m$  under strong stratification also requires that  $a_1^+ = \text{Pr}_t(\text{Ri}_{\text{st}})/\text{Ri}_{\text{st}}$ . Here,  $\text{Ri}_{\text{st}}$  is the asymptotic Richardson number at  $\xi \gg 1/a_1^+$  in the surface layer.

The empirical function (3.37) has been shown to become increasingly inaccurate with stability for  $\xi > 0.5$  (Sorbján and Grachev, 2010; Optis et al., 2016). Moreover,

extending the use of the limiting scale  $l_w$  above the surface layer precludes the use of  $a_1^+ \neq 0$  in stable conditions, since the Obukhov length characterizes stratification only in the constant flux layer near the surface. Although the use of  $l_w$  in expression (3.26) mandates  $a_1^+ = 0$ , the effect of stability in eddy diffusion is still captured by  $l_b$ . In the constant flux layer, the limiting length  $l_b$  is equivalent to the use of the empirical function (3.37) in the strongly stable limit, with

$$a_1^+ = \frac{1}{(\kappa_*^2 c_m c_b)^2} \text{Pr}_t, \quad \xi \gg \frac{1}{a_1^+}. \quad (3.38)$$

Under weaker stratification, turbulence in the surface layer can reach a quasi-steady state (Spalart, 1988). In this case, the limiting scale  $l_w$  should converge to  $l_{\text{tke}}$ . Assuming that entrainment processes are limited to dynamical entrainment by the plumes in the surface layer, the ratio of the two length scales can be written as

$$\left. \frac{l_w}{l_{\text{tke}}} \right|_{\bar{e}_0} = \frac{(1 - \text{Ri}/\text{Pr}_t)^{1/2}}{(c_d c_m)^{1/2} \kappa_*^2}, \quad (3.39)$$

which is constant under neutral stratification and is slowly varying with Ri due to the opposing effect of the Prandtl number (3.36). From (3.39), the convergence of  $l_{\text{tke}}$  to  $l_w$  in the surface layer is satisfied for  $(c_d c_m)^{1/2} \kappa_*^2 \approx 1$ .

The use of a soft minimum function for the mixing length (3.26) allows for a smooth transition from Monin-Obukhov similarity theory near the surface to a local turbulent closure farther away from it, where the use of Monin-Obukhov scaling may be inaccurate (Optis et al., 2016). In addition, the expressions (3.38) and (3.39) show that this transition is asymptotically consistent.

### Master Mixing Length

Finally, the smooth minimum of the three candidate length scales determines the mixing length,

$$l = s_{\min}(l_{\text{tke}}, l_w, l_b), \quad (3.40)$$

where

$$l_{\text{tke}} = -\frac{\mathcal{I}}{2(\mathcal{S}_l + \mathcal{B}_l)} + \frac{\sqrt{\mathcal{I}^2 + 4(\mathcal{S} + \mathcal{B})\mathcal{D}}}{2(\mathcal{S}_l + \mathcal{B}_l)}, \quad (3.41)$$

$$l_b = c_b \frac{(\bar{e}_0)^{1/2}}{N_e}, \quad (3.42)$$

and

$$l_w = \frac{\kappa}{c_m \kappa_* \phi_m(\xi)} z. \quad (3.43)$$

The proposed diffusive closure is implemented using equations (3.40)–(3.43), as well as the prognostic environmental TKE equation (3.8) in flux form (see Eq. (3.51)). The mixing length (3.40) depends on a group of nondimensional parameters  $C$  that must be obtained empirically:

$$C = \{c_m, c_d, c_b, \kappa, \kappa^*, a_1^-, a_2^-, \text{Pr}_{t,0}\}. \quad (3.44)$$

Values for these parameters are reported in studies of boundary layer turbulence, obtained from field observations (Businger et al., 1971) or LES (Nakanishi, 2001). However, the direct use of some of these values in the EDMF scheme is not justified due to the decomposition of the subgrid-scale flow into different subdomains. Because of the large size of the parameter space  $C$  and the presence of other parameters in the EDMF scheme, we limit the parameter optimization process to  $C^* = \{c_m, c_d, c_b\}$  in this study.  $C^*$  contains the parameters that appear in the closures that are most strongly affected by the domain decomposition. All other parameters in  $C$ , except  $\text{Pr}_{t,0}$ , arise from similarity theory arguments for the unresolved surface layer. Here, it is assumed that similarity arguments are valid outside convective updrafts, and all values are taken from Nakanishi (2001). For the simulations reported in the next section, the parameter space used is shown in Table 4.1. The rest of parameters used in the EDMF scheme, which do not appear explicitly in the formulation of the mixing length closure, are reported in Cohen et al. (2020).

Table 3.1: Parameters in the mixing length closure and values used in this study.

Symbol	Description	Value
$c_m$	Eddy viscosity coefficient	0.14
$c_d$	Turbulent dissipation coefficient	0.22
$c_b$	Static stability coefficient	0.63
$\kappa$	von Kármán constant	0.4
$\kappa^*$	Ratio of rms turbulent velocity to friction velocity	1.94
$a_1^-$	Empirical stability function coefficient	−100
$a_2^-$	Empirical stability function coefficient	−0.2
$\text{Pr}_{t,0}$	Turbulent Prandtl number in neutral conditions	0.74

### 3.5 Results for single-column simulations

Here we focus on case studies targeting the simulation of the Arctic stable boundary layer (SBL), stratocumulus clouds, and dry convection. The performance of the extended EDMF scheme in moist-convective cases is explored in Cohen et al. (2020), using the same set of closures and parameters. The extended EDMF scheme is tested

for horizontal resolutions typical of GCMs. Invoking the boundary layer approximation (neglecting horizontal derivatives), we perform simulations in a single-column model (SCM). The SCM is a one-dimensional vertical model that aims to represent a single atmospheric column within a GCM. Results from single-column simulations using the extended EDMF scheme are then compared to horizontal averages obtained from LES over the same domain. LES are set up by further discretizing the atmospheric column horizontally and using horizontal doubly-periodic boundary conditions.

The EDMF scheme used here differs from the one described in Tan et al. (2018) in the parameterizations of the eddy diffusivity  $K_\psi$ , the vertical pressure anomaly gradients in (3.6) and (3.10), entrainment and detrainment, and the addition of turbulent entrainment  $\hat{E}_{ij}$ . The parameterization of the eddy diffusivity follows (3.9) and (3.40)–(3.43). The entrainment parameterization is described in Cohen et al. (2020), and the treatment of the pressure anomaly term is shown in (3.11). In addition, although the theoretical framework presented here allows for the use of downdrafts, the implementation used in this section decomposes the domain solely into one updraft and its turbulent environment.

LES are performed using PyCLES, an anelastic fluid solver in which the subgrid-scale fluxes are treated implicitly by the WENO scheme used to discretize the prognostic equations (Pressel et al., 2015). Implicit LES using WENO numerics have been shown to result in higher effective resolution than other combinations of numerics and explicit SGS closures (Pressel et al., 2017). Finally, LES results from previous model intercomparison projects are also reported where available.

### **Stable Boundary Layer**

Statically stable conditions in the boundary layer inhibit convection, reducing the EDMF scheme to a diffusive closure. In the implementation of the scheme, this translates to conditioning the surface updraft area fraction on the sign of the surface buoyancy flux, such that it becomes zero in conditions of surface cooling. With no updrafts or downdrafts, the only contribution to the subgrid-scale flux (3.2) comes from the environmental downgradient turbulent flux (3.3). This leads to a high sensitivity of SCM simulations to changes in the mixing length formulation. Here we focus on the GEWEX Atmospheric Boundary Layer Study (GABLS), discussed in Beare et al. (2006).

## Simulation Setup

The initial and boundary conditions of the simulation are adapted from observations during the Beaufort and Arctic Seas Experiment (Curry et al., 1997) and follow Beare et al. (2006). The velocity field is initialized as  $(\langle u \rangle, \langle v \rangle) = (u_g, 0)$ , where the geostrophic velocity is  $u_g = 8 \text{ m s}^{-1}$ . The initial temperature sounding is given by a mixed layer with potential temperature  $\theta = 265 \text{ K}$  up to 100 m, overlain by an inversion with a potential temperature gradient of  $10 \text{ K km}^{-1}$ . The surface boundary condition is given by constant cooling,  $\dot{\theta}_{z=0} = -0.25 \text{ K h}^{-1}$ .

For both the SCM and LES, the domain height is 400 m. In the LES configuration, the domain spans 400 m in both horizontal directions as well. The LES data is generated using an isotropic mesh with  $\Delta x_i = 3.125 \text{ m}$  resolution, which translates into  $2 \times 10^6$  degrees of freedom. The full range of LES results from Beare et al. (2006), using the same resolution, is also included for reference. The SCM simulations are performed at vertical resolutions of  $\Delta z = 3.125 \text{ m}$ ,  $12.5 \text{ m}$ , and  $50 \text{ m}$  (128, 32, and 8 degrees of freedom, respectively). This range characterizes the performance of the EDMF scheme both at high resolution and for coarser resolutions typical of regional and global climate models in the lower troposphere. The time steps used in the SCM simulations, in order of increasing  $\Delta z$ , are  $\Delta t = 5, 15, \text{ and } 60 \text{ s}$ .

## Results

Figure 3.1 shows vertical profiles of  $\langle \theta \rangle$ ,  $\langle u \rangle$  and  $\langle v \rangle$  time-averaged over the 9th hour of simulation. The EDMF scheme captures well the boundary layer height and the intensity of the low-level jet, with little resolution dependence of the mean profiles up to  $\Delta z = 12.5 \text{ m}$ . At 50 m resolution, the SCM predicts a slightly deeper boundary layer. The EDMF-simulated TKE follows closely the LES data, as shown in Figure 3.2. The timeseries show periods of TKE growth due to the subgrid momentum flux from the surface layer, and periods of decay due to the increasing stratification. These changes in vertically integrated TKE are much smaller than the integrated TKE production and dissipation terms, as shown in Figure 3.3. The domain-mean TKE budget, which coincides with the environmental budget for stable conditions, is shown in Figure 3.3.

The two main causes of grid-sensitivity at 50 m resolution are the inability to capture the region of maximum shear production close to the surface, and the deterioration of the friction velocity diagnosis. The effect of the former can be observed in Figure

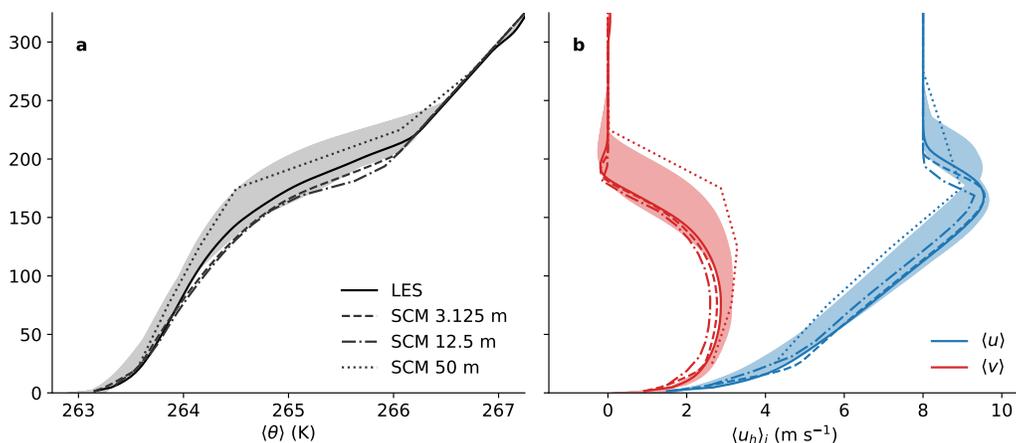


Figure 3.1: Profiles of (a) potential temperature and (b) horizontal velocity averaged over the 9th hour of the GABLS simulation. Results are shown for LES and for the EDMF-based SCM with  $\Delta z = 3.125$  m,  $\Delta z = 12.5$  m, and  $\Delta z = 50$  m. The shaded region represents the spread of LES results with  $\Delta z = 3.125$  m reported in Beare et al. (2006).

3.3. Even if the budget is correctly captured above 50 m, the absence of grid-cells at the lower levels results in a significant reduction of the vertically integrated production and dissipation. In addition, the diagnosis of  $u_*$  based on Byun (1990) overestimates the friction velocity at coarser resolutions. This can be observed by comparing the normalized TKE profile to the vertically integrated timeseries in Figure 3.2.

The dominant mixing length throughout the simulation is shown in Figure 3.2 for all heights. Initially, the wall-limited mixing length  $l_w$  is dominant below the inversion, due to the absence of mean shear and stratification. As the shear and stratification develop, the dominant mixing length profile attains a three-layered structure. Closest to the bottom boundary, the distance to the wall constrains the size of eddies. Farther away from the surface, the mixing length is determined by the local TKE balance. As stratification increases with height, the stratification-limited mixing length  $l_b$  becomes dominant, depleting TKE and limiting turbulent mixing. The eddy diffusivity, shown in Figure 3.2, is maximum near the transition from  $l_{tke}$  to  $l_b$ , where the mixing length is largest. Again, the overestimation of the friction velocity and the absence of grid-points in the lower layers result in an overestimation of the eddy diffusivity at coarse resolutions.

Both the LES and EDMF budgets show the quasi-balance of TKE sources and sinks throughout the boundary layer, even in regions where  $l_{tke}$  is not dominant.

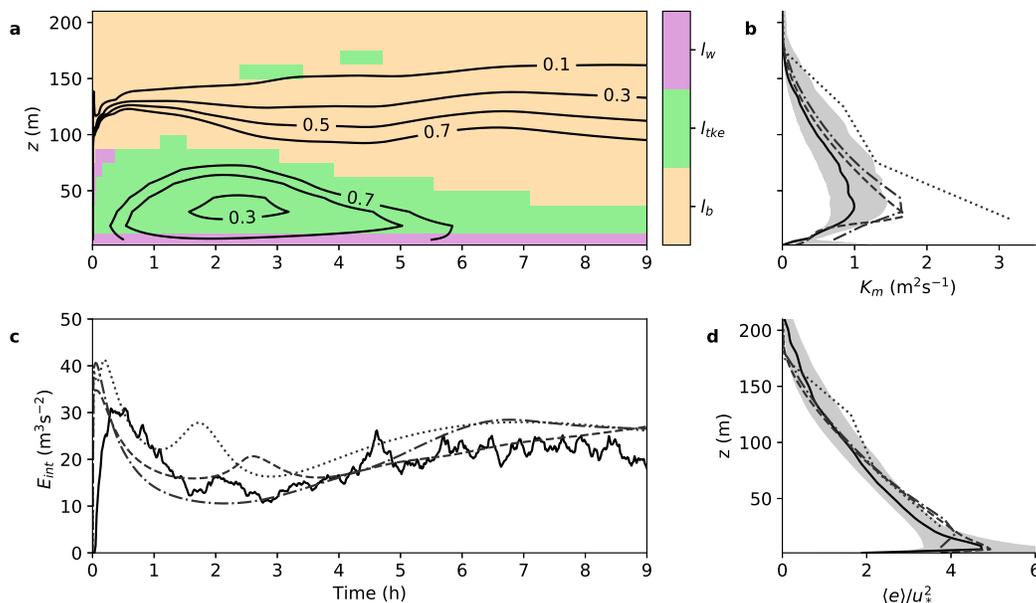


Figure 3.2: (a) Contours of eddy diffusivity  $K_m$  ( $\text{m}^2 \text{s}^{-1}$ ) as a function of time and height for the GABLS simulation using the SCM with  $\Delta z = 12.5$  m. Colors show the dominant (minimum) mixing length. (b) Profiles of eddy diffusivity averaged over the 9th hour. LES diffusivity is diagnosed from the shear production term  $\mathcal{S}$  and the grid-mean shear. (c) Time series of vertically integrated TKE  $E_{int}$ . (d) Profiles of  $u_*$ -normalized TKE averaged over the 9th hour. In (d),  $u_*$  is the average friction velocity during the 9th hour, with  $u_* = 0.25 \text{ m s}^{-1}$  for LES and  $u_* = 0.26$ ,  $0.26$  and  $0.30 \text{ m s}^{-1}$  for the SCM cases, in order of increasing  $\Delta z$ . In (b), (c), and (d), results are shown for LES (solid line), EDMF with  $\Delta z = 3.125$  m (dashed line),  $\Delta z = 12.5$  m (dash-dotted line), and  $z = 50$  m (dotted line). The shaded region represents the spread of LES results with  $\Delta z = 3.125$  m reported in Beare et al. (2006).

The downgradient parameterization of shear production  $\mathcal{S}$ , buoyant production  $\mathcal{B}$ , and the turbulent transport  $\mathcal{T}$  results in profiles that follow closely the LES data, particularly at higher resolution. This validates the assumptions used to model the second-order moments in the extended EDMF scheme under stable stratification.

### Stratocumulus-Topped Boundary Layer

The ability of the extended EDMF scheme to represent the dynamics of the STBL is tested using as a baseline the second Dynamics and Chemistry of Marine Stratocumulus (DYCOMS-II) field study (Stevens et al., 2003), performed near the coast of San Diego, California. In particular, the conditions observed during the first research flight (RF01) are considered, for which precipitation was not observed.

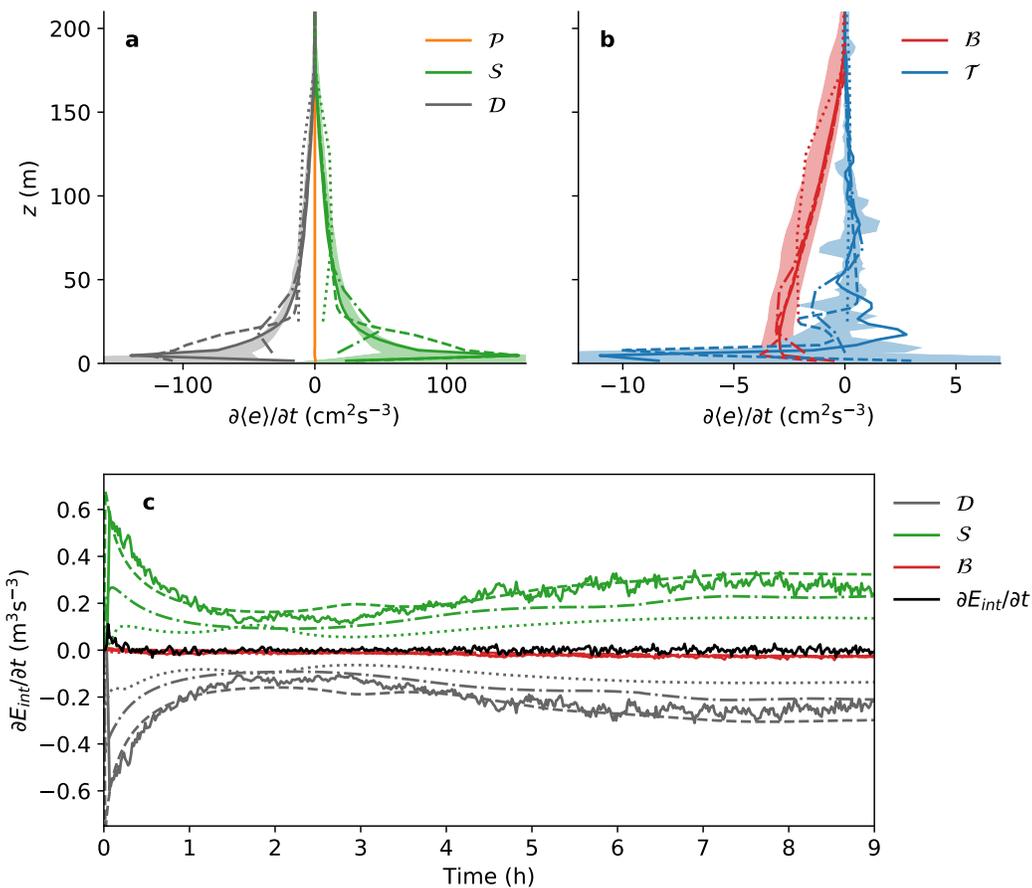


Figure 3.3: Profiles of TKE budget terms averaged over the 9th hour of GABLS simulation. Profiles shown in (a) are for shear production  $\mathcal{S}$ , dissipation  $\mathcal{D}$  and the pressure term  $\mathcal{P}$ . Shown in (b) are the buoyant production term  $\mathcal{B}$  and turbulent transport  $\mathcal{T}$ . (c) Time evolution of the vertically integrated TKE ( $E_{int}$ ) production and dissipation terms. The rate of change of  $E_{int}$  in LES is included as a reference. Results are shown for LES (solid line) and for EDMF with  $\Delta z = 3.125$  m (dashed line),  $\Delta z = 12.5$  m (dash-dotted line), and  $\Delta z = 50$  m (dotted line). The shaded region represents the spread of LES results with  $\Delta z=3.125$  m reported in Beare et al. (2006).

## Simulation Setup

The simulation setup for DYCOMS-II RF01 is reported in Stevens et al. (2005). The base state is initialized with a two-layer structure in  $\theta_l$  and  $q_t$ , separated by a strong inversion at  $z_i = 840$  m. The bottom layer is well-mixed in both conserved variables, with saturation and cloud formation occurring above 600 m. The cloud-top, located at  $z_i$ , is characterized by  $\Delta\theta_l = 8.5$  K and  $\Delta q_t = -7.5$  g kg<sup>-1</sup>. The free troposphere is warmer and drier than the mixed layer, with a  $\theta_l$ -lapse rate of  $(1/3)(z - z_i)^{-2/3}$  K m<sup>-1</sup> and constant  $q_t$ . The surface sensible and latent heat fluxes are set to 15 and 115 W m<sup>-2</sup>, respectively. The vertical water distribution induces radiative cloud-base warming and radiative cooling at cloud-top and in the free troposphere.

The domain height is 1.5 km. In the LES, the horizontal domain extent is set to 3.36 km. The resolution used for the LES is  $\Delta z = 5$  m in the vertical and  $\Delta x = 35$  m in the horizontal. This corresponds to  $2.76 \times 10^6$  degrees of freedom. The SCM simulations are performed with vertical resolutions of  $\Delta z = 5$  m, 20 m, 50 m, and 75 m, or 300, 75, 30, and 20 degrees of freedom, respectively. In the SCM simulations, the time step is diagnosed from a CFL condition based on the maximum updraft velocity in the domain, using a Courant number of 0.9. This results in average time steps of 3 s, 14 s, 39 s, and 63 s, respectively.

## Results

The mean profiles obtained with the extended EDMF scheme display very little resolution sensitivity compared to the spread of results from LES, as shown in Figure 3.4. LES of stratocumulus-topped boundary layers are strongly dependent on the discretization numerics and the treatment of subgrid-scale fluxes (Pressel et al., 2017). Overly diffusive LES models result in excessive cloud-top mixing, reducing the water content of the cloud layer and transitioning to decoupled cumulus-like conditions.

Similarly, the ability of SCM simulations to capture the stratocumulus-cloud layer is contingent upon the cloud-top mixing not being too strong. With large gradients in  $q_t$  and  $\theta_l$  across the inversion, the mixing length is the main limiter of cloud-top diffusive mixing. As shown in Figure 3.5a, the buoyancy scale (3.29) is crucial to limit the cloud-top eddy diffusivity and maintain a sharp inversion over the mixed layer (see 3.10 for details). It is important to note that in our formulation, the mixing

length may be smaller than  $\Delta z$ . This allows to maintain a coupled cloud layer even at coarse vertical resolution.

How the dominant mixing length varies with height in the STBL is shown in Figure 3.5a. Throughout most of the boundary layer, environmental mixing is determined by the minimum-dissipation balance. Mixing is constrained by stratification at cloud top and in the lower part of the cloud, where the environmental cloud fraction  $f_{c,0}$  is less than unity. The vertically integrated TKE obtained in the SCM simulations is similar across resolutions and follows closely the WENO-based LES statistics, as shown in Figure 3.5b. Again, the variation of TKE with resolution in the SCM simulations is significantly lower than the spread of values obtained with different LES, not all of which successfully simulate the presence of a stratocumulus cloud layer.

The liquid water path (LWP) time series from the SCM simulations are in agreement with the LES results. At coarse resolution, cloud-top entrainment of dry air is too low, which leads to an overestimation of  $q_l$  and LWP, as shown in Figures 3.4d and 3.5c. However, even at this resolution, the water content bias obtained with the EDMF scheme is significantly lower than the dry bias of some of the LES models.

The vertical heat and moisture fluxes, as well as the contributions from the turbulent flux (eddy diffusivity) and subdomain-mean terms (mass flux), are shown in Figure 3.4. The SCM simulations slightly overestimate the heat flux in the cloud layer and underestimate the moisture flux throughout the boundary layer. These biases compensate each other to some extent, leading to a small bias in the buoyancy flux. Similar biases are reported for models using the EDMF scheme and different parameterizations (E. Wu et al., 2020).

In the extended EDMF scheme, the environmental turbulent flux is the leading contributor to the buoyancy flux. The context of this decomposition should be considered when comparing these results to LES studies of the dynamics governing the STBL (e.g., Davini et al., 2017). Since we do not consider downdrafts in our SCM simulations, the environment contains all dynamic structures of the flow except updrafts. Therefore, the turbulent flux here also represents the transport due to downdrafts. Although LES studies emphasize the importance of convective transport due to downdrafts in stratocumulus clouds (Davini et al., 2017), we find that their implementation is not necessary to reproduce the STBL using the extended EDMF scheme. This is in agreement with E. Wu et al. (2020), where the authors show that the implementation of downdrafts in an EDMF scheme does not

significantly improve simulations of the STBL.

Vertical profiles of TKE and eddy diffusivity are shown in Figure 3.6. The magnitude of TKE is underestimated by the SCM in the cloud layer, while the SCM maintains similar values to LES in the subcloud-layer. The eddy diffusivity is also diagnosed from LES, with the Prandtl number expression (3.36) used in our model imposed as a constraint. From the closures (3.16), and since the simulations are performed in a single column (i.e., the horizontal derivatives in (3.16) are zero), we can estimate the eddy diffusivity from the environmental shear and buoyancy production as

$$K_m^{\text{les}} = -\frac{1}{2} \left[ \sum_{j=1}^3 \frac{\overline{w'_0 u'_{j,0}}}{\partial \bar{u}_{j,0} / \partial z} + \text{Pr}_t \frac{\overline{w'_0 b'_0}}{\partial \bar{b}_0 / \partial z} \right]. \quad (3.45)$$

In the diagnosis of (3.45), updraft and environment identification is performed using the methodology proposed in Couvreux et al. (2010). The eddy diffusivity in the SCM simulations follows a similar profile to  $K_m^{\text{les}}$  below the cloud layer, underestimating  $K_m^{\text{les}}$  at coarser resolution but converging toward  $K_m^{\text{les}}$  as the resolution is refined. The peak in  $K_m^{\text{les}}$  within the cloud layer is due to a positive  $\overline{w'_0 b'_0}$  under a vanishing  $\bar{b}_0$  gradient. This environmental buoyancy production may be attributed to convective downdrafts, which are considered to be part of the environment in our analysis.

### Dry Convection

The dry convective boundary layer differs from the previous cases in that the mass-flux term is the leading order contribution to the subgrid-scale vertical transport throughout most of the boundary layer. However, an accurate parameterization of the eddy-diffusivity contribution is still necessary for a correct simulation of the dry convective boundary layer.

### Simulation Setup

The simulation setup follows Nieuwstadt et al. (1993). The base state is initialized as a mixed layer with potential temperature  $\theta = 300$  K up to  $z_1 = 1350$  m, above which potential temperature increases at a rate of  $3$  K  $\text{km}^{-1}$ . The flow, which is initialized with a horizontal velocity of  $1$  cm  $\text{s}^{-1}$ , is driven by a constant surface heat flux of  $\langle w^* \theta^* \rangle = 6$  K cm  $\text{s}^{-1}$ .

The simulation is performed in a domain spanning  $3.75$  km in the vertical. For the LES, the horizontal cross-sectional area is  $6.4 \times 6.4$   $\text{km}^2$ , and the resolution is

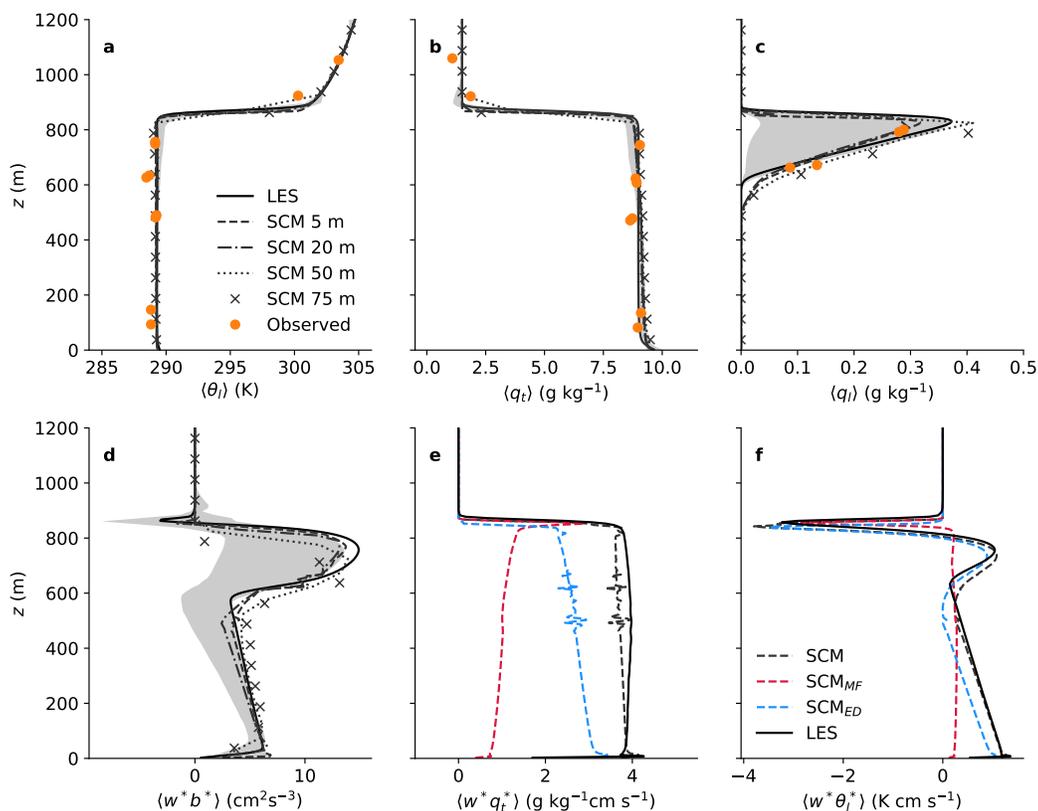


Figure 3.4: Profiles of (a) liquid water potential temperature, (b) total water specific humidity, (c) liquid water specific humidity, (d) vertical buoyancy flux, (e) vertical transport of  $q_t$  and (f) vertical transport of  $\theta_l$ . Profiles averaged over the 4th hour of the DYCOMS-II RF01 simulation. In (e) and (f), the eddy diffusivity ( $SCM_{ED}$ ) and mass flux ( $SCM_{MF}$ ) components of the vertical flux are shown (plotting conventions follow the legend in panel (a)). The shaded region represents the spread of LES results reported in Stevens et al. (2005). Observations are also reported in Stevens et al. (2005).

$\Delta z = 25$  m in the vertical and  $\Delta x = 50$  m in the horizontal. The SCM simulations are performed with vertical resolutions of 25, 50, and 150 m. As for the DYCOMS-II simulations, the time step is diagnosed from a CFL condition. The average time step for these simulations is 14 s, 30 s, and 100 s, respectively.

## Results

Time-averaged profiles of potential temperature and vertical buoyancy flux are shown in Figure 3.7 for the 5th hour of simulation. The potential temperature mixed layer and its associated vertical heat flux are well captured for all resolutions considered, with little resolution sensitivity. The convective heat flux is roughly

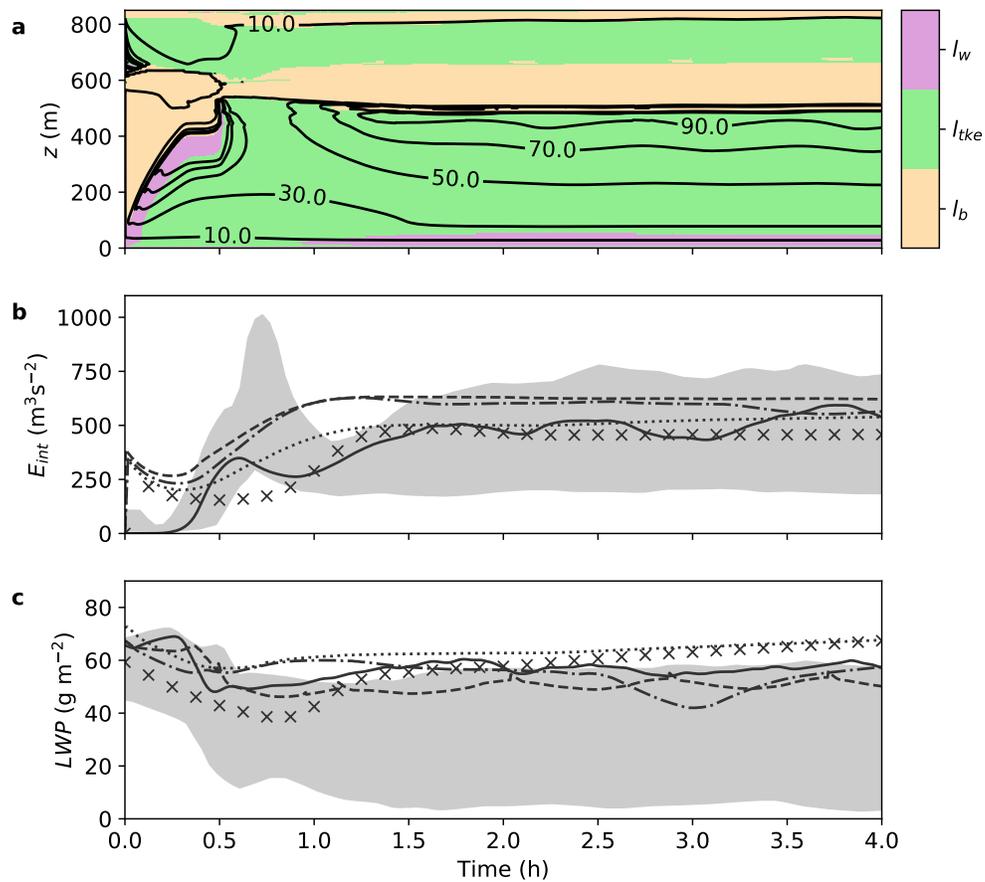


Figure 3.5: (a) Contours of eddy diffusivity  $K_m$  (m<sup>2</sup> s<sup>-1</sup>) as a function of time and height for the DYCOMS-II RF01 simulation using the SCM with  $\Delta z = 5$  m. Colors show the dominant (minimum) mixing length. (b) Time series of vertically integrated TKE  $E_{int}$ . (c) Time series of liquid water path (LWP). In (b) and (c), results are shown for LES (solid line), EDMF with  $\Delta z = 5$  m (dashed line),  $\Delta z = 20$  m (dash-dotted line),  $z = 50$  m (dotted line) and  $z = 75$  m (x).

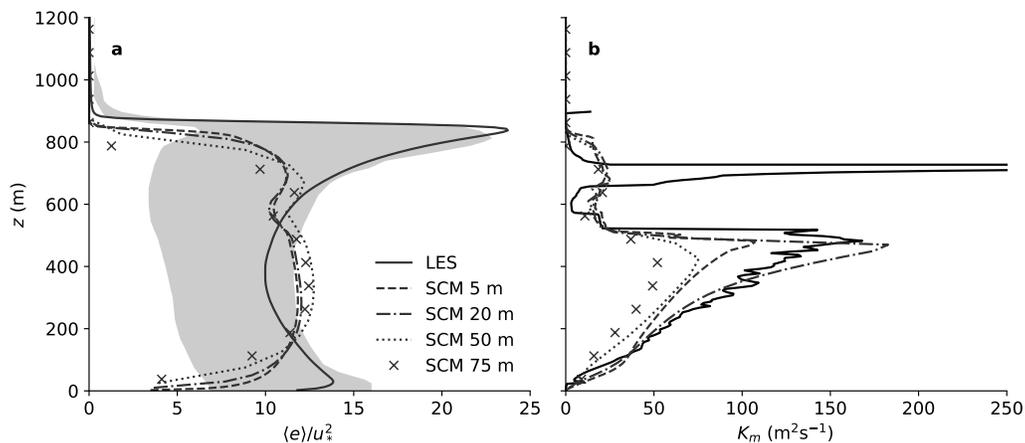


Figure 3.6: Profiles of (a) normalized turbulence kinetic energy and (b) eddy diffusivity. Profiles averaged over the 4th hour of the DYCOMS-II RF01 simulation. In (a),  $u_*$  is the average friction velocity during the 4th hour, with  $u_* = 0.22 \text{ m s}^{-1}$  for LES and  $u_* = 0.26, 0.25, 0.24,$  and  $0.24 \text{ m s}^{-1}$  for the SCM cases, in order of increasing  $\Delta z$ . In (b), the eddy diffusivity diagnosed from LES follows (3.45). The shaded region represents the spread of LES results reported in Stevens et al. (2005). Plotting conventions follow the legend in panel (a).

constant throughout the boundary layer, while the diffusive flux decreases with height.

All simulations show a small cold bias throughout the boundary layer and a warm bias below the inversion. The latter is due to the absence of plume overshooting in the SCM simulations, as shown in Figure 3.7b. The evolution of the boundary layer depth, diagnosed as the height of minimum buoyancy flux (Stevens, 2007), is shown in Figure 3.7d. The boundary layer growth in the SCM simulations is slower than in LES, with the bias decreasing as the vertical resolution is refined.

Reducing these biases with the extended EDMF scheme is possible, albeit with a different set of parameters controlling the pressure closure (3.11). These results are not shown here, since the goal of the model is to simulate all boundary layer regimes with a single set of parameters. Learning a set of parameters that minimizes these and other biases in the results shown here and in Cohen et al. (2020) is left for future work.

Finally, Figure 3.8 shows vertical profiles of TKE and eddy diffusivity. All SCM simulations display a TKE bias above 1500 m due to the absence of plume overshooting, which leads to zero convective TKE following equation (3.5). Eddy diffusion, which is a function of environmental TKE in our model, is not affected by this

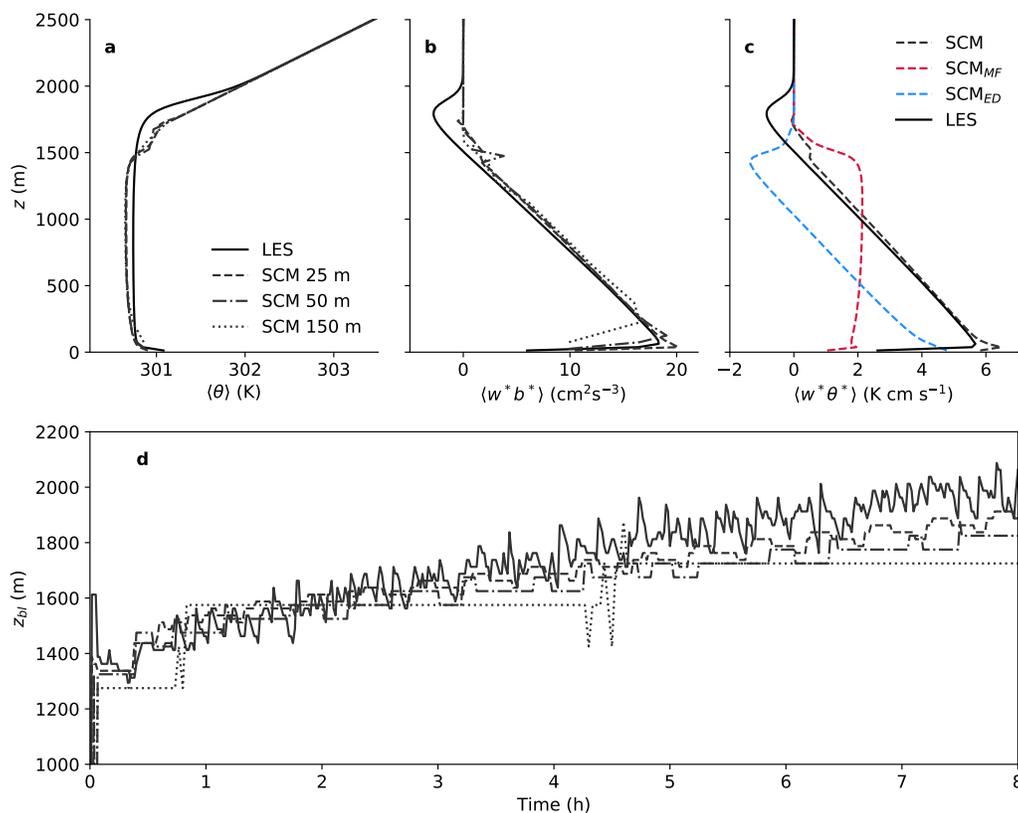


Figure 3.7: Profiles of (a) potential temperature, (b) vertical buoyancy flux, and (c) vertical flux of potential temperature. Profiles averaged over the 5th hour of dry convection simulation. In (c), the eddy diffusivity ( $SCM_{ED}$ ) and mass flux ( $SCM_{MF}$ ) components of the vertical flux are shown separately. (d) Time evolution of the boundary layer depth  $z_{bl}$ , computed as the level of minimum buoyancy flux  $\langle w^* b^* \rangle$ . Plotting conventions follow the legend in panel (a).

convective bias. Indeed, the diffusive closure (3.9) leads to an accurate prediction of the depth of the diffusive layer, as shown in Figure 3.8b.

### 3.6 Summary and discussion

The mixing length formulation proposed in this study provides a regime-independent closure of turbulent fluxes for EDMF schemes. The results for the stable boundary layer, stratocumulus-topped boundary layer, and dry convection demonstrate the ability of EDMF schemes with this mixing length closure to accurately describe the structure of the boundary layer in regimes where existing parameterizations currently in use in climate models fail or are inaccurate.

In the stable boundary layer, where convection and the subdomain decomposition in the EDMF scheme do not play a role, the proposed closure is able to reproduce

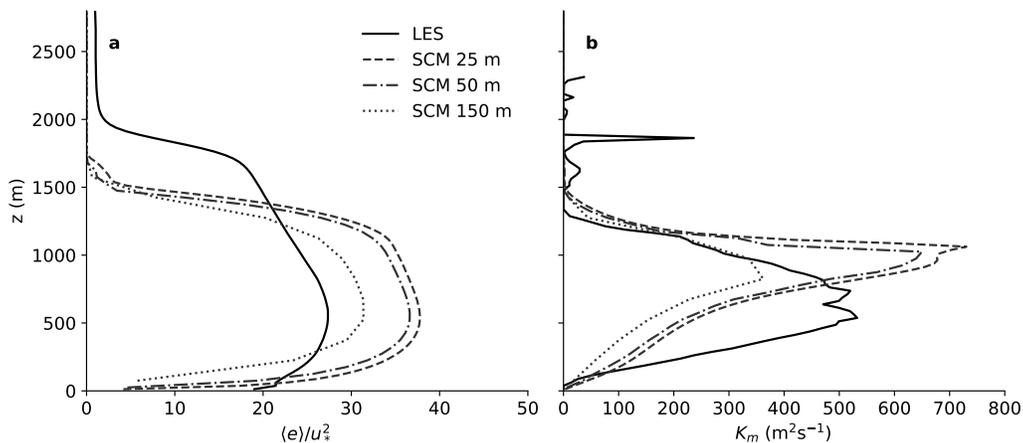


Figure 3.8: Profiles of (a) normalized turbulence kinetic energy and (b) eddy diffusivity for the dry convective boundary layer. Profiles averaged over the 5th hour of simulation. In (a),  $u_*$  is the average friction velocity during the 5th hour, with  $u_* = 0.18 \text{ m s}^{-1}$  for LES and  $u_* = 0.24 \text{ m s}^{-1}$  for all SCM cases. In (b), the eddy diffusivity diagnosed from LES follows (3.45). Plotting conventions follow the legend in panel (a).

the vertical structure and time evolution of turbulence over a range of vertical resolutions, down to  $O(10 \text{ m})$ . In the stratocumulus-topped boundary layer, where convective fluxes do play a role, the transport owing to environmental diffusion still provides the leading-order contribution to the subgrid-scale vertical fluxes in our EDMF scheme. The way in which environmental stratification limits the mixing length seems to be the crucial feature that allows our EDMF scheme to reproduce the sharp inversion at the stratocumulus cloud-top, even at relatively coarse vertical resolution.

Several characteristics differentiate this closure from others proposed in the literature. First, choosing the smooth minimum (3.40) of various candidate mixing lengths is consistent with the idea that estimates of the mixing length arising from different physical arguments should converge to a similar master length scale if they are simultaneously valid. For widely used expressions such as the harmonic mean (3.28), this does not hold, leading to unrealistic reductions in mixing. Second, our formulation explicitly links the eddy diffusivity to the effect of convective cells on the environment, leading to a consistently closed TKE balance. This results, for example, in the TKE injection term  $\mathcal{I}$  appearing in the length scale (3.18), for which TKE production and dissipation are in balance. Third, the mixing length does not depend on integral quantities such as the boundary layer thickness or Dear-

dorff’s convective scale. The inclusion of these terms in other models often leads to regime-dependent nonlocal terms that are non-causal and hence difficult to justify in general. Finally, the closure smoothly connects with Monin-Obukhov similarity theory near the surface with no assumptions about the height at which the transition occurs. This is particularly relevant for climate models with low vertical resolution, for which the use of similarity theory even in the first model level may be inaccurate.

A similar approach to the one shown here may be used to develop increasingly complex closures for high-order turbulence models. As an example, the net dissipation argument used in the TKE production–dissipation inequality (3.13) can also be applied to the temperature variance budget to diagnose the turbulent Prandtl number. The same could be done for other second-moment budgets in models with additional second-order prognostic equations, to obtain independent diffusivities for different tracers.

Finally, the optimization of the full parameter space was beyond the scope of this study and is left for future work. The access to LES data for a wider range of atmospheric conditions is necessary to enable a more comprehensive optimization of the parameter space in the EDMF scheme.

### 3.7 Appendix A: Smooth minimum function

We define as a smooth minimum any function  $f : \mathbb{R}^N \rightarrow \mathbb{R}$  of differentiability class  $C^\infty$  that approximates the  $\min(\mathbf{x})$  operator. Our implementation of (3.40) is based on the softmin function  $\mathbf{s}^\Lambda(\mathbf{x})$ , which is a smooth approximation to the  $\operatorname{argmin}(\mathbf{x})$  function (Titsias, 2016), with

$$\mathbf{s}_j^\Lambda(\mathbf{x}) = \frac{e^{-x_j/\Lambda}}{\sum_{i=1}^N e^{-x_i/\Lambda}}. \quad (3.46)$$

Here,  $\Lambda$  is a regularization parameter. The inner product of  $\mathbf{s}^\Lambda$  with  $\mathbf{x}$  yields an approximation of the  $\min(\mathbf{x})$  operator,

$$\sigma^\Lambda(\mathbf{x}) = \mathbf{x} \cdot \mathbf{s}^\Lambda(\tilde{\mathbf{x}}) = \frac{\sum_{i=1}^N x_i e^{-\tilde{x}_i/\Lambda}}{\sum_{i=1}^N e^{-\tilde{x}_i/\Lambda}}, \quad (3.47)$$

where  $\tilde{x}_j = x_j - \min(\mathbf{x})$ . In (3.47), the translational symmetry  $\mathbf{s}^\Lambda(\mathbf{x}) = \mathbf{s}^\Lambda(\mathbf{x} + \mathbf{c})$  with  $c_j = c$  is used to avoid errors due to finite precision arithmetic. The function  $\sigma^\Lambda(\mathbf{x})$  converges to  $\min(\mathbf{x})$  as  $\Lambda \rightarrow 0$ . In practice, a nonzero regularization parameter is chosen to ensure smoothness. The value of  $\Lambda$  may be obtained by imposing a

monotonically decreasing contribution of each  $x_j$  to  $\sigma^\Lambda(\mathbf{x})$ :

$$\frac{\partial \left( x_j s_j^\Lambda(\mathbf{x}) \right)}{\partial x_j} \leq 0 \quad \forall x_j \text{ if } \Lambda \leq \min(\mathbf{x}). \quad (3.48)$$

Alternatively,  $\Lambda$  may be defined by enforcing an upper bound on the value of  $\sigma^\Lambda(\mathbf{x})$  under certain assumptions about  $\mathbf{x}$ . Let the elements of  $\mathbf{x}$  be ordered such that  $x_j \leq x_{j+1}$  for all  $j$ . Assuming  $x_2 \approx x_n \ll x_{n+1}$ , then

$$\sigma^\Lambda(\mathbf{x}) \leq \min(\mathbf{x})(1 + \epsilon) \quad \text{if } \Lambda \leq \Lambda_0 = \frac{\epsilon}{W\left(\frac{n-1}{e}\right)} \min(\mathbf{x}), \quad (3.49)$$

where  $W(x)$  is the Lambert  $W$  function. In the implementation of (3.40), we use

$$s_{\min}(\mathbf{x}) = \sigma^{\Lambda_0^*}(\mathbf{x}) \quad \text{with } \epsilon = 0.1. \quad (3.50)$$

Here,  $\Lambda_0^* = \max(\Lambda_0, 1 \text{ m})$ , so the smoothing scale is constrained to be at least 1 m. Although a large value of  $n$  results in a closer approximation to the minimum, (3.46) may become difficult to evaluate in finite precision arithmetic. Because of the low dimensionality of  $\mathbf{x}$  in (3.40) and the limitation given by  $\Lambda_0^*$ , finding a compromise is not necessary, and we set  $n = N$ .

### 3.8 Appendix B: Subgrid kinetic energy in the extended EDMF scheme

According to the TKE decomposition (3.5), the grid-mean TKE includes the environmental TKE and the subgrid kinetic energy of the plumes, or convective TKE. The environmental TKE equation in flux form reads

$$\begin{aligned} & \frac{\partial(\rho a_0 \bar{e}_0)}{\partial t} + \nabla_h \cdot (\rho a_0 \langle \mathbf{u}_h \rangle \bar{e}_0) + \frac{\partial(\rho a_0 \bar{w}_0 \bar{e}_0)}{\partial z} = - \frac{\partial(\rho a_0 \overline{w'_0 e'_0})}{\partial z} + \rho a_0 \overline{w'_0 b'_0} \\ & - \rho a_0 \left[ \overline{w'_0 u'_0} \frac{\partial \langle u \rangle}{\partial z} + \overline{w'_0 v'_0} \frac{\partial \langle v \rangle}{\partial z} + \overline{w'_0 w'_0} \frac{\partial \bar{w}_0}{\partial z} \right] + \rho \sum_{i>0} a_i (\bar{w}_i^* - \bar{w}_0^*) \frac{\partial \bar{\Psi}_i^*}{\partial z} \\ & + \sum_j \left[ E_{0j} \frac{1}{2} (\bar{w}_j - \bar{w}_0)^2 - (\Delta_{0j} + \hat{E}_{0j}) \bar{e}_0 - \hat{E}_{0j} \bar{w}_0^* (\bar{w}_j^* - \bar{w}_0^*) \right] - \rho a_0 \mathcal{D} \\ & - \nabla_h \cdot \left( \rho a_0 \overline{\mathbf{u}'_{h,0} e'_0} \right) - \rho a_0 \left( \overline{\mathbf{u}'_{h,0} u'_0} \cdot \nabla_h \langle u \rangle + \overline{\mathbf{u}'_{h,0} v'_0} \cdot \nabla_h \langle v \rangle + \overline{\mathbf{u}'_{h,0} w'_0} \cdot \nabla_h \bar{w}_0 \right). \end{aligned} \quad (3.51)$$

The prognostic equation for the convective kinetic energy in subdomain  $i$  can be obtained as

$$\frac{1}{2} \frac{\partial(\rho a_i \bar{w}_i^{*2})}{\partial t} = \bar{w}_i^* \left[ \frac{\partial(\rho a_i \bar{w}_i)}{\partial t} - a_i \frac{\partial \rho \langle w \rangle}{\partial t} - \langle w \rangle \frac{\partial(\rho a_i)}{\partial t} + a_i \langle w \rangle \frac{\partial \rho}{\partial t} - \frac{\bar{w}_i^*}{2} \frac{\partial(\rho a_i)}{\partial t} \right]. \quad (3.52)$$

Summing over all subdomains, we obtain the subgrid-scale convective TKE balance

$$\begin{aligned}
& \frac{1}{2} \sum_i \left[ \frac{\partial(\rho a_i \bar{w}_i^{*2})}{\partial t} + \nabla_h \cdot (\rho a_i \langle \mathbf{u}_h \rangle \bar{w}_i^{*2}) + \frac{\partial(\rho a_i \bar{w}_i \bar{w}_i^{*2})}{\partial z} \right] = - \sum_i \frac{\partial(\rho a_i \bar{w}_i^* \overline{w'_i w'_i})}{\partial z} \\
& + \sum_i \left[ -\nabla_h \cdot (\rho a_i \bar{w}_i^* \overline{\mathbf{u}'_{h,i} w'_i}) - \rho a_i \bar{w}_i^{*2} \frac{\partial \langle w \rangle}{\partial z} + \rho a_i \overline{w'_i w'_i} \frac{\partial \bar{w}_i^*}{\partial z} + \rho a_i \overline{\mathbf{u}'_{h,i} w'_i} \cdot \nabla_h \bar{w}_i^* \right] \\
& + \sum_i \left( \rho a_i \bar{b}_i^* \bar{w}_i^* - \rho a_i \frac{\partial \bar{\Psi}_i^*}{\partial z} \bar{w}_i^* \right) - \sum_i \sum_{j \neq i} \left[ E_{ij} \frac{(\bar{w}_j^* - \bar{w}_i^*)^2}{2} - \hat{E}_{ij} \bar{w}_i^* (\bar{w}_j^* - \bar{w}_i^*) \right], \tag{3.53}
\end{aligned}$$

where, under the EDMF assumptions, all terms involving within-subdomain covariances are only nonzero in the environment ( $i = 0$ ). The first and second terms on the right-hand side are turbulent transport terms. The third term represents shear production of convective energy. The fourth and fifth terms yield shear production of subdomain TKE by the convective flow, representing an advective sink in the balance (3.53). The sixth and seventh terms are the convective components of the buoyant production and velocity-pressure gradient terms. Finally, the dynamical and turbulent entrainment terms act to transfer subgrid kinetic energy from the plumes to within-subdomain variance. Note that the velocity-pressure gradient term can be rewritten as

$$\sum_{i \geq 0} \rho a_i \frac{\partial \bar{\Psi}_i^*}{\partial z} \bar{w}_i^* = \sum_{i > 0} \rho a_i \frac{\partial \bar{\Psi}_i^*}{\partial z} (\bar{w}_i^* - \bar{w}_0^*), \tag{3.54}$$

since  $\sum_i a_i \partial \bar{\Psi}_i^* / \partial z = 0$ . This yields the definition of the pressure work on the plumes used in expression (3.10).

Some of the terms in budgets (3.51) and (3.53) transfer subgrid energy among the environment and plumes, resulting in a null contribution to the global budget. The grid-mean TKE prognostic equation that results from their sum is

$$\begin{aligned}
& \frac{\partial(\rho \langle e \rangle)}{\partial t} + \nabla_h \cdot (\rho \langle \mathbf{u}_h e \rangle) + \frac{\partial(\rho \langle w e \rangle)}{\partial z} = \rho \langle w^* b^* \rangle - \rho a_0 \mathcal{D} \\
& - \rho \left( \langle w^* u^* \rangle \frac{\partial \langle u \rangle}{\partial z} + \langle w^* v^* \rangle \frac{\partial \langle v \rangle}{\partial z} + \langle w^* w^* \rangle \frac{\partial \langle w \rangle}{\partial z} \right) \\
& - \rho \left( \langle \mathbf{u}_h^* u^* \rangle \cdot \nabla_h \langle u \rangle + \langle \mathbf{u}_h^* v^* \rangle \cdot \nabla_h \langle v \rangle + \langle \mathbf{u}_h^* w^* \rangle \cdot \nabla_h \langle w \rangle \right), \tag{3.55}
\end{aligned}$$

where there is no contribution from pressure-velocity correlations in our model. The evolution of the grid-mean TKE under the net dissipation closure (3.13) can be

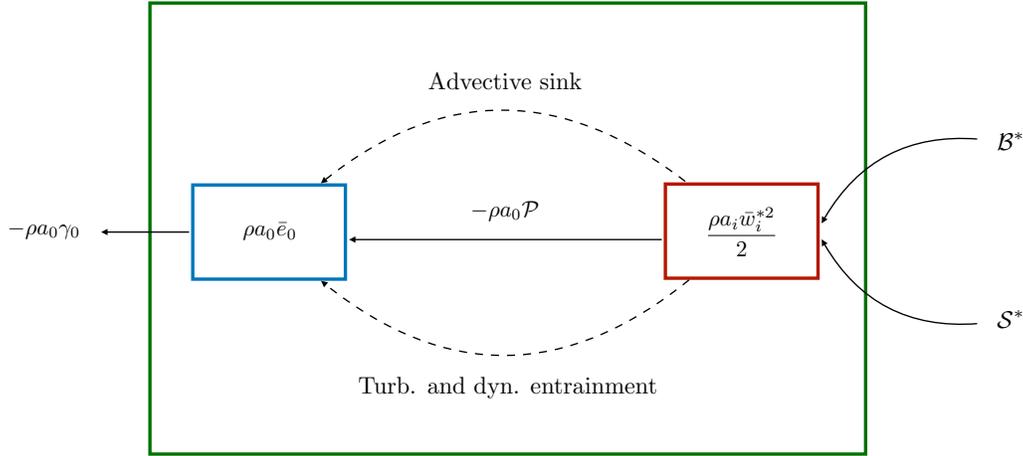


Figure 3.9: Schematic of subgrid kinetic energy reservoirs and pathways in the extended EDMF scheme under the net dissipation closure. Notation follows equation (3.56). Dashed lines represent energy pathways that result in implicit grid-mean TKE dissipation under the net dissipation closure (3.13). Summation over  $i = 0, \dots, n$  is implied.

obtained by subtracting (3.13) from (3.55):

$$\begin{aligned}
 \frac{\partial(\rho\langle e \rangle)}{\partial t} + \nabla_h \cdot (\rho\langle \mathbf{u}_h e \rangle) + \frac{\partial(\rho\langle we \rangle)}{\partial z} &= \underbrace{\sum_i \rho a_i \bar{b}_i^* \bar{w}_i^*}_{\mathcal{B}^*} - \underbrace{\sum_i \rho a_i \bar{w}_i^{*2} \frac{\partial\langle w \rangle}{\partial z}}_{\mathcal{S}^*} - \rho a_0 \gamma_0 \\
 + \underbrace{\rho a_0 \left( \overline{w_0'^2} \cdot \frac{\partial \bar{w}_0^*}{\partial z} + \overline{\mathbf{u}'_{h,0} w_0'} \cdot \nabla_h \bar{w}_0^* \right)}_{\text{Advection sink}} &- \underbrace{\sum_{i>0} \left[ E_{0i} \frac{(\bar{w}_i^* - \bar{w}_0^*)^2}{2} - \hat{E}_{i0} \bar{w}_0^* (\bar{w}_i^* - \bar{w}_0^*) \right]}_{\text{Turb. and dyn. entrainment}},
 \end{aligned} \tag{3.56}$$

where  $\gamma_0$  is the net environmental dissipation. According to (3.56), grid-mean TKE is generated through convective buoyant production  $\mathcal{B}^*$  and the vertical convergence term  $\mathcal{S}^*$ . Both dynamical and turbulent entrainment act as a transfer term from subgrid-scale convective kinetic energy to environmental TKE, resulting in a grid-mean TKE sink under the net dissipation closure. A schematic of the energetic pathways between budgets (3.51) and (3.53) and the overall evolution of grid-mean TKE under the mixing length closure presented here is shown in Figure 3.9.

### 3.9 Appendix C: Probabilistic model for cloud fraction

We consider  $\theta_l$  and  $q_t$  to be log-normally distributed with expected values  $\bar{\theta}_{l,0}$  and  $\bar{q}_{t,0}$ , variances  $\sigma_{\theta_l}^2$  and  $\sigma_{q_t}^2$ , and covariance  $\sigma_{q_t, \theta_l}^2$ . The log-normal distribution

is preferred over the commonly used Gaussian distribution (e.g., Sommeria and Deardorff, 1977) for two reasons: both  $\theta_l$  and  $q_t$  remain non-negative, and positive skewness is allowed. Under the Gaussian assumption, negative values of  $q_t$  may be drawn from the distribution if  $\sigma_{q_t}^2/\bar{q}_{t,0}^2$  is not sufficiently small (Mellor, 1977). In addition, distributions with positive skewness have been shown to capture the development of cumulus convection better (Bougeault, 1981).

The expected value of cloud fraction  $f_{c,0}$  can be computed as (Mellor, 1977)

$$f_{c,0} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(q_l(\theta_l, q_t)) p(\theta_l, q_t) d\theta_l dq_t, \quad (3.57)$$

where  $H(\cdot)$  is the Heaviside function and  $p(\theta_l, q_t)$  is the probability density function (PDF) of the log-normal bivariate distribution with marginal probability density functions (PDFs) given by

$$q_t \sim \text{LN}(\mu_{q_t}, s_{q_t}^2), \quad \mu_{q_t} = \ln \left( \frac{\bar{q}_{t,0}^2}{\sqrt{\sigma_{q_t}^2 + \bar{q}_{t,0}^2}} \right), \quad s_{q_t}^2 = \ln \left( \frac{\sigma_{q_t}^2}{\bar{q}_{t,0}^2} + 1 \right) \quad (3.58)$$

and

$$\theta_l \sim \text{LN}(\mu_{\theta_l}, s_{\theta_l}^2), \quad \mu_{\theta_l} = \ln \left( \frac{\bar{\theta}_{l,0}^2}{\sqrt{\sigma_{\theta_l}^2 + \bar{\theta}_{l,0}^2}} \right), \quad s_{\theta_l}^2 = \ln \left( \frac{\sigma_{\theta_l}^2}{\bar{\theta}_{l,0}^2} + 1 \right). \quad (3.59)$$

The conditional PDF of  $\theta_l$  given  $q_t$  is the log-normal distribution

$$\theta_l | q_t \sim \text{LN}(\mu_c, s_c^2), \quad \mu_c = \mu_{\theta_l} + \frac{s_{\theta_l, q_t}^2}{s_{q_t}^2} (\ln(q_t) - \mu_{q_t}), \quad s_c^2 = s_{\theta_l}^2 - \frac{s_{\theta_l, q_t}^4}{s_{q_t}^2}, \quad (3.60)$$

where

$$s_{\theta_l, q_t}^2 = \ln \left( \frac{\sigma_{q_t, \theta_l}^2}{\bar{q}_{t,0} \bar{\theta}_{l,0}} + 1 \right). \quad (3.61)$$

The cloud fraction (3.57) can be calculated by Gaussian quadrature as

$$f_{c,0} \approx \frac{1}{\pi} \sum_i^{n_i} w_i \sum_j^{n_j} w_j H(q_l(\theta_{l,j}, q_{t,i})), \quad (3.62)$$

where  $w_j$  and  $w_i$  are the Gauss-Hermite weights corresponding to evaluation points  $\theta_{l,j}$  and  $q_{t,i}$ , respectively. The evaluation points  $(\theta_{l,j}, q_{t,i})$  of the log-normal distributions (3.58) and (3.60) are related to the Gauss-Hermite mass points  $(\xi_j, \chi_i)$  through the normal distributions  $x$  and  $y$  with same parameters:

$$\theta_{l,j} = e^{x_j}, \quad x_j = \mu_c + \sqrt{2} s_c \xi_j, \quad x \sim N(\mu_c, s_c^2) \quad (3.63)$$

and

$$q_{t,i} = e^{y_i}, \quad y_i = \mu_{q_t} + \sqrt{2}s_{q_t}\chi_i, \quad y \sim N(\mu_{q_t}, s_{q_t}^2). \quad (3.64)$$

Note that the evaluation points  $\theta_{l,j}$  are drawn from the conditional PDF (3.60). In (3.62), the liquid water specific humidity  $q_l$  is obtained as  $q_l = q_{t,i} - q_s(\theta_{l,j}, q_{t,i})$  for  $q_{t,i} > q_s(\theta_{l,j}, q_{t,i})$ , where  $q_s$  is the equilibrium saturation specific humidity. Thus, supersaturation is not considered and all excess water vapour is immediately converted to liquid water condensate. The equilibrium saturation specific humidity is found iteratively using a saturation adjustment procedure (see Bryan and Fritsch, 2002, for details). Consistent with this approach, the environmental liquid water specific humidity  $\bar{q}_{l,0}$  is computed as

$$\bar{q}_{l,0} = \frac{1}{\pi} \sum_i^{n_i} w_i \sum_j^{n_j} w_j [q_{t,i} - q_s(\theta_{l,j}, q_{t,i})] H(q_{t,i} - q_s(\theta_{l,j}, q_{t,i})). \quad (3.65)$$

In this study, the probabilistic cloud model is implemented using  $n_i = n_j = 3$ .

### 3.10 Appendix D: Results with alternative mixing length formulations

The importance of the different length scale restrictions in the mixing length closure (3.40) is shown in Figures 3.10, 3.11, and 3.12 for the simulation of the SBL, the STBL, and dry convection. We consider three alternative formulations to (3.40): one without wall constraints (no  $l_w$ ), one without stratification constraints (no  $l_b$ ), and another one without entrainment effects ( $\mathcal{I} = 0$  in  $l_{\text{tke}}$ ). These alternative formulations are denoted as NLW, NLB, and NED, respectively.

Simulations of the SBL and STBL are most strongly affected by the omission of stratification constraints, as shown in Figures 3.10 and 3.11. In the case of the STBL, omitting  $l_b$  in formulation (3.40) leads to a decoupling of the cloud layer from the boundary layer. In both dry convection and the stable boundary layer, the absence of  $l_b$  leads to a deeper boundary layer and stronger buoyancy fluxes near the inversion.

The effect of the entrainment term  $\mathcal{I}$  in the formulation of  $l_{\text{tke}}$  is noticeable in the buoyancy flux and LWP of the STBL, both of which show a positive bias with respect to LES in Figure 3.11. The effect of the entrainment term  $\mathcal{I}$  is more significant in the dry convective case, where its absence leads to very low diffusive mixing in the boundary layer and low TKE compared to LES. Finally, the results for NLW show that the implied balance in  $l_{\text{tke}}$  can approximate the wall constraints relatively well, leading only to small biases in the stable boundary layer simulation.

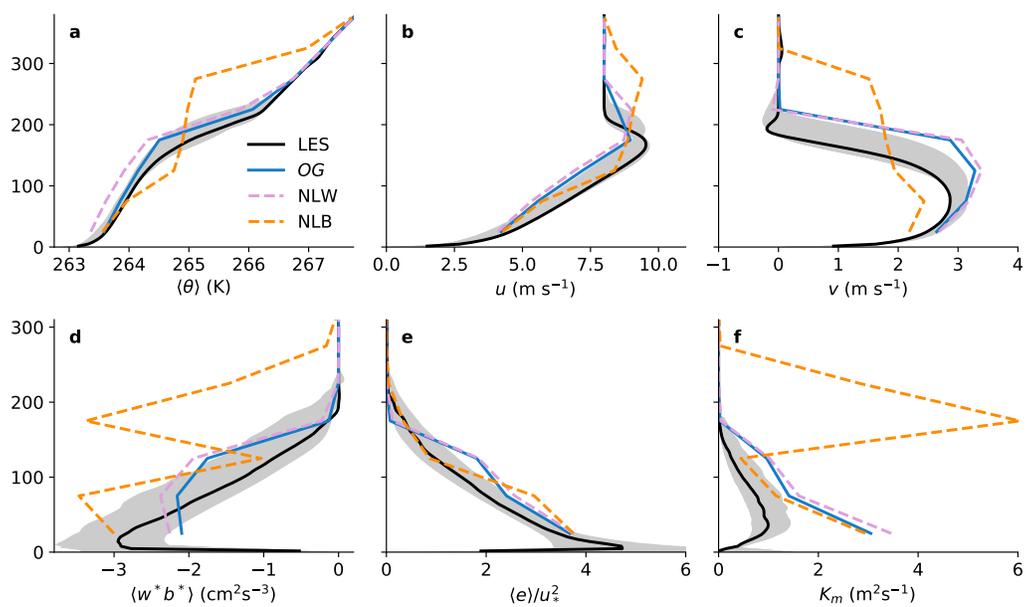


Figure 3.10: Profiles of (a) potential temperature, (b-c) horizontal velocity, (d) buoyancy flux, (e) normalized TKE, and (f) eddy diffusivity for the 9th hour of the GABLS simulation. Results shown for LES and for the SCM ( $\Delta z = 50$  m) with alternative mixing length formulations: original (OG), no  $l_w$  (NLW), and no  $l_b$  (NLB). Shaded region as in Figure 3.1.

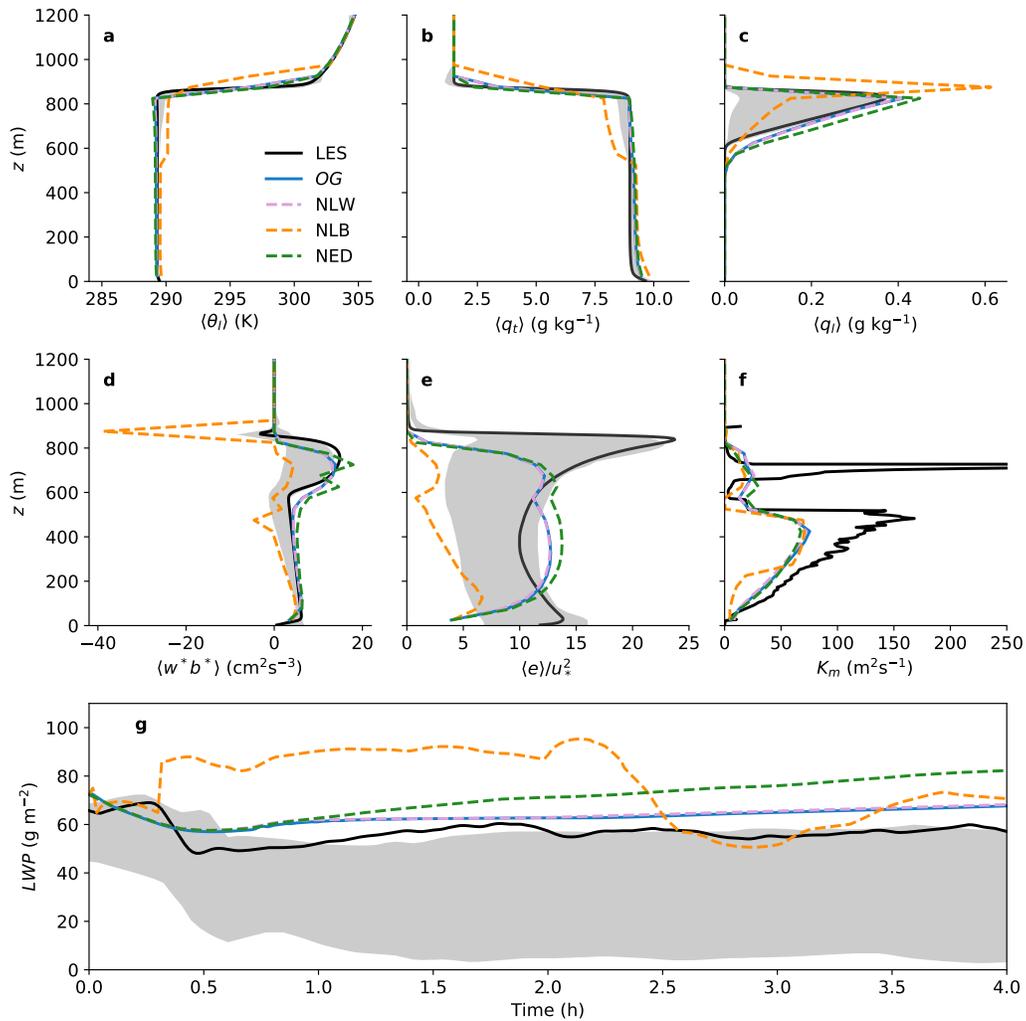


Figure 3.11: Profiles of (a) liquid water potential temperature, (b) total water specific humidity, (c) liquid water specific humidity, (d) buoyancy flux, (e) normalized TKE and (f) eddy diffusivity for the 4th hour of the DYCOMS-II RF01 simulation. The time evolution of the liquid water path is shown in (g). Results shown for LES and for the SCM ( $\Delta z = 50$  m) using alternative mixing length formulations: original (OG), no  $l_w$  (NLW), no  $l_b$  (NLB) and no entrainment term  $\mathcal{I}$  in equation (3.41) (NED). Shaded region as in Figure 3.4.

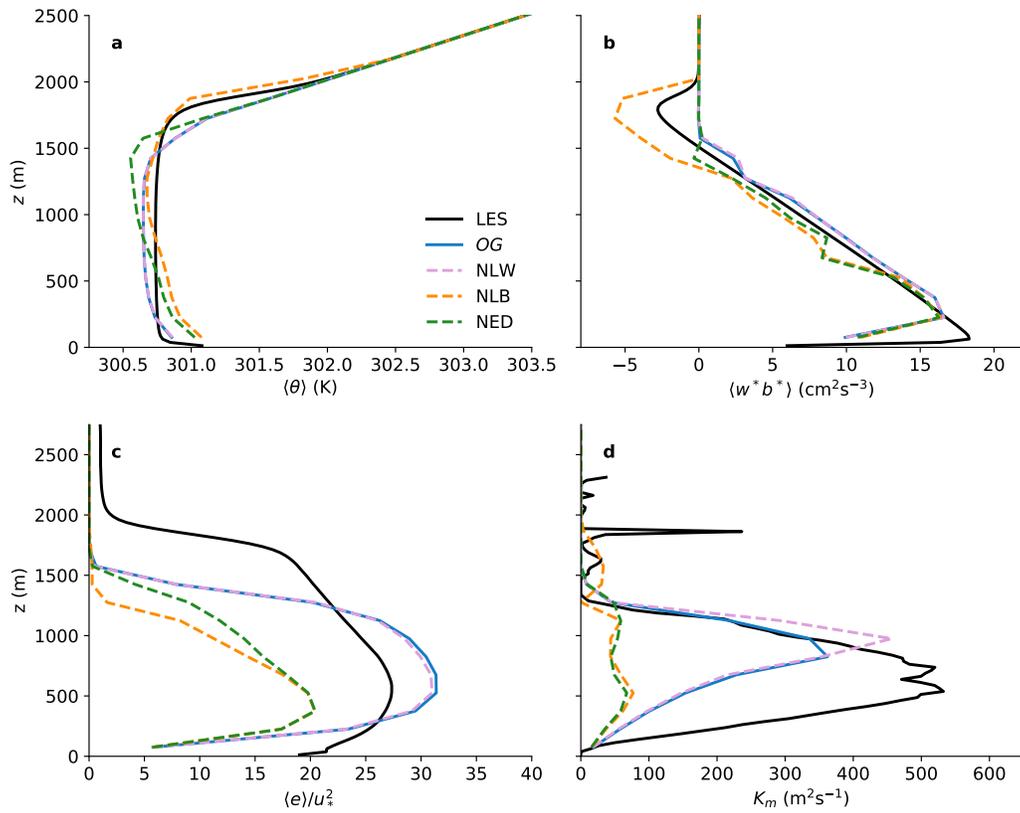


Figure 3.12: Profiles of (a) potential temperature, (b) buoyancy flux, (c) normalized TKE, and (d) eddy diffusivity for the 5th hour of the dry convection simulation. Results shown for LES and for the SCM ( $\Delta z = 150$  m) with alternative mixing length formulations.

*Chapter 4***TRAINING PHYSICS-BASED MACHINE-LEARNING  
PARAMETERIZATIONS WITH GRADIENT-FREE ENSEMBLE  
KALMAN METHODS**

Lopez-Gomez, Ignacio, Costa Christopoulos, Haakon Ludvig Langeland Ervik, Oliver R. A. Dunbar, Yair Cohen, and Tapio Schneider (2022). “Training physics-based machine-learning parameterizations with gradient-free ensemble Kalman methods”. In: *Journal of Advances in Modeling Earth Systems* 14, e2022MS003105. DOI: [10.1029/2022MS003105](https://doi.org/10.1029/2022MS003105).

**4.1 Abstract**

Most machine learning applications in Earth system modeling currently rely on gradient-based supervised learning. This imposes stringent constraints on the nature of the data used for training (typically, residual time tendencies are needed), and it complicates learning about the interactions between machine-learned parameterizations and other components of an Earth system model. Approaching learning about process-based parameterizations as an inverse problem resolves many of these issues, since it allows parameterizations to be trained with partial observations or statistics that directly relate to quantities of interest in long-term climate projections. Here we demonstrate the effectiveness of Kalman inversion methods in treating learning about parameterizations as an inverse problem. We consider two different algorithms: unscented and ensemble Kalman inversion. Both methods involve highly parallelizable forward model evaluations, converge exponentially fast, and do not require gradient computations. In addition, unscented Kalman inversion in its regularized version provides a measure of parameter uncertainty. We illustrate how training parameterizations can be posed as a regularized inverse problem and solved by ensemble Kalman methods through the calibration of an eddy-diffusivity mass-flux scheme for subgrid-scale turbulence and convection, using data generated by large-eddy simulations. We find the algorithms amenable to batching strategies, robust to noise and model failures, and efficient in the calibration of hybrid parameterizations that can include empirical closures and neural networks.

## 4.2 Introduction

The remarkable achievements of machine learning over the past decade have led to renewed interest in informing Earth system models with data (Schneider et al., 2017b; Reichstein et al., 2019). The spotlight is often on creating or improving models of processes that are deemed important for the correct representation of the Earth system as a whole. Examples of these processes include moist convection (Brenowitz et al., 2020), cloud microphysical and radiative effects (Seifert and Rasp, 2020; Villefranque et al., 2021; Meyer et al., 2022), and evapotranspiration (W. L. Zhao et al., 2019), among others.

Processes governed by poorly understood dynamics, such as biological processes, are obvious candidates for representation by purely data-driven models. On the other end of the spectrum are fluid transport processes, which are governed by the Navier-Stokes equations. Uncertain representation of these processes comes from a lack of resolution, not lack of knowledge about the underlying dynamics. Hybrid modeling approaches that incorporate domain knowledge and augment it by learning from data are attractive for such processes, because they reduce what needs to be learned from data.

For processes with known dynamics, data-informed models fall into three broad categories according to their leverage of domain knowledge. In the first category are models that try to learn the entire dynamics using a sufficiently expressive hypothesis set, such as deep neural networks. This approach has proved successful for predicting precipitation over short time horizons (Ravuri et al., 2021), and it has been explored for medium-range weather forecasting (Rasp and Thuerey, 2021; Pathak et al., 2022; Lopez-Gomez et al., 2022a). An advantage of these models is that they are typically easy to implement and cheap to evaluate. They can afford very large time steps (Weyn et al., 2021), or they may learn directly mappings from the initial state to a probability distribution of final states with no need of time marching or ensemble forecasting (Sønderby et al., 2020). A deficiency of these models is that they often require an extreme amount of data to constrain the many (often  $> 10^6$ ) parameters in them and to achieve acceptable performance.

Methods in the second and third categories employ models of subgrid processes to solve the closure problem that arises when coarse-graining the known dynamics, which are retained. Retaining the coarse-grained equations of motion ensures conservation of mass, momentum, and energy, which is more difficult when using models in the first category (Beucler et al., 2021; Brenowitz et al., 2020). The

second category encompasses methods that try to learn the functional form of these closures avoiding the use of empirical laws. For example, Zanna and Bolton (2020) use relevance vector machines to prune a library of functions, resulting in a closed form expression of mesoscale eddy fluxes in ocean simulations; Ling et al. (2016) learn a neural network closure of the Reynolds stress anisotropy tensor while explicitly encoding rotational invariance in the context of  $k - \epsilon$  models of turbulence.

Finally, the third category refers to methods that seek to learn the parameters that arise in empirical closures of subgrid processes. In general, models in the third category are more restrictive, and they may be expected to underperform with respect to those in the second category given sufficient data on the target distributions. However, the limited parametric complexity of these closures makes them amenable to physical interpretation, robust to overfitting, and better suited for learning in the low-data regime. This may be attractive for Earth system models, for which online learning from limited high-resolution data may be a useful strategy to assimilate computationally generated data of the changing climate (Schneider et al., 2017b).

A barrier delimiting data-driven and empirical subgrid-scale closures is the access to practical calibration tools. Neural network parameterizations are easily calibrated using stochastic gradient descent through backpropagation, which limits datasets to those including output labels, and models to those that afford automatic differentiation with respect to their parameters. Empirical closures, which may depend on time-evolving terms with memory (e.g., Lopez-Gomez et al., 2020) or yield unobservable outputs (e.g., turbulent versus dynamical entrainment in Cohen et al., 2020) cannot be trained using this approach. Traditional Bayesian inference techniques, like random walk Metropolis (Metropolis et al., 1953) or sequential Monte Carlo (Moral et al., 2006), can be used in this context if the number of parameters is small and the model to be trained is cheap to evaluate. Such methods additionally provide uncertainty quantification, but they become intractable for expensive models with many parameters (e.g., Cotter et al., 2013; Souza et al., 2020). Model-agnostic tools that enable fast calibration of subgrid-scale closures from diverse data are a necessary step toward the development of hybrid closures that leverage the strengths of all modeling approaches.

With this goal in mind, we present calibration strategies for models of subgrid processes, formulating the learning task as an inverse problem (Kovachki and Stuart, 2019). Solutions to the inverse problem are sought using the ensemble and unscented Kalman inversion algorithms (Iglesias et al., 2013; D. Z. Huang et al., 2022a).

Emphasis is given to practical aspects of this specific inverse problem, which have not previously been explored in the literature. These include the construction of a domain-agnostic loss function from high-dimensional observations, a heuristic a priori estimate of model error, systematic handling of model failures during the training process, and the use of the Kalman inversion algorithms when only noisy evaluations of the loss function are available.

The strategies presented here are designed to have several attractive properties compared to other learning algorithms. First, framing learning as an inverse problem enables the use of partial observations or statistically aggregated data. Second, calibration is performed using gradient-free methods, which are well suited for stochastic models and/or models whose derivatives do not exist or are difficult to obtain. Finally, the strategies presented are amenable to parallelization and the use of high-dimensional correlated observations. The last two properties draw heavily on the use of Kalman inversion algorithms to tackle the inverse problem, which themselves build on the success of the ensemble Kalman filter (EnKF) for data assimilation (Evensen, 1994; Houtekamer and Mitchell, 1998; Burgers et al., 1998) and are closely related to iterative EnKF (Chen and Oliver, 2012; Emerick and Reynolds, 2013; Bocquet and Sakov, 2013). The methods presented here are applicable to models of subgrid-scale processes, within the second and third categories described above. They provide an alternative to learning algorithms that impose stringent requirements on either the model architecture, its computational cost, or the nature of the training data.

The article is organized as follows. Section 4.3 casts learning about parameterizations as an inverse problem, which can be solved through the minimization of a regularized low-dimensional encoding of the data-model mismatch. Section 3 reviews the application of the ensemble and unscented Kalman inversion algorithms to inverse problems and proposes modifications to their update equations that enable training models that may experience failures. Section 4 then applies these ensemble Kalman algorithms to the calibration of closures within an eddy-diffusivity mass-flux (EDMF) scheme of turbulence and convection, using data generated from large-eddy simulations (LES). The robustness of these learning strategies is demonstrated by calibrating the EDMF scheme using noisy loss evaluations and partial information, and their flexibility is emphasized by learning the parameters in a hybrid model containing both empirical and neural network closures. Finally, Section 5 ends with a discussion of the findings and concluding remarks.

### 4.3 Learning about parameterizations as an inverse problem

We consider the problem of learning the parameters  $\phi$  of a dynamical model  $\Psi(\phi)$ , using noisy observations  $y$  of the true dynamical system  $\zeta$  that  $\Psi(\phi)$  seeks to represent. In the context of subgrid parameterizations,  $\Psi(\phi)$  represents a closed version of the coarse-grained dynamical system (e.g., the filtered Navier-Stokes equations), where closures are parameterized by  $\phi$ . The model  $\Psi(\phi)$  maps a user-defined initial state  $\varphi_0$  and a forcing  $F_\varphi(t)$  to a state trajectory  $\hat{\varphi}(t)$ . Thus, our definition of  $\Psi(\phi)$  can be interpreted as the iterative application of the resolvent operator on the initial field  $\varphi_0$  (Brajard et al., 2021). In the following, we denote any set of initial and forcing conditions collectively as the configuration  $x_c = \{\varphi_0, F_\varphi\}_c$ ; the definition of all symbols is summarized in the appendix.

For each configuration  $x_c$ , the dynamical model can be related to the observations  $y_c$  by the observational map  $\mathcal{H}_c$ , which encapsulates all averaging and post-processing operations necessary to yield the model predictions associated with the observations. More precisely, the relationship between the observations  $y_c$ , the true dynamics  $\zeta$ , and the dynamical model  $\Psi(\phi)$  for a given configuration may be expressed as

$$y_c = \mathcal{H}_c \circ \zeta(x_c) + \eta_c = \mathcal{H}_c \circ \Psi(\phi; x_c) + \delta(x_c) + \eta_c, \quad (4.1)$$

where  $\phi \in \mathbb{R}^p$  is the vector of learnable parameters,  $\eta_c$  is the observational noise associated with  $y_c$ , and  $\delta(x_c)$  is the model or representation error, which we define as the mismatch between the denoised observations  $\mathcal{H}_c \circ \zeta(x_c)$  and the output of a best-fitting model  $\mathcal{H}_c \circ \Psi(\phi^*; x_c)$ , following Kennedy and O’Hagan (2001). Thus, the model error is approximated as additive (Cohn, 1997; van Leeuwen, 2015) and defined with respect to the observational map  $\mathcal{H}_c$  and the optimal parameters  $\phi^*$  that minimize its contribution to the data-model relation (4.1).

Observations are taken to come from finite spatial and temporal averages of fields such as temperature. Learning from averages can help prevent overfitting to trajectories in chaotic systems by focusing on the statistics of the dynamics (Morzfeld et al., 2018). It also improves numerical stability when coupling to a parent model (Brenowitz and Bretherton, 2018). Under this definition of observations, it is reasonable to assume the noise  $\eta_c$  to be additive and Gaussian. In the following, we will further consider  $\delta(\cdot)$  to be a centered Gaussian, although this constitutes a significantly stronger assumption (e.g., that the model is unbiased) and may not be appropriate for a detailed characterization of posterior uncertainty (van Leeuwen, 2015; Brynjarsdóttir and O’Hagan, 2014). The construction of more precise error

models remains a challenge beyond the scope of this work. These assumptions enable us to write  $\delta(x_c) + \eta_c \sim \mathcal{N}(0, \Gamma_c)$ .

In general, we are interested in minimizing the mismatch between  $y_c$  and the model output for a wide range of configurations  $C = \{x_c, c = 1, \dots, |C|\}$  that are representative of the conditions in which the model will operate. This defines the global data-model relation

$$y = \mathcal{H} \circ \Psi(\phi) + \delta + \eta, \quad (4.2)$$

where  $y = [y_1, \dots, y_{|C|}]^T \in \mathbb{R}^d$ ,  $\delta = [\delta(x_1), \dots, \delta(x_{|C|})]^T$ ,  $\eta = [\eta_1, \dots, \eta_{|C|}]^T$ ,  $\mathcal{H} \circ \Psi(\phi) = [\mathcal{H}_1 \circ \Psi(\phi; x_1), \dots, \mathcal{H}_{|C|} \circ \Psi(\phi; x_{|C|})]^T$  and  $\delta + \eta \sim \mathcal{N}(0, \Gamma)$ . In addition, implicit in the definition of the dynamical model  $\Psi(\phi)$  is a discrete resolution  $\Delta$ . This dependence may be lifted if the closures are designed to be scale-aware or scale-independent, in which case the relation (4.2) should be augmented by stacking copies of  $y$  and evaluating  $\mathcal{H} \circ \Psi(\phi, \Delta_i)$  for different discretizations  $\Delta_i$ .

In practice, the parameters  $\phi$  are often defined over some subspace  $U \subset \mathbb{R}^P$  outside of which the model trajectories are unphysical or numerically unstable. Examples of these are parameters controlling the diffusion or turbulent dissipation of a scalar field, for which negative values are not physically valid. On the other hand, many algorithms designed to solve inverse problems assume  $\phi \in \mathbb{R}^P$ . This obstacle may be circumvented by defining a transformation  $\mathcal{T} : U \rightarrow \mathbb{R}^P$ , such that the global data-model relation (4.2) can be defined in an unconstrained parameter space,

$$y = \mathcal{G}(\theta) + \delta + \eta, \quad (4.3)$$

where

$$\mathcal{G} := \mathcal{H} \circ \Psi \circ \mathcal{T}^{-1}, \quad \phi = \mathcal{T}^{-1}(\theta). \quad (4.4)$$

In expressions (4.3) and (4.4),  $\theta \in \mathbb{R}^P$  is the parameter vector in unconstrained space and  $\mathcal{G} : \mathbb{R}^P \rightarrow \mathbb{R}^d$  is the map from transformed parameters to model predictions, which represents the forward model. The task of learning a set of model parameters  $\theta$  under relation (4.3) can be cast as the Bayesian inverse problem of finding the posterior (Kaipio and Somersalo, 2006; Tarantola, 2005; D. Z. Huang et al., 2022b)

$$\rho(\theta|y, \Gamma) = \frac{e^{-\mathcal{L}(\theta; y)}}{Z(y|\Gamma)} \rho_{\text{prior}}(\theta), \quad \mathcal{L}(\theta; y) = \frac{1}{2} \|y - \mathcal{G}(\theta)\|_{\Gamma}^2, \quad (4.5)$$

where  $Z(y|\Gamma)$  is a normalizing constant,  $\|\cdot\|_{\Gamma}^2$  denotes the Mahalanobis norm  $\langle \cdot, \Gamma^{-1} \cdot \rangle$ ,  $\mathcal{L}$  is the loss or negative log-likelihood, and  $\rho_{\text{prior}}(\theta)$  is the prior density. We stress that the posterior  $\rho(\theta|y, \Gamma)$  is conditioned on our approximation of the

noise  $\delta + \eta$ ; see Kennedy and O’Hagan (2001) for a discussion on the usefulness and caveats of such an approach. Given the inverse problem (4.3)–(4.5), we may be interested in finding the maximum a posteriori (MAP), approximations of the density  $\rho(\theta|y, \Gamma)$  around the MAP for uncertainty quantification, or simply the maximum likelihood estimator (MLE) if we have no prior information about  $\theta$ . Algorithms to perform these tasks are described in Section 4.4.

The error covariance  $\Gamma_c$  appearing in each model-data relation (4.1), and ultimately defining the inverse problem (4.3)–(4.5), is yet to be defined. In Section 4.3, we suggest an estimate of  $\Gamma_c$  relevant to the calibration of models with an unknown error structure  $\delta(\cdot)$ . In addition, the choice of observational map  $\mathcal{H}_c$  may not be evident when training dynamical models that aim to represent complex dynamical systems  $\zeta$  with many observable fields. Section 4.3 suggests a model-agnostic definition of  $\mathcal{H}_c$  that can be used to construct a regularized inverse problem.

### Estimate of noise covariances

Since the structure of the representation or model error  $\delta$  is unknown a priori, we must either parameterize it and calibrate it as well (Brynjarsdóttir and O’Hagan, 2014), or use a heuristic to capture its magnitude. Here, we follow the second route and offer a heuristic that has worked well for us in practice. If we take  $y_c = y_c(t)$  to be an observation of the true system in configuration  $x_c$  aggregated over a time interval  $[t, t + \tau]$ , we can write equation (4.1) as

$$y_c(t) - y_c(0) = \mathcal{H}_c \circ \Psi(\phi; x_c, t) - y_c(0) + \delta(x_c; t) + \eta_c(t). \quad (4.6)$$

If we further consider a model with no predictive power of the first kind (Lorenz, 1975; Schneider and Griffies, 1999), such that  $\mathcal{H}_c \circ \Psi(\phi; x_c, t) \approx y_c(0)$  for all times  $t$ , the covariance of (4.6) from  $t = 0$  to  $t = t_c \gg \tau$  reads

$$\Gamma_c = \text{Cov}(y_c) \approx \text{Cov}(\delta(x_c)) + \text{Cov}(\eta_c), \quad (4.7)$$

which yields an estimate of the aggregate noise  $\eta_c + \delta(x_c) \sim \mathcal{N}(0, \Gamma_c)$  from the variability of the observation  $y_c$  over a time interval  $[0, t_c]$ . For non-stationary conditions or finite-time averages,  $\Gamma_c$  depends on  $t_c$ . Estimating the magnitude of the aggregate noise from the internal variability of the true dynamics ensures that the loss or negative log-likelihood  $\mathcal{L}(\theta; y)$  penalizes models  $\Psi(\phi)$  that produce unrealistic outputs, and it represents a form of error inflation if the best-fitting model is expected to outperform the aforementioned unskillful model. The heuristic (4.7) is most appropriate when the dynamical model  $\Psi(\phi)$  is expressive enough to closely

replicate the initial observations  $y_c(0)$ , such that any mismatch in the initial condition can be lumped together with the observation error.

## Design of the observational map

### Application to problems with high-resolution data

High-resolution data are becoming increasingly common, from reanalysis products (Muñoz-Sabater et al., 2021), satellite imagery (Schmit et al., 2017), and partial differential equation (PDE) solvers such as LES (Shen et al., 2022). Although computationally generated and thus suffering from their own limitations (e.g., microphysical processes still need to be parameterized in LES), data from PDE solvers have some particularly desirable properties for the calibration of dynamical models:

- All variables appearing in the coarse-grained equations of motion are observable. As a consequence, the nature of the observational map  $\mathcal{H}$  used to constrain the model is largely a design choice.
- Data can be obtained systematically for all configurations  $x_c$  of interest, which may be chosen to minimize parameter uncertainty through active learning (Dunbar et al., 2022). In contrast, data drawn from physical measurements (e.g., field observations) are often sparse in the space of forcing and boundary conditions.

High-resolution data are often high-dimensional, which poses particular difficulties regarding the conditioning and tractability of linear systems of equations when solving inverse problems. The guidelines for the construction of the observational map  $\mathcal{H}$  presented here are tailored to solve these issues, with a focus on data from high-fidelity solvers.

### Model calibration

We define *model calibration* as the minimization of the mismatch between the observed dynamics and the dynamics induced by the model. We will use this definition to construct a domain-agnostic map  $\mathcal{H}$ . As an example, consider a system  $\zeta$  with coarse-grained dynamics

$$\frac{\partial \bar{\varphi}}{\partial t} + \bar{\mathbf{v}} \cdot \nabla \bar{\varphi} + \nabla \cdot (\overline{\mathbf{v}'\varphi'}) = F_\varphi, \quad (4.8)$$

where  $\overline{(\cdot)}$  denotes spatial filtering,  $(\cdot)'$  subfilter-scale fluctuations, and  $F_\varphi$  is the forcing. The field  $\bar{\mathbf{v}}$  is prescribed and  $\overline{\mathbf{v}'\varphi'}$  is the term parameterized in  $\Psi(\phi)$ . Let

$S(t) = [\bar{\varphi}(t), \overline{\mathbf{v}'\varphi'}(t)]^T$  be the true state augmented with subgrid-scale fluxes, and  $\hat{S}(t)$  the augmented state predicted by the model. For an incompressible fluid model,  $S(t)$  would contain the fluid momentum, energy, and the subgrid advective fluxes of these fields.

Model calibration then entails finding the minimizer of the expected state mismatch  $\mathbb{E}[\|\hat{S} - S\|]$  with respect to some norm and time interval, where the expectation is taken to allow for the calibration of stochastic models. Observations of the augmented state  $S(t)$ , which includes subgrid-scale fluxes, are not always available. Therefore, this definition of model calibration is representative of the ideal learning scenario. In scenarios where the full state is not observable, we will consider  $S(t)$  to be an *observed state* formed by all relevant observable spatial fields.

### Observations in physical space

Following our definition of model calibration, we preliminarily define the observations in the model-data relation (4.1) as finite-time averages of the normalized observed state  $s_c$  for a set of configurations  $C$ ,

$$\tilde{y}_c = \frac{1}{T_c} \int_{t_c - T_c}^{t_c} s_c(\tau) d\tau, \quad s_c = \begin{bmatrix} v_{c,1} \\ \dots \\ v_{c,n_c} \end{bmatrix} = \begin{bmatrix} V_{c,1}/\sigma_{c,1} \\ \dots \\ V_{c,n_c}/\sigma_{c,n_c} \end{bmatrix}, \quad c = 1, \dots, |C|, \quad (4.9)$$

where  $T_c$  is the averaging time,  $v_{c,j} \in \mathbb{R}^{h_c}$  are the normalized spatial fields comprising  $s_c$ ,  $V_{c,j}$  are the components of the state  $S_c$  prior to normalization,  $n_c$  is the number of fields observed in configuration  $x_c$ , and  $h_c$  is the number of degrees of freedom of each field. As an example, the first configuration's observed state  $S_1$  may include as fields atmospheric soundings of temperature and specific humidity ( $n_1 = 2$ ) measured at  $h_1$  vertical locations above the surface, and the second configuration's state  $S_2$  may include these fields as well as horizontal velocity profiles ( $n_2 = 4$ ), measured at  $h_2$  different locations. Normalization of the observed state  $S_c$  is performed using the pooled time standard deviation  $\sigma_{c,j}$  of each field  $V_{c,j}$ , with

$$\sigma_{c,j}^2 = h_c^{-1} \text{tr} \left[ \text{Cov}(V_{c,j}) \right]. \quad (4.10)$$

Covariances are computed over a time  $t_c \geq T_c$  following the heuristic of Section 4.3 to capture the expected magnitude of the data mismatch,

$$\text{Cov}(V_{c,j}) = \frac{1}{t_c} \int_0^{t_c} V_{c,j} V_{c,j}^T d\tau - \frac{1}{t_c^2} \left( \int_0^{t_c} V_{c,j} d\tau \right) \left( \int_0^{t_c} V_{c,j} d\tau \right)^T. \quad (4.11)$$

We resort to pooled normalization, instead of normalizing each of the dimensions of the observed state  $S_c$  by their standard deviation, because some of the dimensions of the spatial fields  $V_{c,j}$  may not vary with a given forcing, resulting in zero-variance components. For example, in the atmospheric boundary layer, observations of liquid water specific humidity will always be zero below the lifting condensation level.

Stacking the observations from all configurations together, the full observation vector  $\tilde{y}$  is

$$\tilde{y} = \begin{bmatrix} \tilde{y}_1 \\ \dots \\ \tilde{y}_{|C|} \end{bmatrix} \in \mathbb{R}^{\tilde{d}}, \quad \tilde{d} = \sum_{c=1}^{|C|} \tilde{d}_c = \sum_{c=1}^{|C|} n_c h_c. \quad (4.12)$$

Following again the heuristic in Section 4.3, the noise covariance associated with each observation vector  $\tilde{y}_c \in \mathbb{R}^{\tilde{d}_c}$  is  $\tilde{\Gamma}_c = \text{Cov}(s_c)$ , computed as in equation (4.11). Given that the noise is estimated independently for each configuration, the full noise covariance is the block diagonal matrix

$$\tilde{\Gamma} = \begin{bmatrix} \tilde{\Gamma}_1 & & 0 \\ & \ddots & \\ 0 & & \tilde{\Gamma}_{|C|} \end{bmatrix} \in \mathbb{R}^{\tilde{d} \times \tilde{d}}, \quad \tilde{\Gamma}_c = \text{Cov}(s_c) \in \mathbb{R}^{\tilde{d}_c \times \tilde{d}_c}, \quad (4.13)$$

where  $\tilde{\Gamma}_c$  is the noise covariance matrix of configuration  $c$ .

### Observations in a reduced space

Each covariance matrix  $\tilde{\Gamma}_c$ , possibly associated with high-dimensional observations and a finite sampling interval, is likely to be rank-deficient and have a large condition number  $\kappa = \mu_{c,1}/\mu_{c,r_c}$ , where  $\mu_{c,i}$  is the  $i$ -th largest eigenvalue of  $\tilde{\Gamma}_c$  and  $r_c$  is the approximate rank of the matrix (Hansen, 1998). Numerically rank-deficient problems arise when  $\tilde{d}_c$  is greater than or equal to the number of samples used to construct  $\tilde{\Gamma}_c$ , or when there exist eigenvalues  $\mu_{c,i}$  such that  $\mu_{c,i}/\mu_{c,1} \lesssim \epsilon_m$ , where  $\epsilon_m$  is a measure of data or machine precision. An efficient regularization method for rank-deficient problems is to project the data from each configuration onto a lower-dimensional encoding, adding Tikhonov regularization to limit the condition number of the resulting global covariance matrix. If the lower-dimensional encoding is obtained through principal component analysis (PCA),

$$y_c = P_c^T \tilde{y}_c, \quad \Gamma_c = d_c P_c^T \tilde{\Gamma}_c P_c + \kappa_*^{-1} \mu_1 I_{d_c}, \quad (4.14)$$

where  $y_c \in \mathbb{R}^{d_c}$ ,  $P_c$  is the projection matrix formed by the  $d_c$  leading eigenvectors of  $\tilde{\Gamma}_c$ ,  $I_{d_c}$  is the identity matrix,  $\mu_1$  is the leading eigenvalue of the unregularized global

covariance and  $\kappa_*$  is the limiting condition number of the global covariance, which should be chosen to be  $\kappa_* < \epsilon_m^{-1/2}$ . The encoding dimension  $d_c$  should be chosen such that  $d_c \leq r_c \leq \tilde{d}_c$ , where  $r_c$  is the approximate rank of  $\tilde{\Gamma}_c$ . The actual value of  $d_c$  may be chosen through the discrepancy principle, generalized cross validation, or based on the preservation of a given fraction of the total variance, among other criteria (Reichel and Rodriguez, 2013; Hansen, 1998). The Tikhonov inflation term regularizes problems where PCA is performed between eigenvalues that are close in value, or where the range of configuration variances  $\text{tr}(\tilde{\Gamma}_c)$  is large (Hansen, 1990). In projection (4.14), since the number of retained principal components may differ among configurations for a given truncation criterion, each block covariance matrix is scaled by  $d_c$ .

Projection (4.14) enables the use of arbitrarily correlated observations by regularizing the linear system  $\Gamma^{-1}(\mathcal{G}(\theta) - y)$  that appears in the gradient of the loss

$$\nabla \mathcal{L}(\theta; y) \propto (D\mathcal{G}(\theta))^T \Gamma^{-1}(\mathcal{G}(\theta) - y), \quad (4.15)$$

and lowering its computational cost. Here,  $D\mathcal{G}(\theta) \in \mathbb{R}^{d \times p}$  is the Jacobian matrix of  $\mathcal{G}$  evaluated at  $\theta$ . Although the ensemble Kalman algorithms presented in Section 4.4 do not compute the gradient (4.15) explicitly, they do rely on approximations of it, so this regularization effect still applies.

Since  $\tilde{\Gamma}$  in equation (4.13) is block diagonal, PCA can be performed in parallel for different configurations. The projection (4.14) maximizes the projected variance for each configuration; it is different than performing PCA on  $\tilde{\Gamma}$  in that it does not discriminate based on the total variance of each configuration. Disparities between the two approaches are discussed in 4.7. Finally, the regularized observation vector and noise covariance matrix read

$$y = \begin{bmatrix} y_1 \\ \dots \\ y_{|C|} \end{bmatrix} \in \mathbb{R}^d, \quad \Gamma = \begin{bmatrix} \Gamma_1 & & 0 \\ & \ddots & \\ 0 & & \Gamma_{|C|} \end{bmatrix} \in \mathbb{R}^{d \times d}, \quad (4.16)$$

which define a regularized inverse problem of the form (4.3)–(4.5). A schematic of the inverse problem construction process is given in Figure 4.1. The construction of  $y_c$  from each dynamical system configuration  $\zeta(x_c)$  defines the observational map  $\mathcal{H}_c$ , used to obtain the forward model evaluation  $\mathcal{G}_c : \mathbb{R}^p \rightarrow \mathbb{R}^{d_c}$  for the same configuration from the dynamical model. The construction of each  $(y_c, \Gamma_c)$  pair, and the evaluation of  $\mathcal{G}_c(\cdot)$ , can be done in parallel.

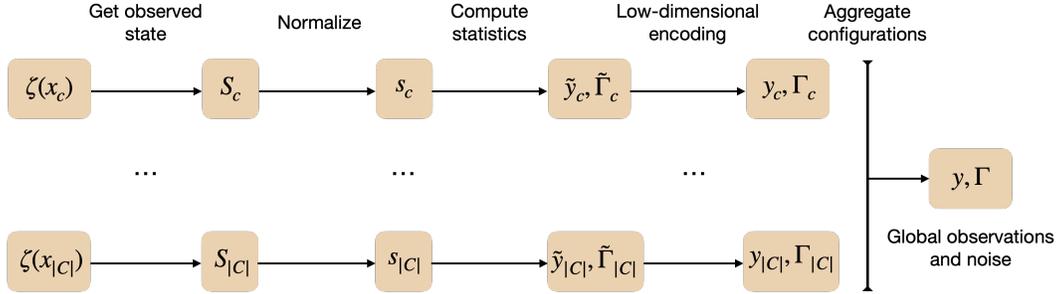


Figure 4.1: Schematic of the strategy used to construct a regularized inverse problem from observations of a dynamical system  $\zeta$ . The two branches represent different configurations of the dynamical system. From left to right: (a) the observed state is obtained following Section 4.3 or from any observable fields for each configuration  $c$ ; (b) the observed state is normalized; (c) mean and covariance of the normalized state are computed; (d)  $\tilde{y}_c$  and  $\tilde{\Gamma}_c$  are projected onto a lower dimension and regularized; (e) the statistical summaries of each configuration are aggregated, defining the global inverse problem (4.3)–(4.5).

### Bayesian interpretation of the loss and batching

Once the data and noise estimate encodings (4.16) have been defined, iterative methods to solve inverse problem (4.3)–(4.5) require evaluating the loss  $\mathcal{L}(\theta; y)$  at each iteration, which entails running the dynamical model in all configurations  $C$  and can be very computationally demanding. A less onerous alternative is to use a mini-batch of configurations  $B \subset C$  to evaluate the average configuration loss,

$$L(\theta; y_B) = \frac{1}{2|B|} \sum_{c=1}^{|B|} \|y_c - \mathcal{G}_c(\theta)\|_{\Gamma_c}^2 = \frac{1}{2} \sum_{c=1}^{|B|} \|y_c - \mathcal{G}_c(\theta)\|_{|B|\Gamma_c}^2, \quad (4.17)$$

which acts as a surrogate of the configuration-averaged loss  $L(\theta; y) = \mathcal{L}(\theta; y)/|C|$ . The use of  $L(\theta; y_B)$  in lieu of  $L(\theta; y)$  may be regarded as using noisy evaluations of the loss for each parameter update. From a Bayesian perspective, using  $L(\theta; y)$  in expression (4.5) leads to the same MAP estimator as  $\mathcal{L}(\theta; y)$  but a wider uncertainty about it, since we no longer consider configurations independent. This is important when interpreting the posterior uncertainty. To employ the loss (4.17), we only need to use the scaling  $\Gamma_c \rightarrow |B|\Gamma_c$ ; to approximate the aggregate loss  $\mathcal{L}(\theta, y)$  when batching, we can use  $\Gamma_c \rightarrow (|B|/|C|)\Gamma_c$  instead.

Batching is widely employed in data assimilation (Houtekamer and Mitchell, 2001) and deep learning, where it has been shown to help avoid convergence to local minima that generalize poorly (M. Li et al., 2014; Keskar et al., 2016). Understanding the behavior of algorithms when using mini-batches is crucial for online learning, where observations become available sequentially and the full loss cannot be sam-

pled. Moreover, it provides insight into the appropriateness of training sequentially on seasonal or geographically sparse data in Earth system modeling applications. We explore the effect of batching on the solution of the inverse problem in Section 4.5, training sequentially on randomly sampled configurations with markedly different dynamics.

#### 4.4 Ensemble Kalman methods

We consider two highly parallelizable gradient-free algorithms to solve the inverse problem defined in Section 4.3: ensemble Kalman inversion (EKI, Iglesias et al., 2013) and unscented Kalman inversion (UKI, D. Z. Huang et al., 2022a). Both algorithms are based on the extended Kalman filter and draw heavily on Gaussian conditioning for their derivation: underlying their update rules is the approximation of the parameter distribution as Gaussian. They afford a Bayesian interpretation when augmented with prior information at every iteration (D. Z. Huang et al., 2022b); how to do this is discussed in Section 4.4. If prior information is not used, which may be desirable when training for instance neural networks, they can be regarded as derivative-free methods to obtain the MLE.

EKI and UKI have been used successfully in a wide variety of inverse problems (Iglesias et al., 2013; Iglesias, 2016; Xiao et al., 2016; Kovachki and Stuart, 2019; D. Z. Huang et al., 2022a). We demonstrate them here in the context of training models that may experience numerical instabilities for a priori unknown parameter combinations, starting with a brief review of the algorithms.

##### Ensemble Kalman inversion (EKI)

Ensemble Kalman inversion searches for the optimal  $\theta^*$  given an inverse problem (4.3)–(4.5) through iterative updates of an initial parameter ensemble  $\Theta_0 = [\theta_0^{(1)}, \dots, \theta_0^{(J)}]$ , used to obtain empirical estimates of covariances between parameters and the model output at each step of the algorithm. We form the initial ensemble by randomly sampling  $J$  parameter vectors  $\theta_0^{(j)} \in \mathbb{R}^p$  from a Gaussian  $\mathcal{N}(m_0, \Sigma_0)$ . The EKI update equation for the ensemble at iteration  $n$  is (Schillings and Stuart, 2017)

$$\Theta_{n+1} = \Theta_n + \text{Cov}(\theta_n, \mathcal{G}_n) [\text{Cov}(\mathcal{G}_n, \mathcal{G}_n) + \Delta t^{-1} \Gamma]^{-1} \varepsilon(\Theta_n), \quad (4.18)$$

where  $\Theta_n \in \mathbb{R}^{p \times J}$ ,  $\Delta t$  is the nominal learning rate of the algorithm, and  $\varepsilon(\Theta_n) \in \mathbb{R}^{d \times J}$  encodes the mismatch between the forward model evaluations and the data,

$$\varepsilon(\Theta_n) = [y_{n+1}^{(1)} - \mathcal{G}(\theta_n^{(1)}), \dots, y_{n+1}^{(J)} - \mathcal{G}(\theta_n^{(J)})], \quad (4.19)$$

where

$$y_{n+1}^{(j)} = y + \xi_{n+1}^{(j)}, \quad \xi_{n+1}^{(j)} \sim \mathcal{N}(0, \Delta t^{-1} \Gamma). \quad (4.20)$$

All covariances in update (4.18) are estimated as sample covariances of the  $J$  ensemble members,

$$\text{Cov}(\theta_n, \mathcal{G}_n) = \frac{1}{J} \left( \Theta_n - \frac{1}{J} \sum_j \theta_n^{(j)} \mathbf{1}^T \right) \left( \mathcal{G}_{\Theta_n} - \frac{1}{J} \sum_j \mathcal{G}(\theta_n^{(j)}) \mathbf{1}^T \right)^T, \quad (4.21)$$

$$\text{Cov}(\mathcal{G}_n, \mathcal{G}_n) = \frac{1}{J} \left( \mathcal{G}_{\Theta_n} - \frac{1}{J} \sum_j \mathcal{G}(\theta_n^{(j)}) \mathbf{1}^T \right) \left( \mathcal{G}_{\Theta_n} - \frac{1}{J} \sum_j \mathcal{G}(\theta_n^{(j)}) \mathbf{1}^T \right)^T, \quad (4.22)$$

where  $\mathcal{G}_{\Theta_n} = [\mathcal{G}(\theta_n^{(1)}), \dots, \mathcal{G}(\theta_n^{(J)})]$ , and  $\mathbf{1} \in \mathbb{R}^J$  is the all-ones vector. Note that the sample covariances (4.21) and (4.22) have at most ranks  $\min(\min(d, p), J - 1)$  and  $\min(d, J - 1)$ , respectively. From definitions (4.14) and (4.16),  $\text{rank}(\Gamma) = d$  by construction, so the linear system in (4.18) is well-defined even for  $J < d$ .

Through iterative application of the update (4.18), the ensemble  $\Theta$  minimizes the projection of the model-data mismatch on the linear span of its  $J$  members. In this study, we limit the use of EKI and UKI to the calibration of dynamical models for which using an ensemble size  $J \sim p$  is feasible. For models with a large number of parameters, localization or sampling error correction techniques can be used to maintain performance with  $J \ll p$  members (Y. Lee, 2021; Tong and Morzfeld, 2022), like in EnKF for data assimilation (J. L. Anderson, 2012). The update (4.18) also drives the ensemble toward consensus, in the sense that  $|\text{Cov}(\theta_n, \mathcal{G}_n)| \rightarrow 0$  as  $n \rightarrow \infty$ ; a popular method to control collapse speed is additive inflation (J. L. Anderson and S. L. Anderson, 1999; Tong and Morzfeld, 2022). This collapse property precludes obtaining information about parameter uncertainties directly from EKI. However, the sequence of parameter-output pairs  $\{\Theta_n, \mathcal{G}_{\Theta_n}\}$  can be used to train emulators for uncertainty quantification (Cleary et al., 2021).

### Addressing model failures within the ensemble

For some parameters  $\theta_f$ , simulations may be physically or numerically unstable. For instance, the Courant–Friedrichs–Lewy condition in fluid solvers may change nonlinearly with model parameters, or the initialized weights from a neural network parameterization may lead to unstable trajectories. In such situations, we need to modify update (4.18) to account for model failures within the ensemble.

Here we propose a novel failsafe EKI update based on the successful parameter ensemble. Let  $\Theta_{s,n} = [\theta_{s,n}^{(1)}, \dots, \theta_{s,n}^{(J_s)}]$  be the successful ensemble, for which each

evaluation  $\mathcal{G}(\theta_{s,n}^{(j)})$  is stable or physically consistent, and let  $\theta_{f,n}^{(k)}$  be the ensemble members for which the evaluation of the forward model  $\mathcal{G}(\theta_{f,n}^{(k)})$  fails. We update the successful ensemble  $\Theta_{s,n}$  to  $\Theta_{s,n+1}$  using expression (4.18), and redraw each failed ensemble member from a Gaussian defined by the successful ensemble

$$\theta_{f,n+1}^{(k)} \sim \mathcal{N}(m_{s,n+1}, \Sigma_{s,n+1}), \quad (4.23)$$

where

$$m_{s,n+1} = \frac{1}{J_s} \sum_{j=1}^{J_s} \theta_{s,n+1}^{(j)}, \quad \Sigma_{s,n+1} = \text{Cov}(\theta_{s,n+1}, \theta_{s,n+1}) + \kappa_*^{-1} \mu_{s,1} I_p \quad (4.24)$$

are the sample mean and regularized sample covariance matrix of the updated successful ensemble. In expression (4.24),  $\kappa_*$  is a limiting condition number and  $\mu_{s,1}$  is the largest eigenvalue of the sample covariance  $\text{Cov}(\theta_{s,n+1}, \theta_{s,n+1})$ . This update has proved very effective for us in practice, even in situations where  $J_s < J/2$ ; we use it throughout Section 4.5. The failsafe update may be combined with other conditioning techniques at initialization. For instance, the initial ensemble  $\Theta_0$  may be drawn recursively until the number of failed members is reduced below an acceptable threshold.

### Bayesian regularization in ensemble Kalman methods

EKI implicitly regularizes the inverse problem by searching for the optimal solution  $\theta^*$  over the finite-dimensional space spanned by the initial ensemble. Although UKI does not share this property, both algorithms can be equipped with Bayesian regularization by considering the augmented data-model relation (Chada et al., 2020)

$$y_a = \mathcal{G}_a(\theta) + \xi := \begin{bmatrix} y \\ m_p \end{bmatrix} = \begin{bmatrix} \mathcal{G}(\theta) \\ \theta \end{bmatrix} + \begin{bmatrix} \hat{\delta} + \hat{\eta} \\ \lambda \end{bmatrix}, \quad (4.25)$$

instead of expression (4.3). Here,  $m_p \in \mathbb{R}^p$  is the parameter prior mean,  $\lambda \sim \mathcal{N}(0, 2\Lambda)$  defines the degree of regularization,  $\hat{\delta} + \hat{\eta} \sim \mathcal{N}(0, 2\Gamma)$ , and  $\xi \sim \mathcal{N}(0, \Gamma_a)$  is the augmented error defined by relation (4.25). In practice, using expression (4.25) amounts to substituting  $\{\mathcal{G}, y, \Gamma\}$  by  $\{\mathcal{G}_a, y_a, \Gamma_a\}$  in both algorithms. The Kalman inversion solution to the inverse problem defined by relation (4.25) then satisfies

$$\theta^* = \arg \min_{\theta} \left[ \mathcal{L}(\theta; y) + \frac{1}{2} \|\theta - m_p\|_{\Lambda}^2 \right]. \quad (4.26)$$

From a Bayesian perspective, the solution (4.26) approximately maximizes the posterior density (4.5) for the Gaussian prior  $\rho_{\text{prior}} \sim \mathcal{N}(0, \Lambda)$ . This is particularly

interesting for UKI, which provides parametric uncertainty estimates (D. Z. Huang et al., 2022b). When using a nominal learning rate  $\Delta t \neq 1$ , the scaling  $\Lambda \rightarrow \Delta t \cdot \Lambda$  must be used to retain the Bayesian interpretation of  $\Lambda$  as the prior variance, due to the fact that  $\Delta t$  effectively modifies the noise in update (4.18) to be  $\Delta^{-1}\Gamma$ . As noted before, if the original data-model relation (4.3) is used instead of the augmented relation (4.25), UKI and EKI provide maximum likelihood estimators.

### Unscented Kalman inversion (UKI)

Unscented Kalman inversion seeks a Gaussian approximation of the posterior  $\rho(\theta|y, \Gamma)$  around the MAP (given relation (4.25)), or an approximation of the likelihood around the MLE (given (4.3)), by deterministically evolving an initial Gaussian estimate  $\mathcal{N}(m_0, \Sigma_0)$  through updates

$$m_{n+1} = m_n + \text{Cov}_q(\theta_n, \mathcal{G}_n) [\text{Cov}_q(\mathcal{G}_n, \mathcal{G}_n) + \Delta t^{-1}\Gamma]^{-1} \varepsilon(m_n), \quad (4.27)$$

$$\Sigma_{n+1} = (1 + \Delta t)\Sigma_n - \text{Cov}_q(\theta_n, \mathcal{G}_n) [\text{Cov}_q(\mathcal{G}_n, \mathcal{G}_n) + \Delta t^{-1}\Gamma]^{-1} \text{Cov}_q(\theta_n, \mathcal{G}_n)^T, \quad (4.28)$$

where  $m_n$  and  $\Sigma_n$  are the mean and covariance estimates of the Gaussian after  $n$  iterations of the algorithm, and  $\varepsilon(m_n) = y - \mathcal{G}(m_n)$  is the data-model mismatch of the mean estimate. The covariances  $\text{Cov}_q(\theta_n, \mathcal{G}_n)$  and  $\text{Cov}_q(\mathcal{G}_n, \mathcal{G}_n)$  in expressions (4.27) and (4.28) are computed through quadratures over  $2p+1$  sigma points defined as

$$\begin{aligned} \hat{\theta}_n^{(j)} &= m_n + a\sqrt{p}[\sqrt{\Sigma_n(1 + \Delta t)}]_j, & 1 \leq j \leq p, \\ \hat{\theta}_n^{(j+p)} &= m_n - a\sqrt{p}[\sqrt{\Sigma_n(1 + \Delta t)}]_j, & 1 \leq j \leq p, \end{aligned} \quad (4.29)$$

where  $[\sqrt{\Gamma}]_j$  is the  $j$ -th column of the Cholesky factor of  $\Gamma$ ,  $a = \min(\sqrt{4/p}, 1)$  is a hyperparameter defined in D. Z. Huang et al. (2022a), and  $\hat{\theta}_n^{(0)} = m_n$  is the central sigma point. The quadratures are then defined as

$$\text{Cov}_q(\theta_n, \mathcal{G}_n) = \sum_{j=1}^{2p} w_j (\hat{\theta}_n^{(j)} - m_n) (\mathcal{G}(\hat{\theta}_n^{(j)}) - \mathcal{G}(m_n))^T, \quad (4.30)$$

$$\text{Cov}_q(\mathcal{G}_n, \mathcal{G}_n) = \sum_{j=1}^{2p} w_j (\mathcal{G}(\hat{\theta}_n^{(j)}) - \mathcal{G}(m_n)) (\mathcal{G}(\hat{\theta}_n^{(j)}) - \mathcal{G}(m_n))^T, \quad (4.31)$$

where  $w_j = (2a^2p)^{-1}$  are the quadrature weights.

A limitation of this algorithm is that the number of sigma points scales linearly with  $p$ , which precludes its use when training models with a large number of parameters.

However, for situations where using an ensemble of  $2p+1$  members is tractable, UKI improves upon EKI by providing uncertainty quantification, instead of collapsing to a point estimate. In particular, when updates (4.27) and (4.28) are applied to the augmented data-model relation (4.25), UKI ensures that  $\Sigma_n$  in the limit  $n \rightarrow \infty$  converges towards a Gaussian estimate of parametric uncertainty (D. Z. Huang et al., 2022b),

$$\Sigma_\infty \approx \text{Cov}_q(\theta_\infty, \mathcal{G}_{a,\infty}) [\Delta t \cdot \text{Cov}_q(\mathcal{G}_{a,\infty}, \mathcal{G}_{a,\infty}) + \Gamma_a]^{-1} \text{Cov}_q(\theta_\infty, \mathcal{G}_{a,\infty})^T, \quad (4.32)$$

which involves the augmented forward model  $\mathcal{G}_a(\cdot)$  and covariance  $\Gamma_a$  defined in Section 4.4.  $\Sigma_\infty$  approximates the covariance of the posterior (4.5) around  $m_\infty$  if the full loss is evaluated at every UKI iteration and  $\Delta t = 1$  (D. Z. Huang et al., 2022b). When batching, an equivalent approximation can be recovered by using  $\Delta t = |C|/|B|$  to compensate for sampling errors in the construction of the empirical covariances (4.30) and (4.31); this is demonstrated in Section 4.5.

Finally, note that the limit (4.32) does not depend on  $\Sigma_0$ , only on the Bayesian prior covariance  $\Lambda$ . This enables using a tight initial guess (i.e.,  $\text{tr}(\Sigma_0) \ll \text{tr}(\Lambda)$ ), which can reduce the fraction of model failures within the ensemble. To ensure robustness to the model failures that may still arise, we propose a modification of the UKI dynamics robust to model failures, similar to the one proposed for EKI, in 4.8.

#### 4.5 Application to an atmospheric subgrid-scale model

In this section, the framework and algorithms discussed in Sections 4.3 and 4.4 are used to learn closure parameters within an EDMF scheme of atmospheric turbulence and convection. The EDMF scheme is derived by spatially filtering the Navier-Stokes equations for an anelastic fluid, and then decomposing the subgrid flow into  $n > 1$  distinct subdomains with moving boundaries (Cohen et al., 2020). In practice, the subdomain decomposition requires the use of  $n - 1$  additional equations per grid-mean prognostic field, and  $n - 1$  additional equations tracking the volume fraction of each subdomain within the grid (Tan et al., 2018). We retain second-order moments for one of the subdomains, the environment. Covariances within the other subdomains (updrafts) are neglected, which circumvents the need for turbulence closures therein. In the end, the EDMF equations require closure for the turbulent diffusivity and dissipation in the environment, and the mass, momentum, and tracer fluxes between environment and updrafts. In what follows, we consider an EDMF scheme with a single updraft ( $n = 2$ ).

We consider the EDMF scheme discussed in Cohen et al. (2020) and Lopez-Gomez et al. (2020), which is implemented in a single-column model (SCM). Within this SCM, we first seek to learn 16 closure parameters: 5 describing turbulent mixing, dissipation, and mixing inhibition by stratification (Lopez-Gomez et al., 2020), 3 describing the momentum exchange between subdomains (J. He et al., 2021), 7 describing entrainment fluxes between updrafts and the environment (Cohen et al., 2020), and another one defining the surface area fraction occupied by updrafts. In Section 4.5, we substitute the empirical dynamical entrainment closure proposed in Cohen et al. (2020) by a neural network, and train the resulting physics-based machine-learning model.

To showcase the versatility of the algorithms, UKI is used for approximate Bayesian inference of empirical parameters (using relation (4.25)), and EKI is used for both MAP estimation of empirical parameters (relation (4.25), Sections 4.5, 4.5) and MLE estimation of neural network parameters (relation (4.3), Section 4.5). In all cases, we employ our failsafe modifications of the algorithms (Section 4.4 and 4.8). The name, prior range  $U$ , and reference to the definition of each empirical parameter in the literature are given in Table 4.1. The prior mean is taken to be equal to the parameter values used in Lopez-Gomez et al. (2020) and Cohen et al. (2020). The prior in unconstrained space  $\mathcal{N}(m_p, \Lambda)$  is obtained from the physical prior mean and range through transformations defined in 4.9. Finally, we initialize EKI ensembles from the prior,  $\mathcal{N}(m_0, \Sigma_0) \equiv \mathcal{N}(m_p, \Lambda)$ , and all UKI sigma points from a tighter initial guess  $\mathcal{N}(m_p, \Lambda/16)$  to demonstrate the ability of UKI to decouple from the initial guess.

### **Description of LES data and model configurations**

The data used for training and testing the EDMF scheme are taken from the LES library described in Shen et al. (2022). This library contains high-resolution simulations of low-level clouds spanning the stratocumulus-to-cumulus transition in the East Pacific Ocean. The large-scale forcing used for these simulations is derived from the cfSites output of the HadGEM2-A model, retrieved from the Coupled Model Intercomparison Project Phase 5 (CMIP5) archive. In particular, the monthly climatology of the cfSites output is computed over the 5-year period 2004-2008, and used to initialize and force large-eddy simulations for a period of 6 days. Radiative forcing is computed interactively using the Rapid Radiative Transfer Model (RRTM, Mlawer et al., 1997).

Table 4.1: Parameters  $\phi$  considered for calibration in this study. The prior mean values are taken from LG2020 (Lopez-Gomez et al., 2020), C2020 (Cohen et al., 2020) and H2021 (J. He et al., 2021), where a physical description of the parameters may be found.

Symbol	Description	Prior range	Prior mean
$c_m$	Eddy viscosity coefficient	(0.01, 1.0)	0.14, LG2020
$c_d$	Turbulent dissipation coefficient	(0.01, 1.0)	0.22, LG2020
$c_b$	Static stability coefficient	(0.01, 1.0)	0.63, LG2020
$\text{Pr}_{t,0}$	Neutral turbulent Prandtl number	(0.5, 1.5)	0.74, LG2020
$\kappa_*$	Ratio of rms turbulent velocity to friction velocity	(1.0, 4.0)	1.94, LG2020
$c_\varepsilon$	Entrainment rate coefficient	(0, 1)	0.13, C2020
$c_\delta$	Detrainment rate coefficient	(0, 1)	0.51, C2020
$c_\gamma$	Turbulent entrainment rate coefficient	(0, 10)	0.075, C2020
$\beta$	Detrainment relative humidity power law	(0, 4)	2.0, C2020
$\mu_0$	Entrainment sigmoidal activation parameter	( $10^{-6}$ , $10^{-2}$ )	$4 \cdot 10^{-4}$ , C2020
$\chi_i$	Updraft-environment buoyancy mixing ratio	(0, 1)	0.25, C2020
$c_\lambda$	Turbulence-induced entrainment coefficient	(0, 10)	0.3, C2020
$a_s$	Updraft surface area fraction	(0.01, 0.5)	0.1, C2020
$\alpha_b$	Updraft virtual mass loading coefficient	(0, 10)	0.12, H2021
$\alpha_a$	Updraft advection damping coefficient	(0, 100)	0.001, H2021
$\alpha_d$	Updraft drag coefficient	(0, 50)	10.0, H2021

The SCM runs are initialized from the coarse-grained LES fields after 5.75 days of simulation and are run for 6 hours. This runtime was chosen to be much longer than the equilibration time of the SCM to the steady forcing; experiments using a runtime of 12 hours resulted in no statistical changes of the results. Large-scale forcing is identical to that of the LES, and the radiative heating rates are given by the horizontal mean of the rates experienced by the high-resolution simulations. The observational map used to define the inverse problem follows the guidelines of Section 4.3, using time and horizontally averaged vertical profiles from the last  $T_c = 3$  hours of simulation, at a vertical resolution of  $\Delta z = 50$  m; this is also the resolution of the SCM simulations, which employ 80 vertical levels. Following the strategy in Figure 4.1, we extract the observations from each configuration as

$$S_c = [\bar{u}, \bar{v}, \bar{s}, \bar{q}_l, \bar{q}_t, \overline{w'q'_t}, \overline{w's'}]^T, \quad (4.33)$$

where  $\overline{(\cdot)}$  denotes time and horizontal averaging,  $\bar{u}$  and  $\bar{v}$  are the horizontal velocity components,  $\bar{s}$  is the entropy,  $\bar{q}_t$  is the total specific humidity,  $\overline{w'q'_t}$  and  $\overline{w's'}$  are vertical fluxes of moisture and entropy, and  $\bar{q}_l$  is the liquid water specific humidity. The pooled variances for normalization and covariance matrix  $\tilde{\Gamma}_c$  associated with the observed state  $S_c$  are obtained from the full 6 day statistics of the LES to

capture the internal variability of the system. Finally, a low-dimensional encoding is obtained from the normalized time-averaged observations through truncated PCA as in equation (4.14), truncating the dimension of the noise covariance matrix so as to preserve 99% of the total noise variance. Calibration results using fewer observed fields at a coarser resolution are discussed in Section 4.5.

As training data we include a total of 60 LES configurations from the Atmospheric Model Intercomparison Project (AMIP) experiment, spanning the months of January, April, July and October, and locations from the coasts of Peru and California to the tropical Pacific. Results are also shown for a validation set, which includes January and July simulations from an AMIP4K experiment, where sea surface temperature is increased by 4 K with respect to AMIP. This temperature increase leads to 10–20% weaker large-scale subsidence, higher cloud tops, and reduced cloud cover; see Shen et al. (2022) for a detailed comparison. Validation results are representative of the generalizability of the trained model for the simulation of a warming climate; the model was not trained on these warmer conditions.

### **Calibration using mini-batch loss evaluations**

To demonstrate the effectiveness of Kalman inversion in settings where evaluating all configurations of interest per iteration may be too expensive or impossible (e.g., due to sequential data availability), we present calibration results using mini-batches. Batching introduces noise in the loss evaluations due to sampling error. For this reason, the behavior of Kalman inversion algorithms using mini-batches is representative of their robustness to other sources of noise, such as noise in the data or stochasticity of the dynamical model. To correct for sampling noise due to batching, we use  $\Delta t = |C|/|B|$  as discussed in Section 4.4.

For training, data are fed to the algorithm by drawing  $|B|$  configurations randomly and without replacement from the training set at every iteration. Configurations are reshuffled at the end of every epoch (i.e., every full pass through the training set). Figure 4.2 shows the evolution of the training and validation errors for UKI and EKI, using training batches of 5 and 20 configurations. Since the total number of configurations in the training set is 60, an epoch requires 12 iterations when using  $|B| = 5$  and 3 when using  $|B| = 20$ . For many geophysical applications, the cost of evaluating an ensemble of long-term statistics  $\mathcal{G}(\cdot)$  from a forward model is significantly higher than performing the inversion updates (4.18) or (4.27). In these situations, a training epoch has similar computational cost for any value of  $|B|$ .

The training error is evaluated in normalized physical space with respect to the current batch,

$$\text{MSE}(\theta; \tilde{y}_B) = \frac{1}{\tilde{d}_B} \|\tilde{y}_B - \tilde{\mathcal{G}}_B(\theta)\|^2 = \frac{1}{\sum_{c=1}^{|B|} \tilde{d}_c} \sum_{c=1}^{|B|} \|\tilde{y}_c - \tilde{\mathcal{G}}_c(\theta)\|^2, \quad (4.34)$$

where  $\tilde{y}_B \in \mathbb{R}^{\tilde{d}_B}$ . The validation error is defined similarly, but it is computed over the entire validation set at every iteration. Thus, variations in the validation error are only due to changes in the model parameters; there is no random data sampling. The training and validation errors decrease sharply during the first epoch (Fig. 4.2). Subsequent epochs fine-tune the model parameters, further reducing the data-model mismatch. It is remarkable and important that the validation error decreases by about the same magnitude as the training error, demonstrating that the parameterization approach that leverages a physical model generalizes successfully out of the present-climate training sample to a warmer climate.

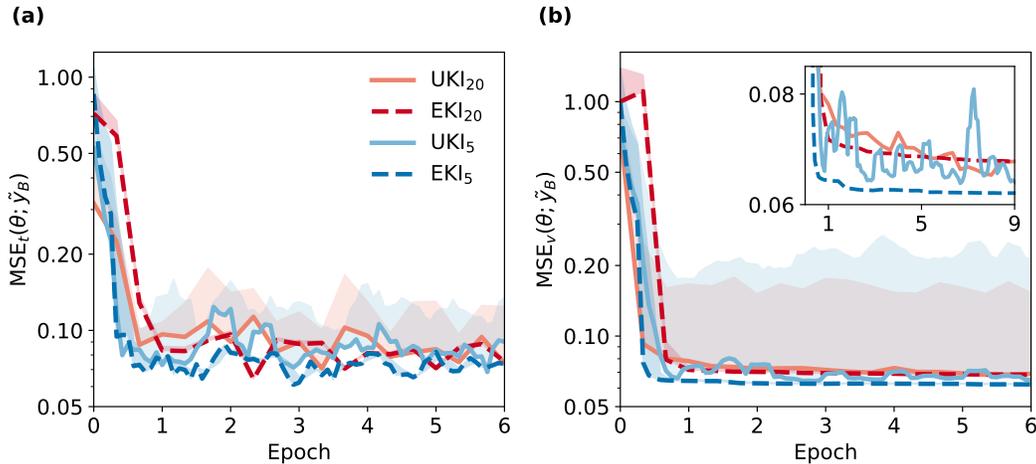


Figure 4.2: Batch (a) training and (b) validation MSE as defined in equation (4.34). Lines represent the error of the ensemble mean  $\bar{\theta}$ ,  $\text{MSE}(\bar{\theta}; \tilde{y}_B)$ , and the shading represents the ensemble standard deviation of  $\text{MSE}(\theta; \tilde{y}_B)$  around the optimal point estimate  $\bar{\theta}$ . All errors are normalized with respect to the largest initial  $\text{MSE}_v(\bar{\theta}; \tilde{y}_B)$ , so they can be compared. Results are shown for EKI and UKI, using  $J = 2p + 1$  and training batch sizes  $|B| = 5, 20$ . Errors for  $|B| = 5$  are averaged using a rolling mean of 20 configurations to enable comparison with  $|B| = 20$ . In (b), the inset focuses on the validation error evolution for a longer training period.

Both EKI and UKI display efficient training in the low batch-size regime: the validation error tends to be lower for smaller batches after a fixed number of epochs. Hence, decreasing batch size in EKI and UKI can help reduce the computational cost of training models. The optimal batch size will depend on the CPU and wall-clock

time constraints of the user. Although using smaller batches reduces CPU time, it requires more serial operations, so using larger batches can reduce wall-clock time.

The sampling noise due to the use of different configurations (e.g., stratocumulus versus cumulus regimes) increases for smaller batches. Although both algorithms achieve convergence for a wide range of batch sizes, we find that EKI is more robust than UKI to high levels of noise. This is shown in the inset of Figure 4.2b for  $|B| = 5$ , and in 4.10 for  $|B| = 1$ . Other differences between UKI and EKI are observed in Figure 4.2. The consensus property of EKI leads to a collapse of the model error spread after a few iterations, converging to a point estimate. On the other hand, the UKI ensemble converges to an MSE spread characteristic of the parameter uncertainty as approximated by the distribution  $\mathcal{N}(m_n, \Sigma_n)$ .

The evolution of the parameter estimate  $(m_n, \Sigma_n)$  is depicted in Figure 4.3 for the turbulent dissipation  $c_d$ , updraft advection damping  $\alpha_a$  and surface area fraction  $a_s$ . The initial parameter estimate depends on the stochastic initialization for EKI. The UKI estimate provides information about parameter uncertainty, whereas EKI only provides a point estimate (i.e.,  $m_n$ ). From the UKI estimate, we observe that the training set constrains the likely values of the turbulent dissipation ( $c_d$ ) and surface area fraction ( $a_s$ ) to a significantly smaller region than the prior. However, the magnitude of updraft advection damping ( $\alpha_a$ ) is not identifiable using this dataset: the corresponding diagonal element of  $\Sigma_n$  converges to the prior variance used in the regularized problem (4.25) (Figure 4.3b).

The covariance estimate  $\Sigma_n$  also provides information about correlations between model parameters and total reduction of uncertainty (Figure 4.4). For the current stratocumulus-to-cumulus transition dataset, our EDMF model shows moderate correlations between parameters regulating the turbulence kinetic energy budget in the boundary layer ( $c_b, c_m, c_d$ , see Lopez-Gomez et al., 2020). We also find entrainment to be negatively correlated with surface updraft area fraction, detrainment and drag. These correlations can be used to improve parameterizations at the process level by identifying or developing a set of uncorrelated parameters. Figure 4.4b shows how  $\Sigma_n$  converges to a quasi-steady state estimate of the posterior covariance after  $\sim 30$  iterations.

Vertical profiles of  $\bar{q}_l, \overline{w'q'_l}$  and  $\bar{u}$  from the validation set are compared to the reference LES profiles in Figure 4.5. The calibrated model yields smoother and more accurate profiles than the model before training. In particular, calibration significantly reduces biases in liquid water specific humidity and moisture transport

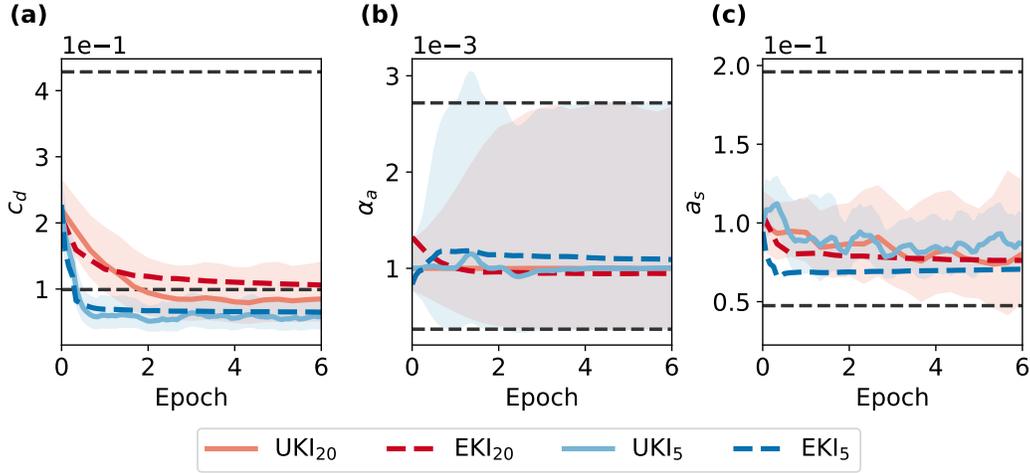


Figure 4.3: Parameter evolution of the turbulent dissipation (a), updraft advection damping (b), and updraft surface area fraction (c). All values are given in physical space. The solid lines describe the trajectories of the mean estimate,  $\mathcal{T}^{-1}(m_n)$ . For UKI, the marginal  $\pm\sigma$  uncertainty band is included in shading. This uncertainty is equal to  $\pm\mathcal{T}^{-1}(\sqrt{(\Sigma_n)_{i,i}})$  for the  $i$ -th parameter. The black dashed lines are the  $\pm\sigma$  uncertainty bands of the prior used for regularization. Legend as in Figure 4.2.

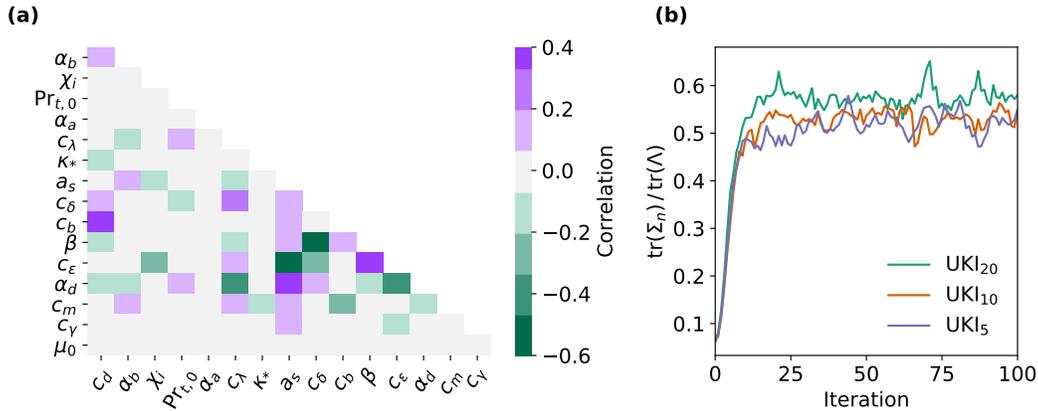


Figure 4.4: Parameter correlations estimated from UKI using  $|B| = 20$  (a), and evolution of the total parameter variance from UKI using  $|B| = 20, 10$  and  $5$ , normalized by the prior variance  $\text{tr}(\Lambda) = 16$  (b). Note that the initial covariance estimate used in UKI (with  $\text{tr}(\Sigma_0) = 1$ ) is decoupled from the prior. Symbols follow Table 4.1.

for both stratocumulus and cumulus cloud regimes in the 4 K-warmer AMIP4K experiment. These results confirm that the dynamical model can be trained using a low-dimensional encoding of the time statistics, as proposed in Section 4.3. They also highlight the generalizability of sparse physics-based models.

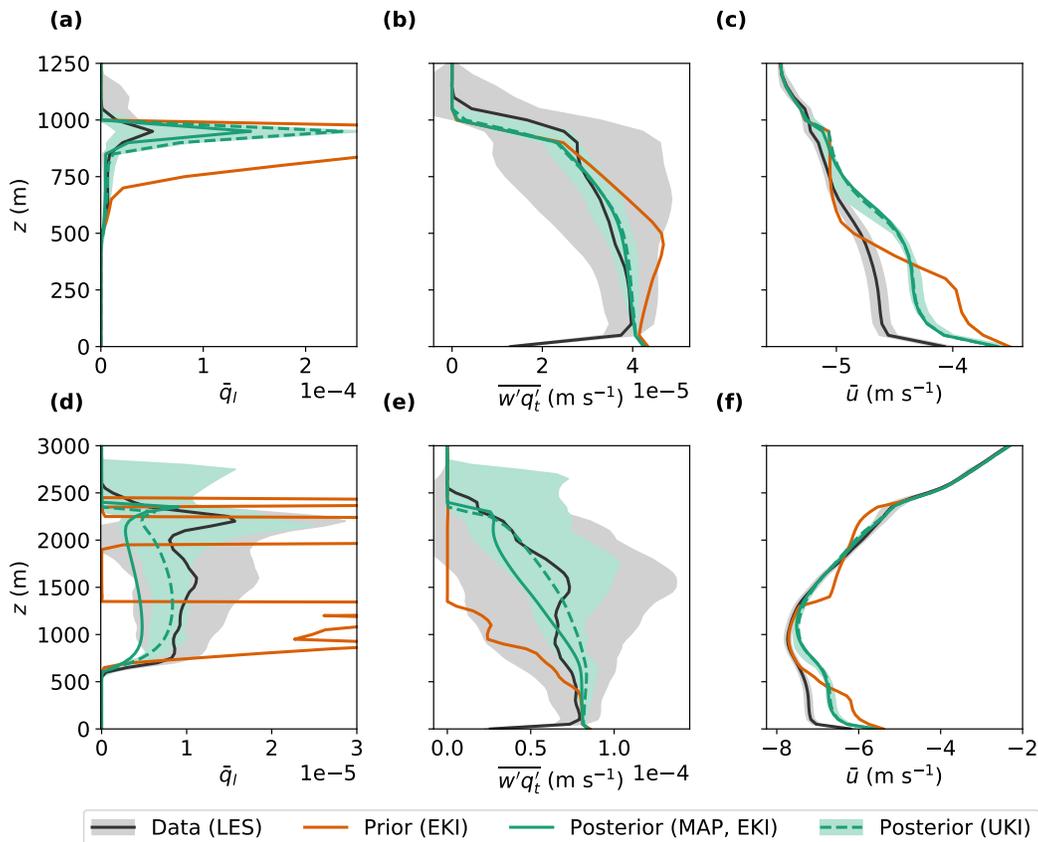


Figure 4.5: Prior, posterior and LES profiles of liquid water specific humidity ( $\bar{q}_l$ ), subgrid-scale moisture flux ( $\overline{w'q'_l}$ ) and zonal velocity ( $\bar{u}$ ) for cfSites 5 (top) and 14 (bottom) using July forcing from the AMIP4K experiment as in Shen et al. (2022). The gray shading represents the internal variability of the LES simulations over 6 days of steady forcing, and the full lines represent 3-hour time-averaged profiles. EKI prior and posterior results are point estimates evaluated at the parameter vector closest to the ensemble mean. The UKI posterior shading spans the central 68% of the profile posterior distribution. All Kalman methods used  $|B| = 5$  and  $J = 2p + 1$ .

### Calibration using partial observations

Another application of synthetic high-resolution data is the study of calibration sensitivity to data resolution and partial loss of information. Such sensitivity studies can inform the technical requirements of future observing systems or field campaigns (Suselj et al., 2020), and are easily implemented with ensemble and unscented Kalman inversion through modifications of the observational map  $\mathcal{H}$ .

Here, we employ the EKI and UKI algorithms for this task by using coarser training data at a vertical resolution of  $\Delta z = 200$  m. In addition, we consider only a subset of fields for which observational data may be obtained in practice: the liquid water potential temperature  $\bar{\theta}_l$ , the total water specific humidity  $\bar{q}_t$  and the liquid water

specific humidity  $\bar{q}_l$  (National Academies of Sciences, Engineering, and Medicine, 2018; Suselj et al., 2020). Figure 4.6 compares calibration results using this reduced setup with the results obtained using the full high-resolution observations of Section 4.5. The loss of information is evident in the inability of the algorithms to find the same minimum reached with richer observations. Nevertheless, Kalman inversion significantly reduces the validation error from the prior even with sparser data and a limited number of fields.

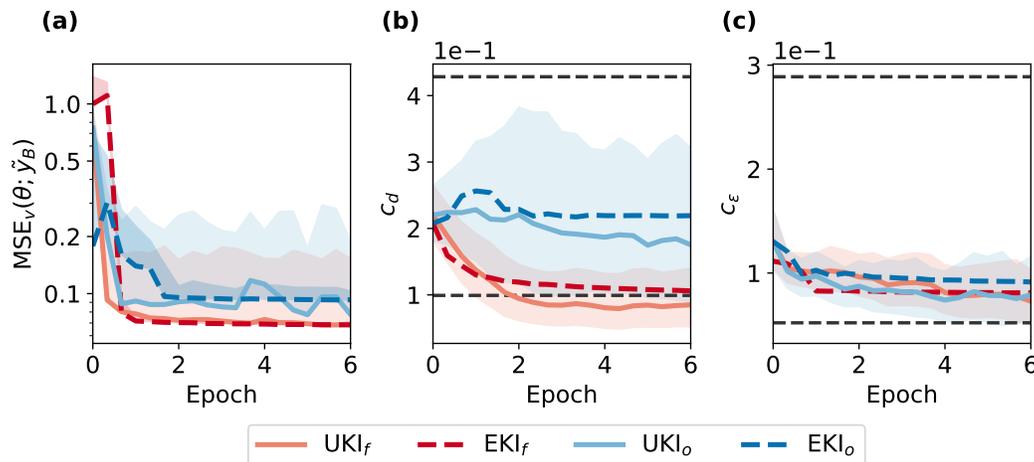


Figure 4.6: Evolution of the validation error (a) and estimates of the turbulent dissipation (b) and entrainment coefficient (c) for calibration processes using observations of the state (4.33) at 50 m resolution (UKI<sub>f</sub>, EKI<sub>f</sub>), or from  $\bar{\theta}_l$ ,  $\bar{q}_t$  and  $\bar{q}_l$  at 200 m resolution (UKI<sub>o</sub>, EKI<sub>o</sub>). All inversion processes use  $|B| = 20$ . Shading is defined as in Figures 4.2 and 4.3.

The identifiability of individual parameters as a function of the observational map  $\mathcal{H}$  can be inferred from the UKI  $\Sigma_n$  diagnostic. Figure 4.6 shows that the partial observations of temperature and humidity are enough to constrain the entrainment coefficient in the EDMF scheme. However, the loss of information with respect to the original observations leads to much poorer constraints on the turbulent dissipation coefficient. The same comparison can be performed for any parameter of interest to inform observational requirements to constrain models at the process level. This diagnostic is an important advantage of UKI over EKI; identifiability is not directly inferable from ensemble Kalman inversion due to the ensemble collapse. However, this information can be recovered through the emulation of the forward map (Cleary et al., 2021).

The use of partial observations also highlights the benefits of learning from time statistics instead of tendencies. Learning from statistics not only ensures that the

calibrated dynamical model is stable, which requires a leap of faith when training on instantaneous tendencies (Bretherton et al., 2022). It also couples the evolution of thermodynamic and dynamical fields, which can improve the forecast of fields unseen during training. An example is shown in Figure 4.7. The model calibrated using thermodynamic profiles improves upon the prior model in the forecast of horizontal velocities within the boundary and cloud layers. A common reason to use tendencies for calibration is that they enable the use of supervised learning techniques, which are easy to implement for neural network architectures (e.g., Bretherton et al., 2022). In the next subsection, we demonstrate the power of UKI and EKI to calibrate hybrid models with embedded neural network parameterizations.

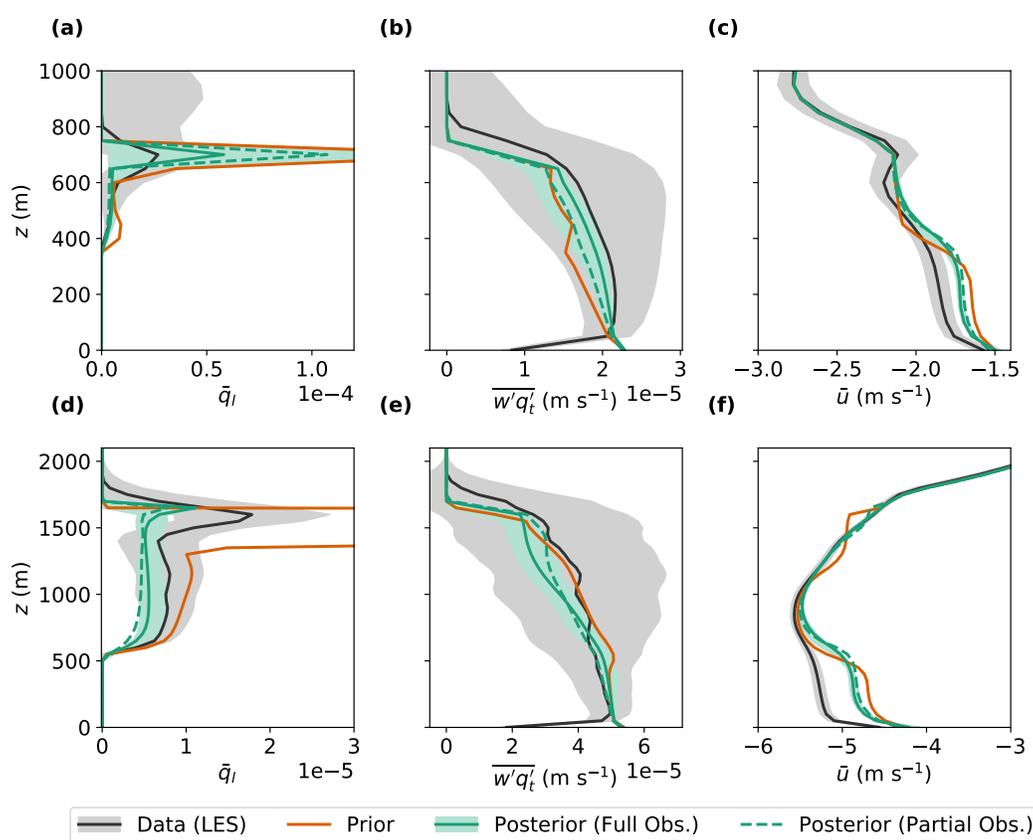


Figure 4.7: Prior, posterior and LES profiles of liquid water specific humidity ( $\bar{q}_l$ ), vertical moisture flux ( $\overline{w'q'_l}$ ) and zonal velocity ( $\bar{u}$ ) for cfSite 3 using July forcing (top) and cfSite 14 using January forcing (bottom) from the AMIP4K experiment (Shen et al., 2022). Posterior results are shown for a model calibrated using the high-resolution state (4.33) (Full Obs.), and coarse-resolution observations of  $\bar{\theta}_l$ ,  $\bar{q}_l$  and  $\bar{q}_l$  (Partial Obs.). Shadings and legend as in Figure 4.5. Results obtained using UKI with  $|B| = 20$ .

### Calibration of a hybrid model with embedded neural network closures

We consider now a hybrid EDMF scheme that substitutes the dynamical entrainment and detrainment closures proposed by Cohen et al. (2020) with a three-layer dense neural network. We define the fractional entrainment ( $\epsilon$ ) and detrainment ( $\delta$ ) rates as

$$\begin{bmatrix} \epsilon \\ \delta \end{bmatrix} = \frac{1}{z} \text{NN}_3(\Pi_1, \dots, \Pi_6), \quad (4.35)$$

where  $z$  is the height, and the hidden layers of  $\text{NN}_3$  have 5 and 4 nodes, from input to outputs. Our closure (4.35) seeks to learn local expressions for the  $z$ -normalized entrainment/detrainment rates, which have been shown to vary weakly in empirical studies of shallow cumulus convection (Siebesma, 1996; de Roode et al., 2000). The neural network inputs  $\Pi_1, \dots, \Pi_6$  are 6 nondimensional groups on which entrainment and detrainment may depend, defined as

$$\Pi_1 = \frac{z(b_{up} - b_{en})}{(w_{up} - w_{en})^2 + w_d^2}, \quad (4.36a)$$

$$\Pi_2 = \frac{a_{up}w_{up}^2 + (1 - a_{up})w_{en}^2}{2(1 - a_{up})e_{en} + a_{up}w_{up}^2 + (1 - a_{up})w_{en}^2}, \quad (4.36b)$$

$$\Pi_3 = \sqrt{a_{up}}, \quad (4.36c)$$

$$\Pi_4 = \text{RH}_{up} - \text{RH}_{en}, \quad (4.36d)$$

$$\Pi_5 = z/H_{up}, \quad (4.36e)$$

$$\Pi_6 = gz/R_d T_{\text{ref}}. \quad (4.36f)$$

In expressions (4.36),  $w_d = (H_{\text{inv}} \overline{w'b'}|_s)^{1/3}$  is the Deardorff convective velocity,  $H_{\text{inv}}$  is the inversion height,  $\overline{w'b'}|_s$  is the surface buoyancy flux,  $g$  is the gravitational acceleration,  $R_d$  is the ideal gas constant for dry air and  $T_{\text{ref}}$  is a reference temperature. The subscripts  $up$  and  $en$  denote updraft and environment:  $a_{up}$  is the updraft area fraction,  $H_{up}$  the updraft top height, and  $e_{en}$  the environmental turbulence kinetic energy. The relative humidity RH, vertical velocity with respect to the grid mean  $w$ , and buoyancy  $b$  are defined for both updraft and environment.

The neural network closure (4.35) introduces 63 additional coefficients with respect to the entrainment and detrainment closure calibrated in Sections 4.5 and 4.5, for a total of 79 parameters. As the closure complexity increases, it is most practical to use EKI for calibration, since it enables the use of ensembles with  $J < 2p + 1$  members. In Figure 4.8, we present training and validation errors for the hybrid

model using ensemble sizes  $J = 50, 100,$  and  $159,$  and for the empirical EDMF scheme with  $J = 2p + 1 = 33$  ensemble members. We initialize the neural network weights as  $\theta_{\text{NN}} \sim \mathcal{N}(\theta_{\text{NN}}^0, I)$  with  $\theta_{\text{NN}}^0 \sim U(-0.05, 0.05)$ . In all cases, we use Bayesian regularization as discussed in Section 4.5 for all model parameters except for the neural network weights. We calibrate all parameters of the empirical and hybrid models, to compare the optimal performance of both closures.

Both the empirical and hybrid EDMF schemes generalize well to the validation set, with training and validation errors reaching levels of about 5% of the largest a priori validation error. The strong generalization to 4 K-warmer cloud regimes contrasts with results from approaches that try to learn unresolved tendencies directly, without encoding the physics (Rasp et al., 2018). Using a physics-based hybrid approach, all learned closures are consistently placed within the coarse-grained dynamics of the system (Cohen et al., 2020), which also vastly reduces data requirements. Further, targeting closure terms that isolate a single physical process lends itself to interpretability in a manner difficult for purely machine-learning based parameterizations that simultaneously model many physical processes. After training, relationships between EDMF variables and targeted physical quantities like entrainment can be teased out using partial dependence plots or ablation studies. In addition, the learned relationships are point-wise and causal.

The inset in Figure 4.8b shows how the higher-complexity hybrid model moderately overfits to the training set after  $\sim 10$  epochs, a behavior that is not observed with the empirical model. Hence, in the low-data regime ( $d \lesssim p$ ), adoption of techniques such as early stopping (Prechelt, 1998) or sparsity-inducing regularization (Schneider et al., 2020) becomes necessary. The compact support property of EKI, which mandates that the solution be in the linear span of the initial ensemble, also regularizes the learned hybrid model with decreasing  $J$ ; for  $J = 50 < p$  overfitting is significantly reduced. Thus, reducing the ensemble size is an efficient regularization technique when training large machine-learning models that tend to overfit, at the expense of reduced expressivity. Additional EKI-specific regularization techniques for deeper networks are discussed in Kovachki and Stuart (2019).

Another difference between the empirical and the hybrid models is that for the latter, we do not know a priori the parameter ranges for which the model trajectories remain physical. During the training sessions shown in Figure 4.8, the hybrid models experienced a maximum of 25 ( $J = 50$ ), 30 ( $J = 100$ ) and 72 ( $J = 159$ ) failures in a single iteration, all occurring during the first epoch. The use of the failsafe update

proposed in Section 4.4 proved crucial to enable training in the presence of model failures, and it reduced the number of failures to a small fraction of the  $J$  ensemble members after a few EKI iterations.

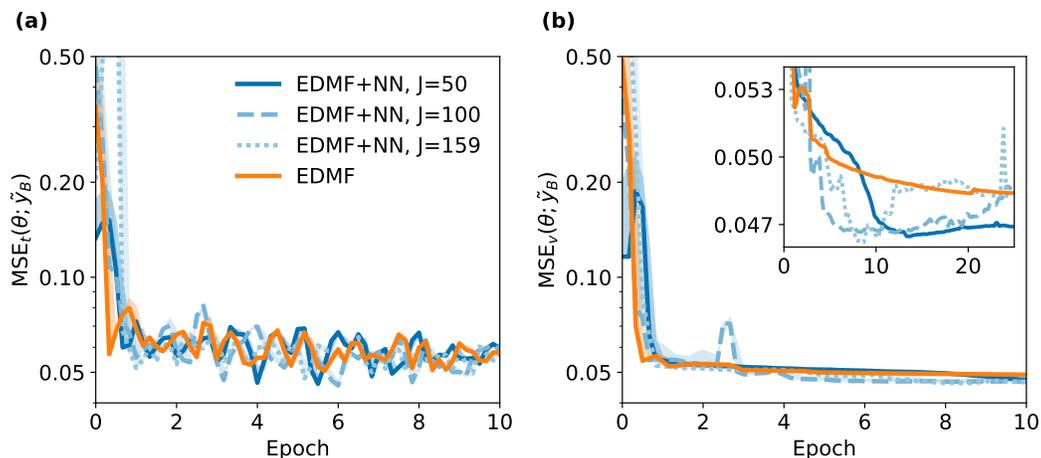


Figure 4.8: Batch (a) training and (b) validation normalized MSE for the hybrid (EDMF+NN) and empirical (EDMF) models. Lines, shading and inset as in Figure 4.2. Results are shown for calibration with EKI, using  $J = 50, 100$  and  $2p + 1 = 159$  ensemble members for the hybrid model. The empirical model training uses  $J = 2p + 1 = 33$ . All inversion processes use batch size  $|B| = 10$ .

Profiles of  $\bar{q}_l$ ,  $\bar{q}_t$  and  $\overline{w'q'_t}$  are shown in Figure 4.9 for the trained empirical and hybrid EDMF models. To produce the profiles with the hybrid model, we retain the parameters learned at the iteration with lowest validation error from a training session spanning 25 epochs, effectively similar to early stopping. As expected from the validation error, the hybrid model slightly improves upon the skill of the empirical model, predicting more accurate profiles of  $\bar{q}_l$  within the cloud layer. This is, of course, at the cost of a significantly higher parameter complexity of the closure.

As shown here, ensemble Kalman inversion allows for rapid prototyping and comparison of closures within an overarching black-box model. Importantly, this comparison can be done during training in terms of the online performance of the fully calibrated dynamical model.

#### 4.6 Discussion and conclusions

Ensemble Kalman methods such as ensemble and unscented Kalman inversion are powerful tools for training possibly expensive models. By leveraging covariances between the model output and its parameters, they do not impose any constraint

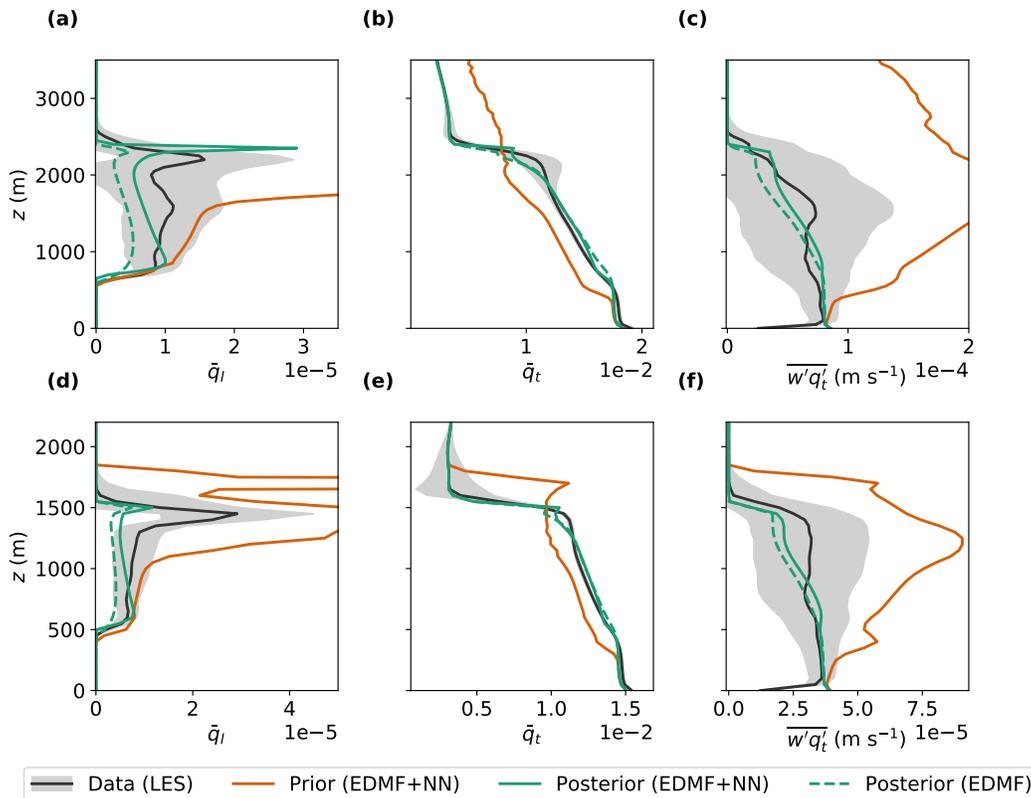


Figure 4.9: Prior, posterior and LES profiles of liquid water specific humidity ( $\bar{q}_l$ ), total water specific humidity ( $\bar{q}_t$ ) and vertical moisture flux ( $\overline{w'q'_t}$ ) for cfSite 14 using July forcing (top) and cfSite 8 using January forcing (bottom) from the AMIP4K experiment (Shen et al., 2022). Definitions of prior, posterior and shading as in Figure 4.5. Posterior results are shown for the EDMF model with empirical closures (EDMF), and with the neural network entrainment closure (4.35) (EDMF+NN), using early stopping and 25 epochs of training. Results obtained using EKI with  $|B| = 10$ .

on the data used for learning, or the architecture of the closures to be calibrated. This means that ensemble Kalman methods can be used to learn all parameters within complex overarching models, regardless of where those parameters appear in the formulation of the model. Furthermore, the Gaussian approximation of the parameter distribution makes them far more efficient than standard Bayesian inference techniques, at the cost of neglecting uncertainty beyond the second moment of the posterior, and the possible convergence to local minima (as for stochastic gradient descent and other optimization methods).

This enables training physics-based machine-learning parameterizations, as demonstrated here by substituting an internal component of the EDMF model by a neural network, which required no change in the data or framework used for training. The

benefits of combining physics and data are demonstrated by the performance of our trained hybrid closure in simulations of clouds typical of conditions 4 K warmer than the clouds in the training set.

To use these algorithms, parameter learning must be framed as an inverse problem. This allows great flexibility, but raises the problem of choosing a reasonable observational map  $\mathcal{H}$  and noise covariance  $\Gamma$  to define an inverse problem. Through a domain-agnostic strategy and a reasonable heuristic about the expected model error, we have demonstrated a systematic way of constructing a well-defined inverse problem from high-dimensional data. This strategy is designed to maximize the information content through a lossy principal component encoding  $\mathcal{H}$  and to allow the use of time averages as observations, making it amenable to harnessing, e.g., satellite observations in addition to computationally generated data. The success of this strategy is demonstrated in a variety of settings, using empirical and hybrid models.

The flexibility of the inverse problem allows to define the observational map  $\mathcal{H}$  through any observable diagnostic of the model, be it differentiable or not. For instance, Barthélémy et al. (2021) use a neural network as the mapping  $\mathcal{H}$ , to train a low-resolution dynamical model directly from features at high resolution. One could also envision the construction of  $\mathcal{H}$  through other statistics of the model dynamics, such as the variance or skewness. These choices may be preferable for particular tasks, such as the prediction of extreme events or the correct representation of emergent phenomena.

Given an inverse problem, we have shown that EKI and UKI are robust to noise and amenable to batching strategies. This establishes the ability of the Kalman algorithms to train models using sequentially sampled data. The same robustness can be expected for other sources of noise, such as stochasticity in the model (Schneider et al., 2021a). In addition, we have proposed modifications of the EKI and UKI updates that enable calibrating models that may fail during training, which is often the case for Earth system models.

Although similar, each ensemble Kalman algorithm presents its own relative strengths in our analysis. Calibration through EKI appears to be more robust to noise, and the number of ensemble members may be chosen to be lower than for UKI when the parameter space is high-dimensional. Indeed, Kovachki and Stuart (2019) show successful results for EKI when the number of parameters (e.g.,  $p \sim 10^6$ ) is two orders of magnitude higher than the ensemble size. Using fewer ensemble members

than parameters also introduces a regularization effect. On the other hand, UKI provides information about parametric uncertainty and correlations, which can be used to improve models at the process level, and to rapidly compare the added value of increasingly precise observing systems. Other ensemble Kalman methods, such as the sparsity-inducing EKI (Schneider et al., 2020) or the ensemble Kalman sampler (Garbuno-Inigo et al., 2020), can provide solutions to the inverse problem with other useful properties. In addition, all these ensemble methods generate parameter-output pairs that can be used to train emulators for uncertainty quantification that can capture non-Gaussian posteriors (Cleary et al., 2021).

Finally, ensemble Kalman methods may be used for the rapid comparison of parameterizations in terms of the online skill of an overarching Earth system model. The same framework could be used to train Gaussian processes, random feature models (Nelsen and Stuart, 2021), Fourier neural operators (Z. Li et al., 2020), or stochastic closures (Guillaumin and Zanna, 2021), for example. These are some of the exciting research avenues that we will be exploring in the future.

#### **4.7 Appendix A: Configuration-based principal component analysis**

Performing PCA on each configuration allows retaining principal modes from low-variance configurations while filtering out trailing modes from high-variance configurations. The importance of this is demonstrated in Figure 4.10 for three configurations of our LES solver (Pressel et al., 2015) based on observational campaigns of a stable boundary layer, a stratocumulus-topped boundary layer, and shallow cumulus convection (Beare et al., 2006; Stevens et al., 2005; Siebesma et al., 2003). Performing global PCA is equivalent to using a cutoff  $\mu_{c,i} > \mu_c^*$  in Figure 4.10a, where we need to choose between neglecting most modes from certain configurations (e.g., GABLS in Figure 4.10a) or retaining highly oscillatory modes from others (e.g., Bomex), as measured by the number of zero-crossings of the eigenmode (Hansen, 1998). Highly oscillatory modes may have a disproportionate contribution to the loss when calibrating imperfect models. On the other hand, performing PCA on each  $\tilde{\Gamma}_c$  alleviates this problem by aligning the eigenspectra before applying the cutoff, as shown in Figure 4.10b. Appropriate conditioning of the global covariance matrix is still enforced when applying configuration-based PCA through the Tikhonov regularizer in equation (4.14).

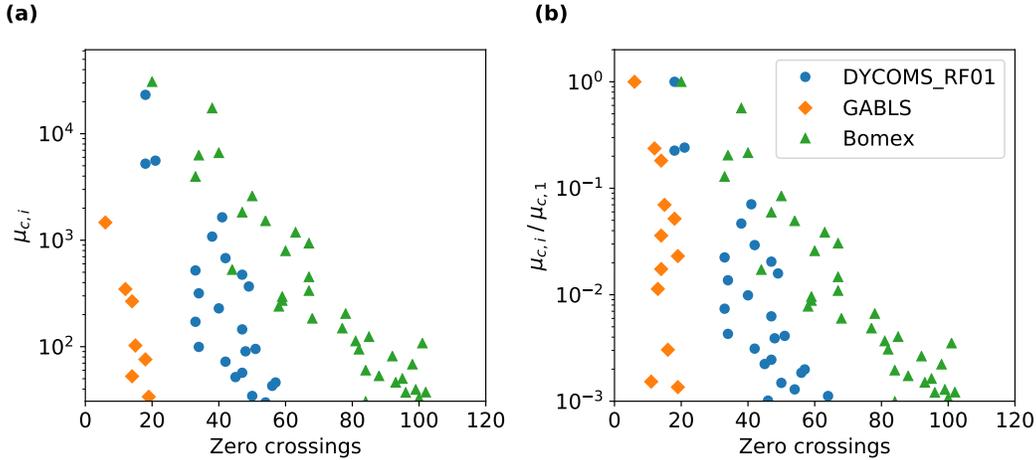


Figure 4.10: (a) Scatter plot of covariance eigenvalues  $\mu_{c,i}$  and the number of zero-crossings of their corresponding eigenmode for three different configurations of an LES solver. (b) The same plot, with eigenvalues normalized by the leading eigenvalue of each configuration ( $\mu_{c,1}$ ). Trailing eigenvalues are associated with high-wavenumber oscillatory modes with frequent sign changes.

#### 4.8 Appendix B: Addressing model failures with unscented Kalman inversion

In the presence of model failures, we perform the UKI quadratures over the successful sigma points. Consider the set of off-center sigma points  $\{\hat{\theta}\} = \{\hat{\theta}_s\} \cup \{\hat{\theta}_f\}$  where  $\hat{\theta}_s^{(j)}$ ,  $j = 1, \dots, J_s$  are successful members and  $\hat{\theta}_f^{(k)}$  are not. For ease of notation, consider an ordering of  $\{\hat{\theta}\}$  such that  $\{\hat{\theta}_s\}$  are its first  $J_s$  elements, and note that we deal with the central point  $\hat{\theta}^{(0)}$  separately. We estimate the covariances  $\text{Cov}_q(\mathcal{G}_n, \mathcal{G}_n)$  and  $\text{Cov}_q(\theta_n, \mathcal{G}_n)$  from the successful ensemble,

$$\text{Cov}_q(\theta_n, \mathcal{G}_n) \approx \sum_{j=1}^{J_s} w_{s,j} (\hat{\theta}_{s,n}^{(j)} - \bar{\theta}_{s,n}) (\mathcal{G}(\hat{\theta}_{s,n}^{(j)}) - \bar{\mathcal{G}}_{s,n})^T, \quad (4.37)$$

$$\text{Cov}_q(\mathcal{G}_n, \mathcal{G}_n) \approx \sum_{j=1}^{J_s} w_{s,j} (\mathcal{G}(\hat{\theta}_{s,n}^{(j)}) - \bar{\mathcal{G}}_{s,n}) (\mathcal{G}(\hat{\theta}_{s,n}^{(j)}) - \bar{\mathcal{G}}_{s,n})^T, \quad (4.38)$$

where the weights at each successful sigma point are scaled up, to preserve the sum of weights,

$$w_{s,j} = \left( \frac{\sum_{i=1}^{2p} w_i}{\sum_{k=1}^{J_s} w_k} \right) w_j. \quad (4.39)$$

In equations (4.37) and (4.38),  $\bar{\theta}_{s,n}$  and  $\bar{\mathcal{G}}_{s,n}$  must be modified from the original formulation if the central point  $\hat{\theta}^{(0)} = m_n$  results in model failure,

$$\bar{\theta}_{s,n} = \begin{cases} m_n & \text{if } \hat{\theta}^{(0)} \text{ successful,} \\ \frac{1}{J_s} \sum_{j=1}^{J_s} \hat{\theta}_{s,n}^{(j)} & \text{otherwise,} \end{cases} \quad (4.40)$$

$$\bar{\mathcal{G}}_{s,n} = \begin{cases} \mathcal{G}(m_n) & \text{if } \hat{\theta}^{(0)} \text{ successful,} \\ \frac{1}{J_s} \sum_{j=1}^{J_s} \mathcal{G}(\hat{\theta}_{s,n}^{(j)}) & \text{otherwise.} \end{cases} \quad (4.41)$$

These modified UKI quadrature rules are used throughout Section 4.5 to deal with model failures. Since UKI can be initialized from a tighter prior than EKI, due to the absence of ensemble collapse, failures are much easier to avoid than with EKI.

#### 4.9 Appendix C: Parameter transformation and prior

Given a prior range  $[\phi_i, \phi_f]$  for a parameter  $\phi \in \mathbb{R}$ , we define the transformation

$$\theta = \mathcal{T}(\phi) = \ln \frac{\phi - \phi_i}{\phi_f - \phi}, \quad (4.42)$$

such that the interval midpoint is mapped to  $\theta = 0$ , and the bounds to  $\pm\infty$ . An unconstrained Gaussian prior may then be defined for  $\theta$  given the prior mean in physical (constrained) parameter space  $\phi_p$  as

$$\theta_0 \sim \mathcal{N}(\mathcal{T}(\phi_p), \sigma_0^2), \quad (4.43)$$

where  $\sigma_0^2$  is a free parameter controlling the size of the region within the interval  $[\phi_i, \phi_f]$  containing most of the probability. This means that the magnitude of  $\sigma_0$  is already normalized with respect to the prior range, so we will generally choose  $\sigma_0 \sim \mathcal{O}(1)$ . The  $p$ -dimensional prior  $\mathcal{N}(m_0, \Sigma_0)$  is then constructed as an uncorrelated multivariate normal with marginal distributions given by expression (4.43). The normalization induced by (4.42) also enables the use of isotropic regularization in equations (4.25)–(4.26), even though the physical parameters  $\phi$  may differ in order of magnitude. For more examples of parameter transformations in the context of EKI and UKI, see D. Z. Huang et al. (2022a), Schneider et al. (2021b), and Dunbar et al. (2022).

#### 4.10 Appendix D: Calibration using very noisy loss evaluations

The Kalman inversion results are expected to deteriorate above some noise threshold, as the signal-to-noise ratio in the training process decreases. We explored the

sensitivity of UKI and EKI to noise by sampling a single configuration per iteration from the training set described in Section 4.5. As shown in Figure 4.11, UKI fails to converge to the minimum found with larger batches in this limit. The validation error is characterized by large oscillations due to strong changes in the value of model parameters like the entrainment coefficient  $c_\epsilon$  or the eddy diffusivity coefficient  $c_m$ . On the other hand, EKI proves robust to noise even in this limit, converging to the minimum found by UKI employing larger batches.

In the context of Kalman inversion, decreasing the step size  $\Delta t$  is equivalent to increasing the noise variance, as shown in updates (4.18) and (4.27). We investigate the time step role in the small batch limit by performing the ensemble Kalman inversion with  $\Delta t = |C|^{-1} = 1/60$ . The smaller time step increases the parameter uncertainty, which leads to a reduction in parameter oscillations and estimates closer to the prior. This is accompanied by a moderate reduction in validation error oscillations. We performed additional inversions using even smaller time steps, which resulted in a convergence of the parameter estimates towards the prior and a minor reduction in validation error with respect to the initialization. We conclude that decreasing  $\Delta t$  in UKI can reduce oscillations due to high levels of noise, but it does not result in the same robustness as EKI.

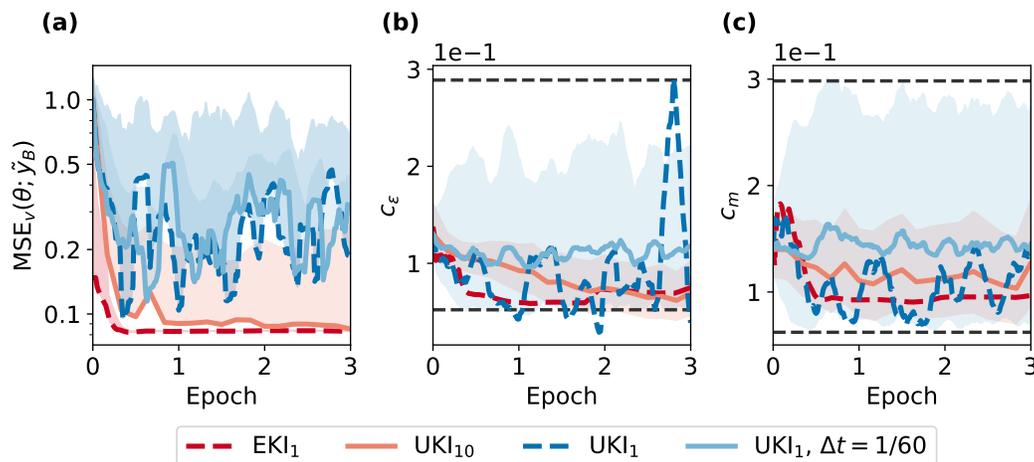


Figure 4.11: Evolution of the validation error (a) and estimates of the entrainment (b), and eddy diffusivity (c) coefficients. Results shown for UKI using batch sizes of 10 and 1, and EKI using a batch size of 1. Parameter uncertainty only shown for UKI<sub>10</sub> and UKI<sub>1</sub>,  $\Delta t = 1/60$  for clarity. All results shown use  $\Delta t = |C|/|B|$  unless otherwise specified. Shading as in Figures 4.2 and 4.3.

#### 4.11 List of Symbols

Table 4.2: List of all symbols used in the chapter, their dimensions when applicable, and a short description.

Symbol	Description
$\phi \in \mathbb{R}^p$	Learnable parameters, in physical space
$\theta \in \mathbb{R}^p$	Transformed learnable parameters, in unconstrained space
$\theta^* \in \mathbb{R}^p$	Optimal unconstrained parameter estimate (MAP or MLE)
$\varphi_0$	Initial dynamical state
$F_\varphi$	Dynamical forcing
$x_c = \{\varphi_0, F_\varphi\}_c$	Configuration of the dynamical system
$\zeta(x_c) : \varphi_0 \rightarrow \varphi(t)$	True dynamical system evolution
$\Psi(\phi; x_c) : \varphi_0 \rightarrow \hat{\varphi}(t)$	Dynamical model evolution
$\mathcal{H}_c$	Observational map for configuration $c$
$y_c \in \mathbb{R}^{d_c}$	Observation vector for configuration $c$
$\eta_c \in \mathbb{R}^{d_c}$	Observation error for map $\mathcal{H}_c$
$\delta(x_c) \in \mathbb{R}^{d_c}$	Model or representation error for configuration $c$
$\Gamma_c \in \mathbb{R}^{d_c \times d_c}$	Covariance of the Gaussian noise $\eta_c + \delta(x_c)$
$\mathcal{G}_c : \mathbb{R}^p \rightarrow \mathbb{R}^{d_c}$	Forward model for configuration $c$
$C = \{x_c, c = 1, \dots,  C \}$	Set of configurations
$y = [y_1, \dots, y_{ C }]^T \in \mathbb{R}^d$	Global observation vector
$\delta = [\delta(x_1), \dots, \delta(x_{ C })]^T$	Global representation error
$\eta = [\eta_1, \dots, \eta_{ C }]^T$	Global observation error
$\Gamma \in \mathbb{R}^{d \times d}$	Global noise covariance matrix
$\mathcal{T} : U \rightarrow \mathbb{R}^p$	Parameter transformation to unconstrained space
$\mathcal{G} : \mathbb{R}^p \rightarrow \mathbb{R}^d$	Forward model
$\rho(\theta y, \Gamma)$	Parameter posterior probability density, given $\Gamma$ and $y$
$\rho_{\text{prior}}(\theta)$	Parameter prior probability density, independent of $\Gamma$
$\mathcal{L} : \mathbb{R}^p \times \mathbb{R}^d \rightarrow \mathbb{R}$	Loss or negative log-likelihood given $\Gamma$
$S_c(t) \in \mathbb{R}^{\tilde{d}_c}$	Observed state
$V_{c,j}(t) \in \mathbb{R}^{h_c}$	Spatial field $j$ within the observed state $S_c$
$s_c(t) \in \mathbb{R}^{\tilde{d}_c}$	Normalized observed state
$v_{c,j}(t) \in \mathbb{R}^{h_c}$	Forward model for configuration $c$
$\sigma_{c,j} \in \mathbb{R}$	Pooled time standard deviation of $V_{c,j}$
$T_c \in \mathbb{R}$	Time-averaging window used in map $\mathcal{H}_c$
$\tilde{y}_c \in \mathbb{R}^{\tilde{d}_c}$	Counterpart of $y_c$ prior to encoding
$\tilde{y} \in \mathbb{R}^{\tilde{d}}$	Global observation vector prior to encoding
$\tilde{\Gamma}_c \in \mathbb{R}^{\tilde{d}_c \times \tilde{d}_c}$	Counterpart of $\Gamma_c$ prior to encoding
$\tilde{\Gamma} \in \mathbb{R}^{\tilde{d} \times \tilde{d}}$	Counterpart of $\Gamma$ prior to encoding

<sup>a</sup> Continued on the next page.

Table 4.3: (cont.) List of all symbols used in the chapter.

Symbol	Description
$I_d \in \mathbb{R}^{d \times d}$	Identity matrix of size $d \times d$
$\mu_{c,i} \in \mathbb{R}$	$i$ -th largest eigenvalue of $\tilde{\Gamma}_c$
$\kappa \in \mathbb{R}$	Approximate condition number of a matrix
$r_c \in \mathbb{R}$	Approximate rank of matrix $\tilde{\Gamma}_c$
$\epsilon_m \in \mathbb{R}$	Machine or data precision
$\kappa_* < \epsilon_m^{-1/2}$	Limiting matrix condition number
$P_c \in \mathbb{R}^{\tilde{d}_c \times d_c}$	Truncated PCA projection matrix
$D\mathcal{G}(\theta) \in \mathbb{R}^{d \times p}$	Jacobian of forward model at $\theta$
$B = \{x_c, c = 1, \dots,  B \}$	Mini-batch of configurations
$L : \mathbb{R}^p \times \mathbb{R}^d \rightarrow \mathbb{R}$	Configuration-averaged loss
$y_B \in \mathbb{R}^{d_B}$	Observation vector for batch $B$
$\tilde{y}_B \in \mathbb{R}^{\tilde{d}_B}$	Counterpart of $y_B$ prior to encoding
$\tilde{\mathcal{G}}_B : \mathbb{R}^p \rightarrow \mathbb{R}^{\tilde{d}_B}$	Forward model corresponding to observations $\tilde{y}_B$
$\Theta_n \in \mathbb{R}^{p \times J}$	Parameter ensemble at iteration $n$
$m_n \in \mathbb{R}^p$	Mean parameter estimate at iteration $n$
$\Sigma_n \in \mathbb{R}^{p \times p}$	Parameter covariance estimate at iteration $n$
$\mathcal{G}_{\Theta_n} \in \mathbb{R}^{d \times J}$	Forward model evaluation ensemble at iteration $n$
$\mathcal{E}(\Theta_n) \in \mathbb{R}^{d \times J}$	Data-model mismatch ensemble at iteration $n$
$\Delta t \in \mathbb{R}$	Nominal learning rate
$\Theta_{s,n} \in \mathbb{R}^{p \times J_s}$	Successful parameter ensemble at iteration $n$
$\theta_{f,n}^{(k)} \in \mathbb{R}^p$	$k$ -th failed parameter vector at iteration $n$
$m_p \in \mathbb{R}^p$	Parameter prior mean
$\Lambda \in \mathbb{R}^{p \times p}$	Gaussian prior covariance
$y_a \in \mathbb{R}^{d+p}$	Observation vector augmented with $m_p$
$\mathcal{G}_a(\theta) \in \mathbb{R}^{d+p}$	Forward model augmented with $\theta$
$\xi \in \mathbb{R}^{d+p}$	Aggregate noise in the augmented data-model relation
$\Gamma_a \in \mathbb{R}^{(d+p) \times (d+p)}$	Covariance of the aggregate noise $\xi$
$\hat{\theta}_n^{(j)} \in \mathbb{R}^p$	$j$ -th sigma point for UKI quadrature
$\Pi_j$	$j$ -th nondimensional input to neural network

*Chapter 5***SKILLFUL REPRESENTATION OF LOW CLOUD REGIMES  
WITH THE EXTENDED EDDY-DIFFUSIVITY MASS-FLUX  
SCHEME****5.1 Abstract**

Unresolved tropospheric mixing processes remain one of the leading sources of uncertainty in climate projections, due to their strong influence on low-cloud cover. Low clouds cool the Earth, due to their high albedo and small net contribution to the greenhouse effect. Although there is growing consensus about the positive sign of the low-cloud climate feedback, its magnitude remains highly uncertain due in part to the inability of current climate models to represent atmospheric turbulence and convection accurately. Here, we demonstrate the ability of an extended eddy-diffusivity mass-flux (EDMF) scheme to model these atmospheric transport processes, and to capture their aggregate effect on the thermodynamic structure of the low troposphere over a wide range of tropical and subtropical low lying cloud regimes. The extended EDMF scheme is calibrated using large-eddy simulations (LES) forced by present day simulations from a single climate model, and its performance is evaluated against LES of clouds forced by multiple climate models, and representative of different baseline climates. We find that the extended EDMF scheme is able to accurately reproduce important aspects of the stratocumulus-to-cumulus transition over the eastern Pacific Ocean in a single-column basis, and to strongly generalize to forcing conditions outside the training data set. We further show that the response of the extended EDMF scheme to a uniform 4 K warming of Earth's surface is consistent with the LES response for subgrid-scale transport and cloud vertical extent, but less so for cloud fraction and liquid water path.

## 5.2 Introduction

Turbulent mixing and convective overturning in the lower troposphere, together with aerosols and microphysical processes, remain the leading source of uncertainty in climate projections, largely due to their prominent role in the dynamical balance that sustains tropical and subtropical low lying clouds (Bony and Dufresne, 2005; Boucher et al., 2013; Sherwood et al., 2014; Schneider et al., 2017a). Low clouds shield Earth’s surface from a sizable fraction of the incoming shortwave radiation, while contributing little to greenhouse warming due to their relatively high emission temperature (Hartmann et al., 1992). There is growing consensus about the fact that low clouds represent a positive climate feedback; higher surface temperatures lead to low-cloud cover thinning, which then fosters a further increase in surface temperatures as more sunlight reaches the surface (Bretherton et al., 2013; Bretherton, 2015; Zelinka et al., 2016; Zelinka et al., 2017). However, the magnitude of this feedback mechanism remains highly uncertain due in part to the inability of current models to predict changes in low tropospheric mixing accurately (Sherwood et al., 2014; M. Zhao, 2014; Webb et al., 2015; Gettelman and Sherwood, 2016).

For this reason, considerable effort is still devoted to improving and unifying parameterizations of turbulent and convective processes that act beyond the necessarily coarse resolution of current climate models (Han and Bretherton, 2019; Suselj et al., 2019a; Cohen et al., 2020; Lopez-Gomez et al., 2020; E. Wu et al., 2020; Thuburn et al., 2022a; Thuburn et al., 2022b; Neggers and Griewank, 2022; Shin and Baik, 2022). One particularly promising parameterization approach is the extended eddy-diffusivity mass-flux (EDMF) scheme, which generalizes EDMF schemes (Siebesma and Teixeira, 2000; Siebesma et al., 2007; Neggers, 2009) by considering updrafts with variable area fraction, and retaining convective memory through transient terms (Tan et al., 2018; Cohen et al., 2020).

The extended EDMF scheme, coupled with generalized formulations of eddy diffusion, entrainment, and detrainment, has been shown to skillfully reproduce the structure of Arctic stable boundary layers, marine stratocumulus clouds, shallow cumuli, and deep cumulonimbus clouds; all with a single set of  $\sim 20$  parameters (Lopez-Gomez et al., 2020; Cohen et al., 2020). Aside from demonstrations of its ability to capture the dynamics of these archetypical regimes, whether the model can simulate transitions between them or perform just as well in less idealized settings remains to be seen.

In this context, the library of large-eddy simulations (LES) of Shen et al. (2022)

provides a useful testbed to validate the predicted response of subgrid-scale (SGS) parameterizations to large-scale conditions typical of the stratocumulus-to-cumulus transition (SCT) in the eastern Pacific Ocean, where cloud cover biases in climate models are particularly large (J.-L. Lin et al., 2014; Brient et al., 2019; Konsta et al., 2022; H.-H. Lee et al., 2022). The library includes simulations of the SCT characteristic of all seasons in the current climate, and simulations over the same geographical region forced by large-scale conditions representative of a climate with 4 K warmer surface temperatures. The library simulations are each forced by one of three operational climate models, spanning two generations of the Coupled Model Intercomparison Project (HadGEM-2A and CNRM-CM5 from CMIP5, and CNRM-CM6-1 from CMIP6).

Here, we leverage the library of Shen et al. (2022) to investigate the skill of the extended EDMF scheme at reproducing the thermodynamic structure of the lower troposphere under forcings typical of the eastern Pacific SCT, as simulated by LES. Since the large-scale forcings are derived from climate model output, this study represents a more realistic assessment of the performance of the scheme than tests that focus on idealized cases (Siebesma et al., 2003; Stevens et al., 2005). To evaluate the robustness and generalization of the results, the extended EDMF scheme is trained on simulations forced by current-climate large-scale conditions from a single climate model, and evaluated against current and warmer climate conditions simulated by all three models available in the library. This strategy also enables comparing the response of the scheme and LES to climate perturbations not seen during training; a particularly difficult task for data-driven parameterizations of turbulence and convection (O’Gorman and Dwyer, 2018; Rasp et al., 2018).

We analyze the response in a single-column basis, which has both advantages and disadvantages with respect to performing global simulations. Single-column simulations enable comparing the SGS scheme output to LES, where the only parameterized processes are those related to cloud microphysics. This methodology is only possible at the single-column scale due to computational constraints (Schneider et al., 2017a). In addition, single-column models (SCMs) represent a controlled environment for targeted analysis of the merits and deficiencies of the extended EDMF scheme; free of interactions with confounding sources of error, such as large-scale circulation feedbacks or parameterized radiative transfer processes (Pincus et al., 2003; Cole et al., 2005; Klinger et al., 2017; Singer et al., 2021).

From this advantage also comes their most important limitation, namely the po-

tential mismatch between the response of the scheme in isolation and in a fully coupled Earth System Model (M. Zhang et al., 2012; M. Zhang et al., 2013). We acknowledge this and suggest interpreting our results as characterizing the potential of the extended EDMF scheme to track the true LES response to a wide range of realistic large-scale forcings. Beyond this, Brient et al. (2019) demonstrated through short term hindcasts that a large fraction of the bias in the representation of marine stratocumulus in CNRM-CM6-1 is due to local parameterized processes and independent of feedbacks with the large-scale circulation; Dal Gesso and Neggers (2018) found that the single-column response to a climate perturbation of the EC-EARTH model largely agrees with the response of the same column in a global model; and Hwang et al. (2022) recently showed that the single-column response of 5 popular convection schemes is representative of their multi-column response in the absence of convective aggregation. These studies suggest that in the SCT region, the single-column response may be a good proxy for the expected behavior of the scheme in a global simulation.

The article is organized as follows. Section 5.3 describes the LES data used as the ground truth, the methodology used to drive a single-column model (SCM) implementation of the EDMF scheme using the forcing seen by the LES, and our calibration and evaluation strategy. Section 5.4 compares the EDMF and LES response across regimes within the SCT regions in the eastern Pacific. The climate change response of the EDMF scheme is then analyzed in Section 5.5, and Section 5.6 ends with a discussion of our findings and some conclusions.

### **5.3 Data and methods**

We focus our analysis on a set of marine locations along two eastern Pacific transects, one in the subtropics connecting the coasts of California and Hawai'i (Teixeira et al., 2011), and another one in the South Tropical Pacific between the Chilean-Peruvian border and French Polynesia (Wood et al., 2011). These cross sections, shown in Figure 5.1, span the dynamical regime transition from the stratocumulus-topped boundary layers found off the west coast of continents to the shallow cumulus convective regions farther away from the coast. High-frequency climate model data at these locations are available as part of the cfSites product of CMIP, starting from Phase 5 (Webb et al., 2015).

We consider data from the AMIP and AMIP4K experiments, which are representative of present-day and a 4 K uniformly warmer climate, respectively. The large-scale

conditions along these transects are characterized by surface latent heat fluxes on the order of  $50 - 200 \text{ Wm}^{-2}$  increasing from east to west, subsidence reaching up to  $1 \text{ cm s}^{-1}$  in the stratocumulus regions, and advective cooling and drying of the boundary layer (Shen et al., 2022).

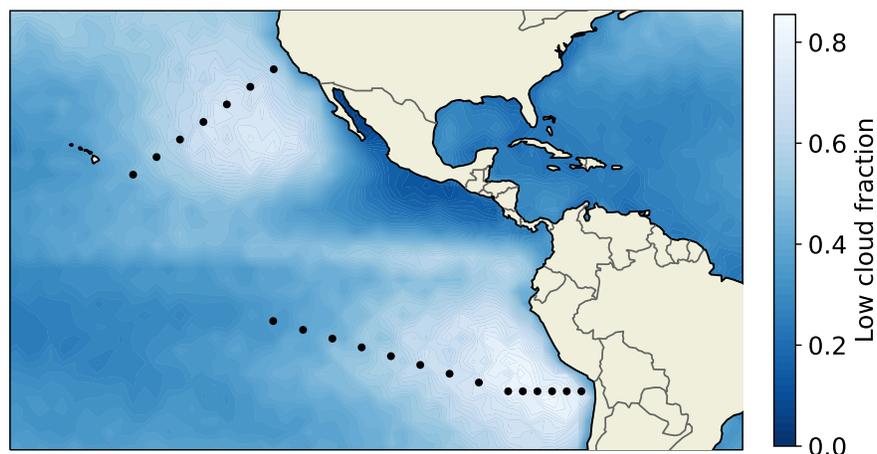


Figure 5.1: Location of cfSites considered in this study (black dots), and average observed low-level cloud fraction over the ocean during the period 2007-2012. Low-level cloud fraction is derived from CALIPSO-GOCCP for clouds below 680 hPa (Chepfer et al., 2010).

### LES Library

The data used as ground truth come from limited-area high-resolution simulations described in detail in Shen et al. (2022), and generated using the PyCLES solver (Pressel et al., 2015; Pressel et al., 2017). PyCLES is an anelastic fluid dynamics solver with weighted essentially non-oscillatory numerics that uses specific entropy and total water specific humidity as prognostic variables. Cloud microphysical processes are parameterized in PyCLES through the one-moment warm-rain scheme of Kessler (1995). The LES are run using time-averaged large-scale forcing from the cfSites output of three climate models: HadGEM-2A, CNRM-CM5, and CNRM-CM6-1. The first two are chosen as representative of models with a high and low tropical low-cloud albedo response to global warming, respectively (Brient and Schneider, 2016); CNRM-CM6-1 is chosen as representative of changes across CMIP model generations (Saint-Martin et al., 2021).

For each model and location, 8 high-resolution simulations are performed: 4 using forcings derived from the AMIP experiment, and another 4 forced using the AMIP4K experiment. The 4 simulations per experiment are each driven by 5-year (2004-2008)

averaged large-scale forcings corresponding to the climatology of January, April, July, and October. Using time-averaged forcing allows for a shorter equilibration time to statistically steady state, and it results in reduced internal variability, so that climatological solutions can be computed using much shorter integration times (Shen et al., 2022). Radiative forcing is computed interactively using the Rapid Radiative Transfer Model (RRTM, Mlawer et al., 1997), and surface fluxes are diagnosed from a bulk scheme based on Monin-Obukhov similarity theory (Byun, 1990). To focus our analysis on low clouds, only simulations that resulted in clouds with tops below 3 km are considered; yielding a grand total of 406 unique simulations.

### SCM Experimental Setup

The extended EDMF scheme is implemented in an SCM, covering the same vertical extent as the LES, with a domain top at 4 km. The prognostic thermodynamic variables of the SCM implementation are liquid water potential temperature and total specific humidity. The closures for eddy diffusion, entrainment fluxes, and perturbation pressure gradients are described in Lopez-Gomez et al. (2020), Cohen et al. (2020) and J. He et al. (2021), respectively. The closures used for microphysical processes are identical to the ones used in LES (Kessler, 1995).

We consider two SCM configurations, which we denote ESM-EDMF and ISO-EDMF. The first configuration is characteristic of the vertical resolution of CMIP6 Earth System Models (ESMs), with 55 levels from the surface to  $z_t = 45$  km (Ingram and Bushell, 2021), a near-surface resolution of  $\Delta z_s = 30$  m, and a top-of-atmosphere resolution of  $\Delta z_t = 8$  km. The definition of the grid stretching and the truncation procedure used to reach the 4 km domain top exactly are described in Section 5.7. After truncation, the SCM domain in the ESM-EDMF configuration has 28 degrees of freedom within the bottom 4 km of the atmosphere.

The ISO-EDMF configuration characterizes the response of the extended EDMF scheme at a significantly higher isotropic resolution of 20 m, or 200 degrees of freedom within the domain. Comparison between the two configurations quantifies the fraction of structural error that may be reduced by increasing the vertical resolution, which has been shown to be substantial when using the CLUBB parameterization scheme (H.-H. Lee et al., 2022). In addition, this configuration enables comparison with studies that employ a similar resolution to evaluate the performance of turbulence and convection schemes (Smalley et al., 2022).

The SCM simulations are initialized from horizontal and time averages of the LES profiles of liquid water potential temperature  $\langle \theta_l \rangle$ , total water specific humidity  $\langle q_t \rangle$ , and horizontal velocity  $\langle \mathbf{u}_h \rangle$ ;  $\langle \cdot \rangle$  denotes horizontal averaging. The initial profiles are averaged using a 1 hour window well after a statistically steady state is reached in the LES. SCM simulations are run for 6 hours after initialization. The SCM forcing follows closely the LES forcing described in Shen et al. (2022), with differences regarding the surface and radiative forcing:

- Horizontal advection and vertical eddy advection are identical to the tendencies used in the LES, and come from the time-averaged cfSites output.
- The same constant large-scale subsidence  $\tilde{w}$  is applied to the moist adiabatic variables  $\langle \theta_l \rangle$  and  $\langle q_t \rangle$ . Here,  $\tilde{(\cdot)}$  represents a time-averaged prescribed value.
- Due to the absence of momentum forcing in the cfSites product, horizontal winds are relaxed to time-averaged profiles from the cfSites output at a frequency of  $4 \text{ day}^{-1}$ . The temperature and specific humidity in the free troposphere above  $z_r = 3.5 \text{ km}$  are relaxed to the cfSites profiles at a frequency of  $1 \text{ day}^{-1}$ , and with a sinusoidally decreasing frequency below  $z_r$  vanishing at  $z_r - 0.5 \text{ km}$ , following Shen et al. (2022). All relaxation profiles are the same as those used to force the LES.
- The radiative tendencies and surface fluxes are constant and equal to the final 6-hour averaged LES values, which ensures that these terms have the same net energetic contribution in the LES and SCM runs.

This LES-SCM forcing framework, originally established in Lopez-Gomez et al. (2022b), is designed to investigate whether the SGS fluxes and microphysical sources given by the extended EDMF scheme can achieve a dynamical equilibrium with the large-scale forcing similar to the one attained by the LES. Prescribing the radiative fluxes facilitates this comparison by eliminating contributions from structural errors in radiative transfer parameterizations (Pincus et al., 2003; Singer et al., 2021). The steady-state dynamical balances enforced by this framework below the free troposphere may be written as

$$\tilde{w} \frac{\partial \langle q_t \rangle}{\partial z} + \frac{1}{\rho} \frac{\partial}{\partial z} (\rho \langle w^* q_t^* \rangle) - S_{\text{mp}} = \tilde{S}_{\text{hadv}} + \tilde{S}_{\text{veddy}}, \quad (5.1)$$

$$\tilde{w} \frac{\partial \langle \theta_l \rangle}{\partial z} + \frac{1}{\rho} \frac{\partial}{\partial z} (\rho \langle w^* \theta_l^* \rangle) - \Pi^{-1} Q_{\text{mp}} = \Pi^{-1} (\tilde{Q}_{\text{rad}} + \tilde{Q}_{\text{hadv}} + \tilde{Q}_{\text{veddy}}), \quad (5.2)$$

$$\frac{1}{\rho} \frac{\partial}{\partial z} (\rho \langle w^* \mathbf{u}_h^* \rangle) - \Gamma_m (\tilde{\mathbf{u}}_h - \langle \mathbf{u}_h \rangle) = 0. \quad (5.3)$$

On the left-hand side of equations (5.1)–(5.3) are all terms that depend on the subgrid-scale dynamics of the EDMF scheme.  $S_{\text{mp}}$  and  $Q_{\text{mp}}$  are moistening and heating rates due to microphysical processes,  $(\cdot)^*$  represents a deviation from the grid-mean  $\langle \cdot \rangle$ ,  $\Pi$  is the Exner function,  $\mathbf{u}_h$  is the horizontal velocity,  $\tilde{\mathbf{u}}_h$  is the reference horizontal velocity from the forcing climate model, and  $\Gamma_m = 4 \text{ day}^{-1}$  is the momentum relaxation frequency.

The EDMF-dependent terms must balance moistening and heating rates due to large-scale horizontal advection ( $\tilde{S}_{\text{hadv}}$ ,  $\tilde{Q}_{\text{hadv}}$ ), vertical eddy advection ( $\tilde{S}_{\text{veddy}}$ ,  $\tilde{Q}_{\text{veddy}}$ ), and radiation ( $\tilde{Q}_{\text{rad}}$ ); equations (5.1)–(5.3) provide a relationship between the grid-mean state, the SGS fluxes, and the microphysical tendencies. In LES, the actual equilibrium state depends on the resolved dynamics, the microphysics closure, and their interaction. In the SCM implementing the EDMF scheme, it depends on the SGS and microphysics closures. Therefore, this forcing framework provides a controlled testbed for the analysis of the EDMF scheme and its closures.

### Calibration of the EDMF scheme

To ensure robustness of the results and assess the degree of overfitting, we calibrate the free parameters within the extended EDMF scheme using LES data from a single model (HadGEM2-A) and climate (AMIP). This partitioning results in a training dataset of 76 unique configurations. We calibrate 16 free parameters related to entrainment processes, environmental mixing, and SGS pressure gradients; the closures where they appear are described in Cohen et al. (2020), Lopez-Gomez et al. (2020) and J. He et al. (2021). Only one parameter differs from those calibrated in Lopez-Gomez et al. (2022b). We no longer calibrate the updraft advection damping coefficient, which was shown to be non-identifiable in the aforementioned study. Instead, we introduce a turbulent Lewis number, defined as the ratio between turbulent heat and moisture diffusion (Chakraborty and Cant, 2009).

Calibration is performed using ensemble Kalman inversion, following Lopez-Gomez et al. (2022b). The loss function is constructed using  $\langle \theta_l \rangle$ ,  $\langle q_l \rangle$ , the liquid water specific humidity  $\langle q_l \rangle$ , the liquid water path (LWP), and the SGS fluxes of entropy  $\langle w^* s^* \rangle$  and moisture  $\langle w^* q_l^* \rangle$ , all averaged over the last 3 hours of simulation for both the LES and the SCM.

#### 5.4 Dynamics of low cloud regimes: LES vs. EDMF

We compare the EDMF and LES results using a set of scalar key performance indicators (KPIs), defined in Table 5.1. The vertically integrated SGS fluxes of moisture and entropy quantify the aggregate effect of turbulent and convective transport on the thermodynamic structure of the lower troposphere, and they are a crucial parameterization output (Siebesma, 1998; Arakawa, 2004). They are independent of the domain top because they vanish in the free troposphere. The cloud-top height  $z_{ct}$  and cloud thickness  $h_c$  are important indicators of cloud regime (Bretherton, 2015). Finally, LWP and the average cloud fraction  $f_{cl}$  have a strong influence on the radiative properties of the simulated cloud layer (Stephens, 1978; Pincus et al., 2003). We define  $f_{cl}$  as

$$f_{cl} = \frac{1}{h_c} \int_{z_{cb}}^{z_{ct}} f dz, \quad (5.4)$$

where  $z_{cb}$  is the cloud-base height. We choose this KPI instead of cloud cover or maximum cloud fraction, because it is independent of cloud overlap assumptions and less sensitive to model resolution.

Table 5.1: Symbol and description of key performance indicators (KPIs) of the extended EDMF scheme in the stratocumulus-to-cumulus transition. Each KPI is associated with a predictor that explains a large fraction of its variance across regimes. LHF stands for latent heat flux, and  $D$  is the decoupling parameter (Bretherton and Wyant, 1997). All KPIs computed over the lower 4 km of atmosphere.

KPI	Description	Predictor
$\langle w^* q_t^* \rangle_I$	Vertically integrated subgrid-scale moisture flux	LHF
$\langle w^* s^* \rangle_I$	Vertically integrated subgrid-scale entropy flux	LHF
$z_{ct}$	Cloud-top height	LHF
$h_c$	Cloud thickness	LHF
LWP	Liquid water path	$D$
$f_{cl}$	Average cloud fraction	$D$

To compare dynamical regime transitions between LES and the extended EDMF scheme, we aggregate simulations into bulk regimes defined by a set of predictors that explain a large fraction of the variance of each KPI across the transects. The integrated SGS fluxes, cloud-top height and thickness are shown in Figure 5.2 as a function of the surface latent heat flux (LHF) for the ESM-EDMF configuration. The EDMF scheme captures well the increase in turbulent transport and cloud height with LHF, compared to LES. The variance of the multi-model distribution is also similar in LES and the ESM-EDMF configuration.

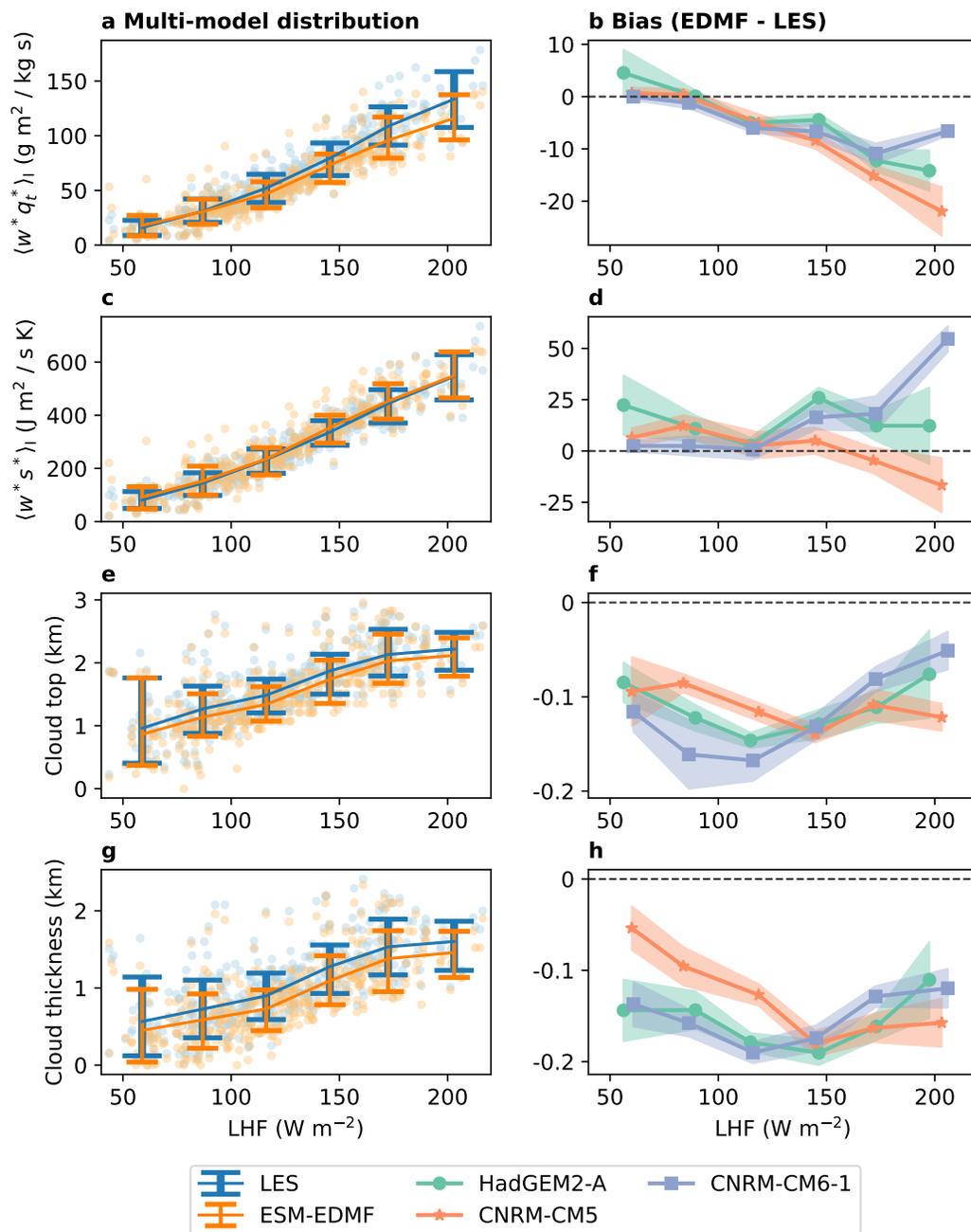


Figure 5.2: Bin-averaged values (left) and biases per forcing model (right) of the vertically integrated subgrid-scale fluxes of (a, b) moisture and (c, d) entropy, (e, f) cloud-top height, and (g, h) cloud thickness. Results for all simulations in the library are shown (dots), binned in  $30 \text{ W m}^{-2}$  intervals of surface latent heat flux (LHF). Error bars define the central 68% of the bin distribution, using the median-unbiased estimator of Hyndman and Fan (1996). Shading shows the bootstrapped standard deviation of the bin-averaged bias.

The ESM-EDMF configuration displays a negative bias in moisture flux, cloud height, and cloud thickness. The sign and magnitude of these biases is robust across models, as shown in Figure 5.2. Biases in moisture and entropy SGS fluxes increase with LHF, and can represent about 15% and 10% of the LES values, respectively. Biases in cloud height and thickness are relatively insensitive to the forcing and on the order of 100 m.

Biases in the moisture SGS flux are significantly reduced in the ISO-EDMF configuration, at the cost of a slight increase in the entropy flux bias, still under 10% of the LES value (Fig. 5.3a-d). Overall, the extended EDMF scheme shows a tendency to either underpredict moisture transport or overpredict entropy transport. Since turbulent transport is modulated by a learnable Lewis number in our model, we attribute this to potential deficiencies in convective transport. For instance, turbulent entrainment processes may result in different moisture and heat fluxes between updrafts and the environment, even though we model them as identical (Cohen et al., 2020).

As expected, biases in cloud top height and thickness are significantly reduced at the ISO-EDMF resolution (Fig. 5.3e-h). From all these KPIs, cloud top-height has been shown to be the easiest to match to LES in other single-column studies (M. Zhang et al., 2013). To verify the sensitivity of these KPIs to the SGS dynamics, we include results using a miscalibrated model in Section 5.8, where it is shown that the goodness of fit is not constrained by the large-scale forcing.

LWP and  $f_{cl}$  are plotted in Figure 5.4 as a function of the decoupling parameter (Bretherton and Wyant, 1997), diagnosed from LES as

$$D = \log \left( \frac{\text{LHF}}{\Delta L} \frac{h_c}{z_{ct}} \right), \quad (5.5)$$

where  $\Delta L$  is the net cloud radiative cooling, computed as in Schneider et al. (2019). Negative values of the decoupling parameter are associated with stratocumulus-topped boundary layers, and positive values are characteristic of cumulus clouds decoupled from the well-mixed boundary layer below. The multi-model distribution of these KPIs as diagnosed from the forcing climate models is also shown as a reference, but it should be noted that in this case the models experienced the fully coupled and transient dynamics of the climate system (Fig. 5.4a, c).

The ESM-EDMF configuration reproduces the transition of both LWP and  $f_{cl}$  as  $D$  changes sign. This is not the case for the ensemble mean response of the forcing

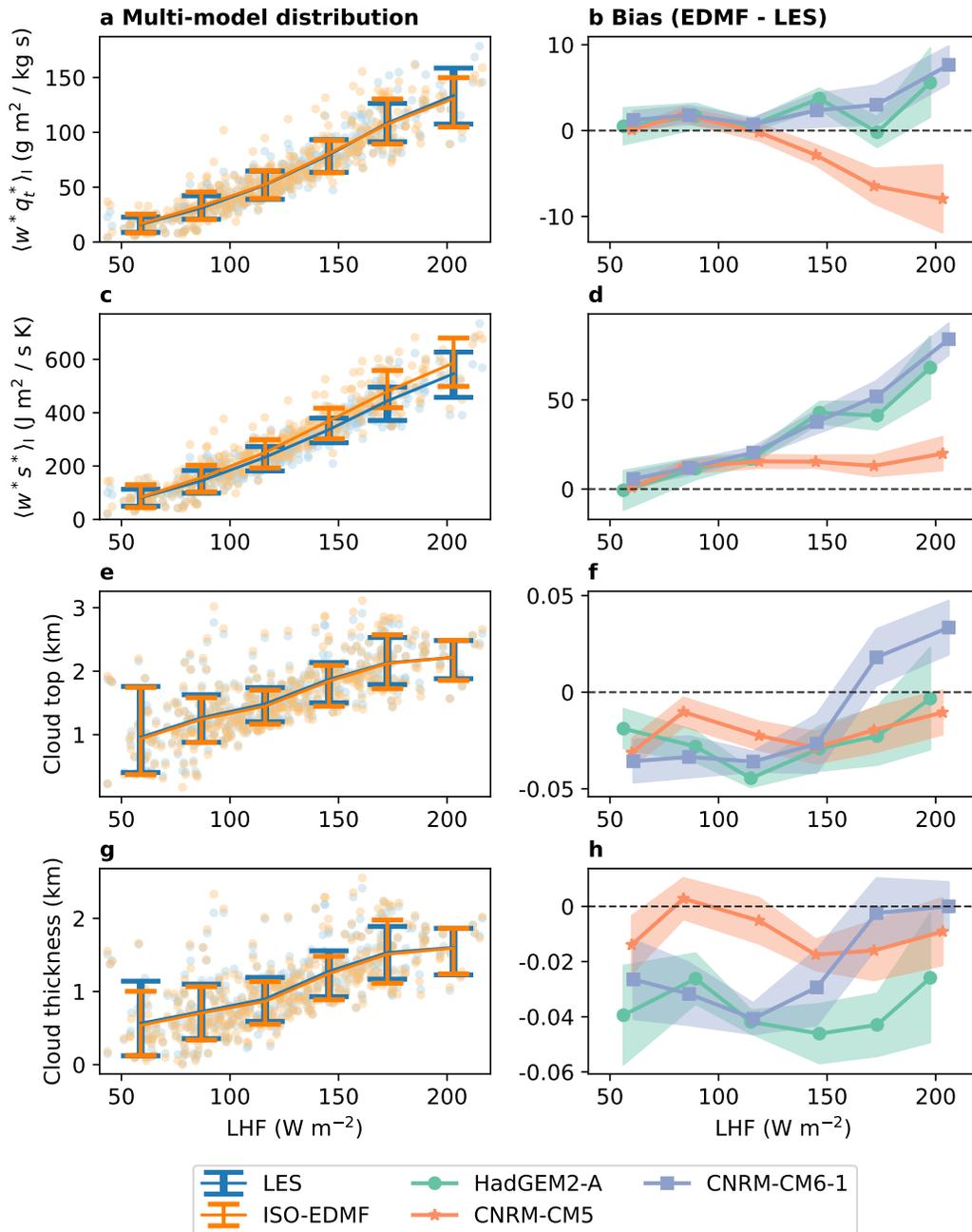


Figure 5.3: Same as Figure 5.2, but for the ISO-EDMF configuration.

models, which predicts a slight increase in both quantities as  $D$  increases from  $D < 0$  to  $D \sim 3$ . The ESM-EDMF configuration tends to forecast more extreme values of LWP and  $f_{cl}$  than LES in the coupled regime ( $D < 0$ ); this regime is also where the bias is highest (Figure 5.4). The sign of the LWP and  $f_{cl}$  biases in the stratocumulus regime varies across forcing models; the ESM-EDMF scheme overpredicts these quantities under CNRM-CM5 forcing, but underpredicts them

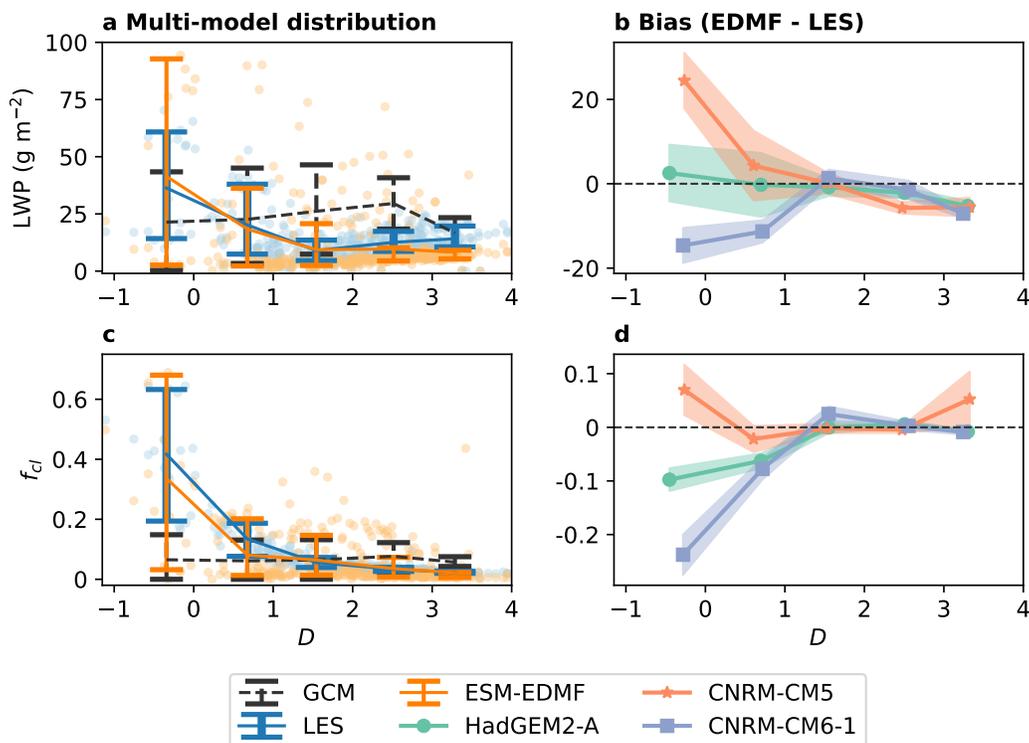


Figure 5.4: Bin-averaged values (left) and biases per forcing model (right) of (a, b) liquid water path and (c, d) average cloud fraction. Individual simulations (dots) are binned in decoupling parameter intervals with bounds  $[-1.15, 0, 1, 2, 3, 4]$ . Bin-averaged values from the forcing climate models are also shown (GCM). Error bars and shading as in Figure 5.2.

under CNRM-CM6-1. This is also the case in the ISO-EDMF configuration, which has a tendency to predict higher LWP and  $f_{cl}$  than ESM-EDMF (Fig. 5.5).

In the strongly decoupled regime, both the ESM-EDMF and ISO-EDMF configurations underpredict LWP across all models, although the bias is lower than in the coupled regime, as shown in Figure 5.6. Bias patterns are similar for  $f_{cl}$  in the ESM-EDMF configuration, with the largest biases in the stratocumulus regime. The biases of the extended EDMF scheme in both configurations are significantly lower than the biases of the forcing models (Fig. 5.6). Biases in cloud fraction are about 25% (of the total area) lower in the stratocumulus regime, and biases in LWP can be as much as  $20 \text{ g m}^{-2}$  lower. Overall, the increase in vertical resolution leads to a significant reduction in cloud fraction bias, but has little effect on the LWP bias of the scheme.

To investigate in more detail the response of the ESM-EDMF configuration, the vertical structure of the boundary layer at two HadGEM2-A AMIP4K cfSites is

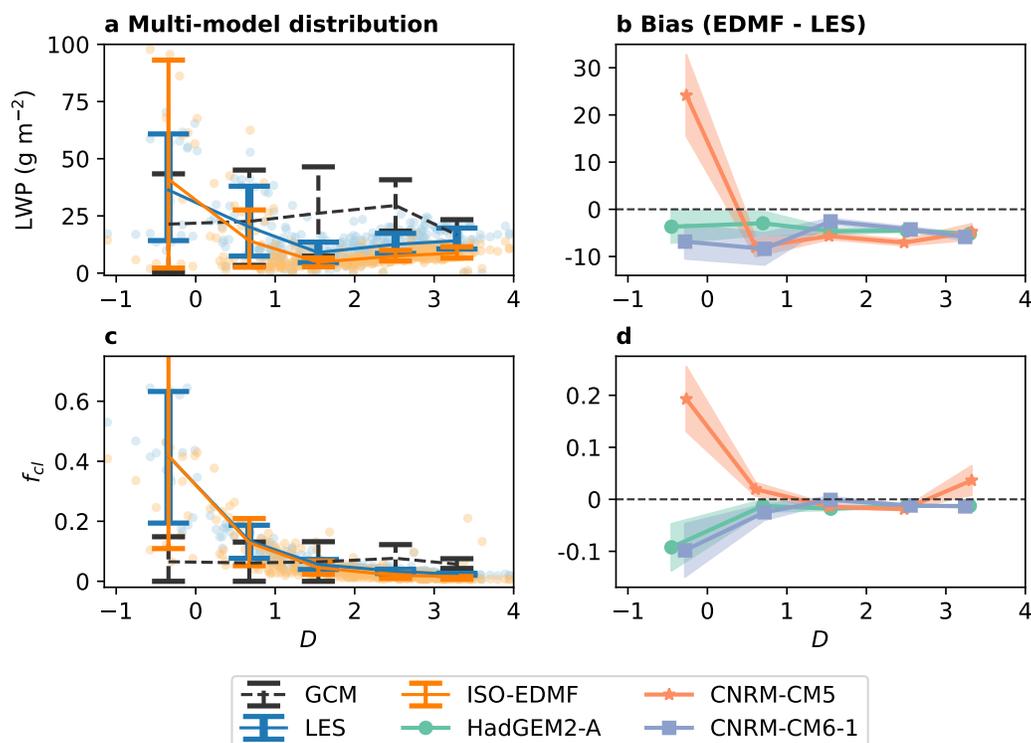


Figure 5.5: Same as Figure 5.4, but for the ISO-EDMF configuration.

shown in Figure 5.7. The time-averaged profiles from the forcing model are included as a reference, although we emphasize that these are the result of unsteady forcing conditions, unlike those used in LES and the SCM. The ESM-EDMF configuration captures the strong inversion around  $z \sim 1$  km at cfSite 17 off the coast of California (Fig. 5.7a, b), enabling it to sustain a stratocumulus cloud layer. The ability of the extended EDMF scheme to maintain sharp inversions at coarse resolution is mostly due to the mixing length parameterization introduced in Lopez-Gomez et al. (2020). The forcing model is unable to capture the sharp inversion, resulting in a thicker cloud layer with lower condensate concentration (Fig. 5.7c). In the lower half of the boundary layer, both the HadGEM2-A model and the ESM-EDMF configuration show a similar moist and cold bias.

At cfSite 23, biases in  $\theta_l$  and  $q_l$  are smaller in both models, although the EDMF scheme still maintains a sharper inversion. The cloud extent is well captured by the EDMF scheme, but the amount of condensate is underpredicted, especially near the anvil. This bias is opposite to the one found in the original HadGEM2-A model.

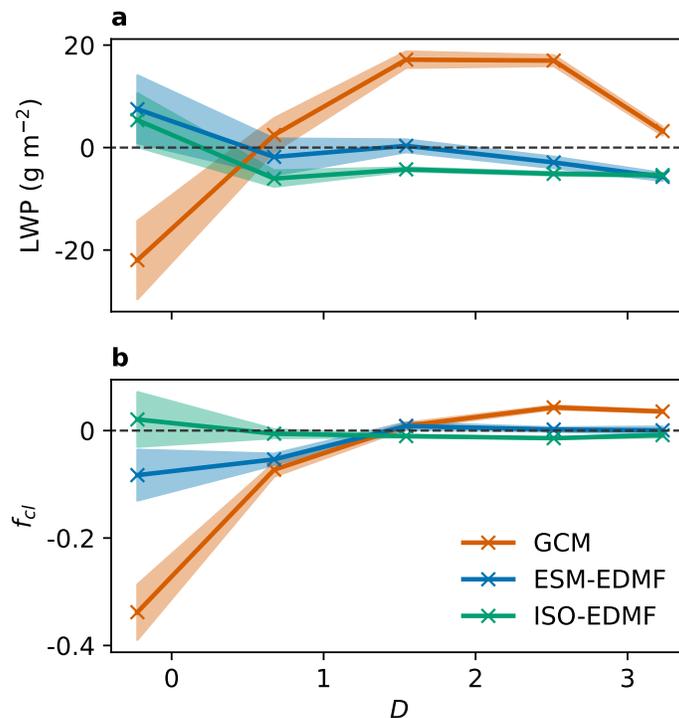


Figure 5.6: Bin-averaged biases in (a) liquid water path and (b) average cloud fraction, computed over the entire library. Results shown for the forcing climate models (GCM), and the extended EDMF scheme in ESM and ISO configurations. Shading as in Figure 5.4.

### 5.5 Response to a uniform 4 K warming: LES vs. EDMF

The response of a cloud-topped boundary layer to a climate perturbation is the result of a balance between competing mechanisms arising from changes in surface temperature, radiative cooling efficiency, wind speed, subsidence, and inversion strength (Bretherton et al., 2013; Bretherton, 2015). The 203 AMIP/AMIP4K simulation pairs in the LES library facilitate analyzing the response to perturbations that combine a realistic balance between these mechanisms, as simulated by the forcing climate models.

The changes in average cloud fraction and decoupling resulting from the 4 K climate perturbation for each LES AMIP/AMIP4K pair are shown in Figure 5.8a. The uniform increase in surface temperature leads to regime transitions from coupled stratocumulus to decoupled cumulus clouds in many of the AMIP stratocumulus sites. Regime transitions are characterized by sharp increments in the decoupling parameter as it changes sign, fostered by cloud thickening ( $\Delta(h_c/z_{ct}) > 0$ ) and reduced radiative cooling due to a diminishing cloud fraction.

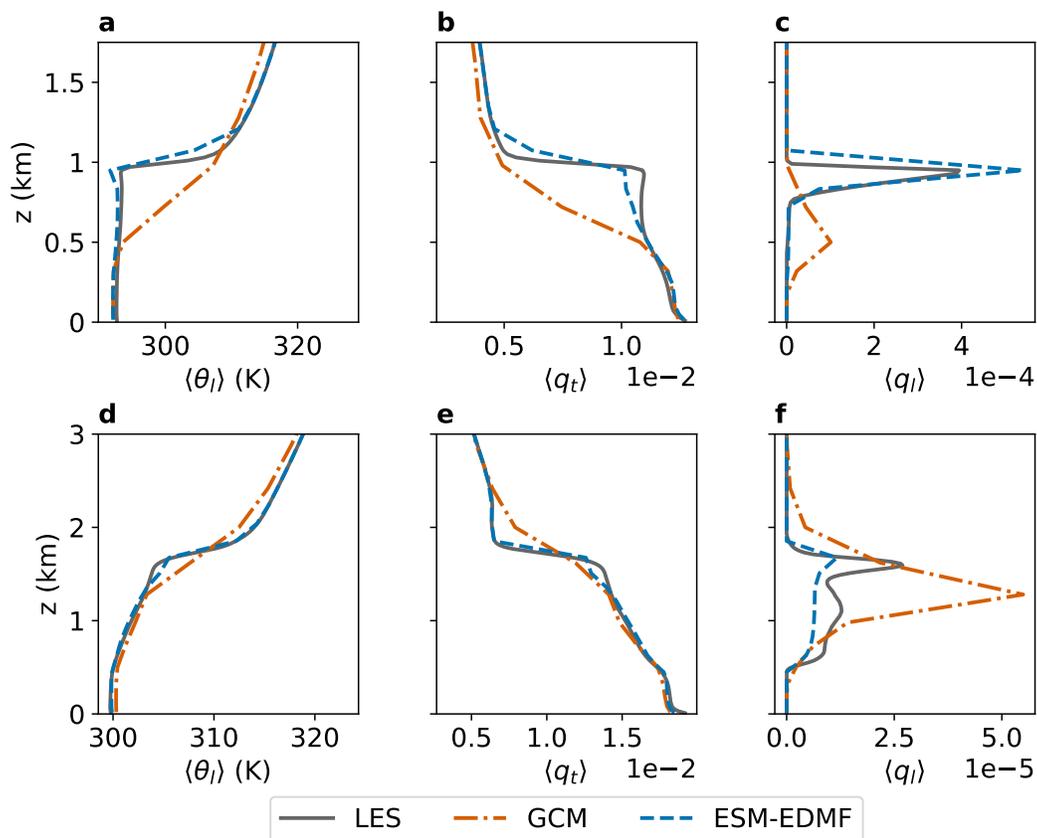


Figure 5.7: Vertical profiles of (a,d) liquid water potential temperature, (b,e) total water specific humidity, and (c,f) liquid water specific humidity from LES, the EDMF scheme at ESM resolution, and HadGEM2-A. Results shown for cfSites 17 (top) and 23 (bottom) of the AMIP4K experiment run with HadGEM2-A and with July forcing.

The EDMF scheme is able to reproduce this decoupling at similar values of  $D$ , given the true radiative tendencies used in our experimental setup (Fig. 5.8b, c). However, the resulting decoupled states in the ESM-EDMF configuration concentrate in two  $f_{cl}$  bands; a bistability that is not observed in LES. This bistability, whose signature is also visible in LWP, leads to errors in the prediction of the climate change response of the EDMF scheme at individual sites (Fig. 5.9e, f). Matching the response to climate perturbations has recently been reported to be more challenging than reducing errors in individual climates, using data-driven corrective tendencies trained on both current and warmer climates (Clark et al., 2022). The same seems to be the case for the EDMF scheme response.

The bistability and cloud fraction gap between the two regions disappears as the vertical resolution is increased; the phase space of the ISO-EDMF configuration is

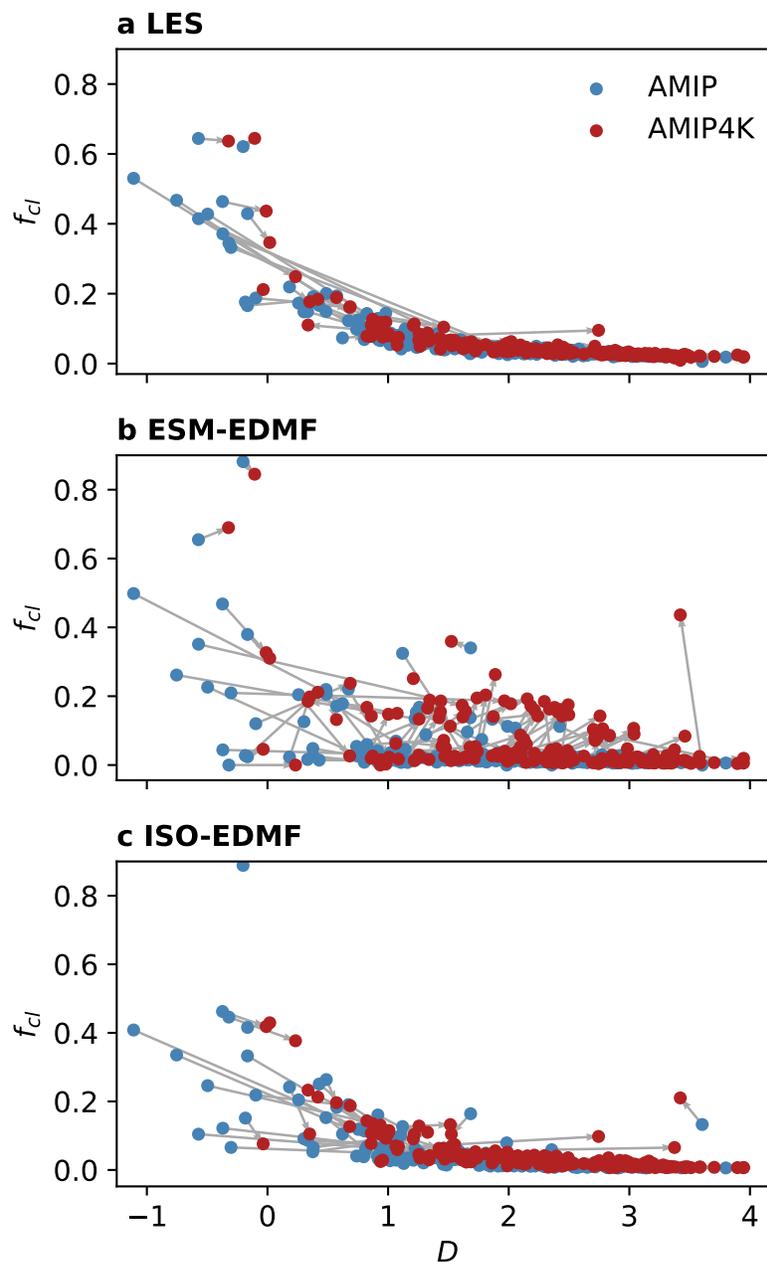


Figure 5.8: Cloud fraction and decoupling phase space of low cloud equilibrium states simulated by (a) LES, (b) ESM-EDMF and (c) ISO-EDMF. Arrows connect the AMIP and AMIP4K states of each cfSite and season. Simulations forced by all three climate models in the library are shown.

very similar to the LES phase space (Fig. 5.8c). This significantly increases the correlation between the LES and EDMF climate response, as shown in Figure 5.10. Even in the presence of bistability at coarse resolution, the extended EDMF scheme

captures the sign of changes in cloud fraction and LWP when these are large in LES, although it slightly underestimates them.

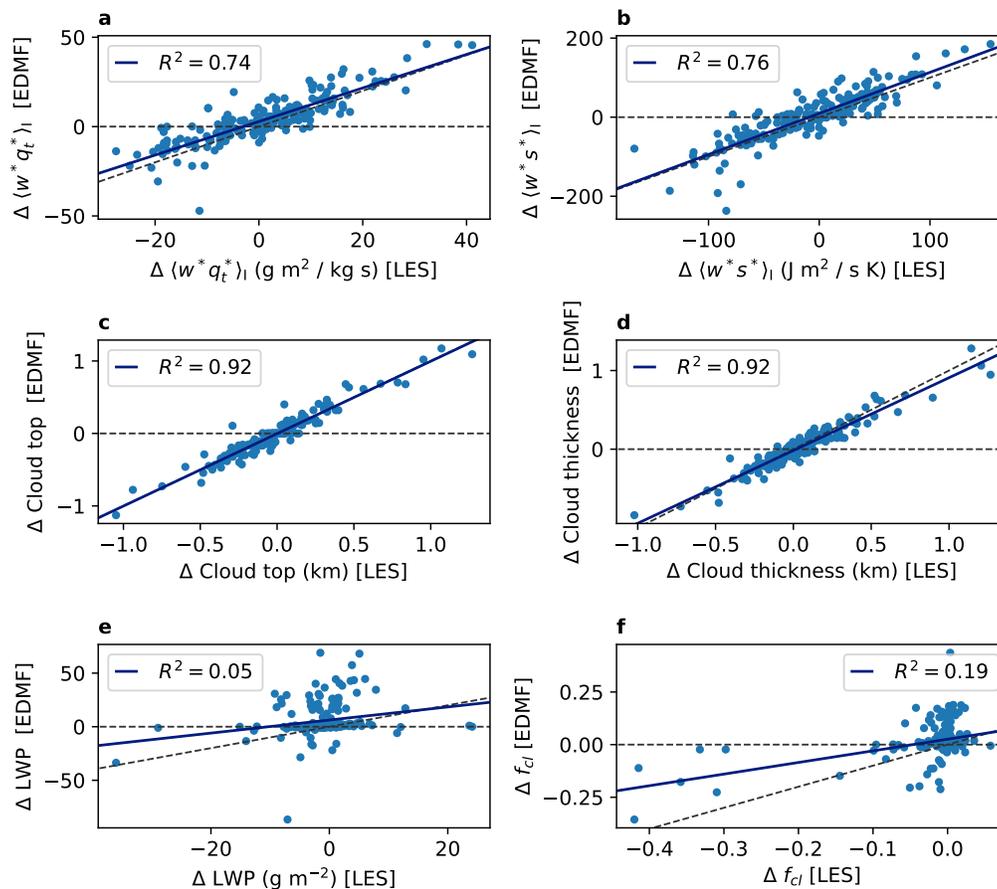


Figure 5.9: Scatter plot of AMIP4K-AMIP differences in KPI values as simulated by LES (x-axis), and ESM-EDMF (y-axis). The legend of each panel shows the coefficient of determination of the EDMF change in terms of the LES change; the linear regression model is shown as a blue line. Dashed lines show the zero-line and the identity.

The observed bistability in the shallow cumulus regime does not have a signature in the integrated vertical SGS fluxes, nor in the cloud top and thickness. As shown in Figure 5.9, the EDMF response to a climate perturbation measured by these KPIs tracks the LES response very well; the EDMF scheme captures over 70% of the LES integrated flux variance and about 90% of the LES variance in cloud thickness and top height. The magnitude of the forecast changes also agree on average. In the ISO-EDMF configuration, the correlations increase by about 10% in these KPIs and 30% in the case of LWP and  $f_{cl}$  (Fig. 5.10).

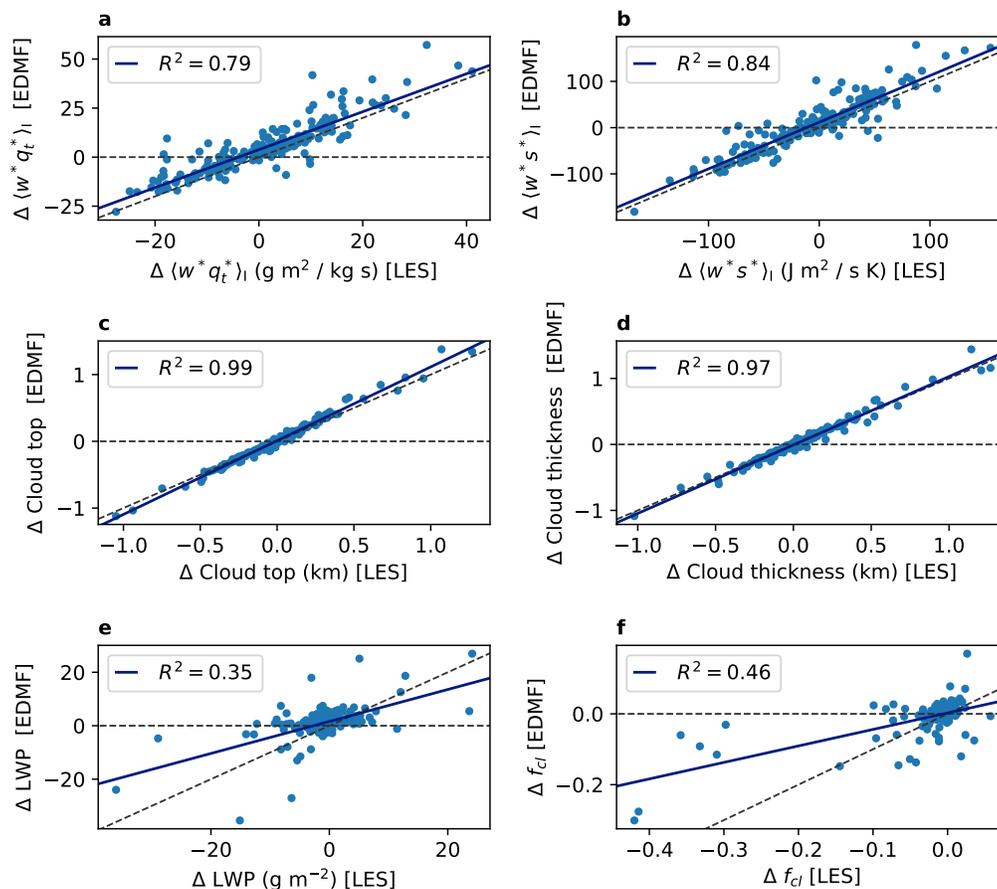


Figure 5.10: Same as Figure 5.9, but for the ISO-EDMF configuration.

## 5.6 Discussion

The dynamical equilibria reached by the extended EDMF scheme across a wide range of low-cloud regimes, spanning subtropical and tropical stratocumulus-to-cumulus transition transects in the eastern Pacific Ocean, are largely consistent with the resolved LES dynamics. This has been verified both at CMIP6-like resolution and at high isotropic resolution. Particularly, the vertically integrated transport of entropy and moisture, the cloud thickness, and the cloud top, all follow closely the corresponding LES values. The bulk dependence of liquid water path and cloud fraction on the decoupling parameter of Bretherton and Wyant (1997) predicted by the scheme is consistent with the LES response; this is far from being the case in the forcing climate models.

The analysis of the response to realistic climate perturbations shows that the extended EDMF scheme can robustly predict changes in subgrid-scale fluxes, cloud top height, and cloud thickness. On the other hand, the response of liquid water path and cloud

fraction when the EDMF is used at a CMIP6-like resolution only agree with the LES on average, and not necessarily at individual sites. This is due to a bistability of the decoupled shallow cumulus states when using the scheme at coarse resolution, and disappears as the resolution is increased. Since the LES climate change response of shallow cumuli is weak (Bretherton et al., 2013; Bretherton, 2015; Shen et al., 2022), this bistability suggests a higher climate change sensitivity of the parameterized response at a regional level. Understanding the climate implications of the larger sensitivity of the shallow cumulus clouds simulated by the extended EDMF scheme will require further analysis using global coupled simulations. Nevertheless, the scheme is able to predict regime transitions from coupled to decoupled states at both tested resolutions, and spans a similar phase space, as shown in Figure 5.8.

Biases in all variables are similar in order of magnitude across simulations forced by three different climate models; two of them covering the lower and higher ends of low-cloud feedback responses in the CMIP5 model ensemble, and another one from the newer CMIP6 generation. In this study, the extended EDMF scheme was calibrated using present climate data from one of the models, which highlights the robustness and generalization of the scheme to unseen conditions. This also serves as an example of the usefulness of sparse physics-based models in the presence of distributional shifts between training and testing sets; such strong generalization properties are difficult to match when large data-driven SGS models are used (Rasp et al., 2018; O’Gorman and Dwyer, 2018).

Overall, our analysis demonstrates that the closures within the extended EDMF scheme adequately model the unresolved dynamics of the low-troposphere in regimes spanning tropical and subtropical stratocumulus-to-cumulus transitions in the eastern Pacific. Only then can a turbulence and convection scheme achieve a dynamical balance similar to LES in our setup, as discussed in Section 5.3. An example of the response of a poorly calibrated scheme in Section 5.8 further shows that the exact quasi-steady equilibrium is largely unconstrained by the large-scale forcing and depends on the internal EDMF dynamics.

Some deficiencies of the current scheme are apparent in the results, related to the amount and horizontal distribution of condensate in the cloud layer, particularly at CMIP6-like resolution. These deficiencies are exacerbated when the climate change response at individual sites is considered. Cloud fraction and liquid water path largely determine the radiative properties and feedbacks of clouds (Stephens, 1978; Bretherton, 2015), so improvements of the scheme on this front are of paramount

importance.

Embedding more expressive data-driven closures within the EDMF scheme, as demonstrated in Lopez-Gomez et al. (2022b), is a potential way to reduce biases. The resulting hybrid EDMF scheme could then be trained across climates to further reduce errors in the climate change response. We leave this for future work here, since it would require expanding our LES library to include simulations across more than two climates, so that validation against unseen climates is still possible. Retaining unseen climates for validation is important because the two-way coupling between large-scale and parameterized dynamics can introduce similar distributional shifts in the large-scale forcing when the trained model is integrated in global simulations. Bias-correcting machine learning could also provide a path forward in the faithful representation of these properties at coarse resolution (Clark et al., 2022). Algorithmic advances that enable the use of localized higher vertical resolution in climate models, as proposed by H.-H. Lee et al. (2022), may provide another alternative path forward.

We reiterate that the scope of our conclusions is limited to the assessment of structural errors in the dynamical SGS and microphysical closures of the EDMF scheme. The effects of large-scale forcing feedbacks or radiative flux biases are not included in this analysis, which enabled a process-based assessment of the scheme but precludes the interpretation of our results as proof of the skill of the scheme in fully coupled global simulations. An analysis of global climate simulations using the extended EDMF scheme as the unified closure for turbulent and convective processes in the atmosphere is left for future work.

## 5.7 Appendix A: Stretched grid definition

The vertical discretization employed in the SCM follows the functional form

$$z(\xi; z_t, \Delta z_s, \Delta z_t) = z_t \cdot \Phi(\xi, h(\Delta z_s, \Delta z_t)) \quad (5.6)$$

$$\Phi(\xi, h) = -h \cdot \log[1 - (1 - e^{-1/h})\xi], \quad (5.7)$$

where  $\xi \in [1/n_z, 2/n_z, \dots, 1]$  is a nondimensional coordinate,  $n_z$  is the number of vertical levels, and  $\Delta z_s$  and  $\Delta z_t$  are the vertical resolution near the surface and near  $z_t$ , respectively. In equation (5.7),  $h$  is the linear stretching function

$$h(\xi) = h_s + \left(\xi - \frac{1}{n_z}\right) \left(\frac{h_t - h_s}{1 - 2/n_z}\right), \quad (5.8)$$

and  $h_s, h_t$  are defined implicitly as

$$\Phi\left(\frac{1}{n_z}, h_s\right) = \frac{\Delta z_s}{z_t}, \quad \Phi\left(1 - \frac{1}{n_z}, h_t\right) = 1 - \frac{\Delta z_t}{z_t}, \quad (5.9)$$

such that the discretization is uniquely defined by parameters  $\{z_t, n_z, \Delta z_s, \Delta z_t\}$ .

We first define a discretization  $\tilde{z}$  using equation (5.6),  $z_t = 45$  km, and values of  $\Delta z_s, \Delta z_t, n_z$  representative of current ESMs or NWP models, as specified in Section 5.3. The grid is then truncated to match the LES domain by preserving the number of degrees of freedom  $n_l$  covering the bottom  $z_l = 4$  km, such that the final discretization is defined as  $z = z(\xi; z_l, \Delta z_s, \Delta z_l)$  with

$$\Delta z_l = \tilde{z} \left( \frac{n_l}{n_z} \right) - \tilde{z} \left( \frac{n_l - 1}{n_z} \right). \quad (5.10)$$

## 5.8 Appendix B: Results for miscalibrated models

To demonstrate the sensitivity of the KPIs to the SGS closures within the turbulence and convection scheme, we analyzed the response of a version of the extended EDMF scheme where the entrainment closures of Cohen et al. (2020) were substituted by a poorly calibrated random feature model (Nelsen and Stuart, 2021). The resulting model produces dynamical equilibria with fog near the surface in most simulations, and dramatically underpredicts the turbulent and convective transport of entropy and moisture, as shown in Figure 5.11.

We also assessed the climate perturbation response of the extended EDMF scheme without substituting any of the SGS closures, but modifying parameters related to diffusion, entrainment and updraft pressure drag. Figure 5.12 shows that in this case the LES and EDMF response to the climate perturbation are weakly correlated, in terms of the vertically integrated fluxes, cloud top height and cloud thickness. This shows that the goodness-of-fit of the perturbation response for these KPIs is not constrained by the forcing.

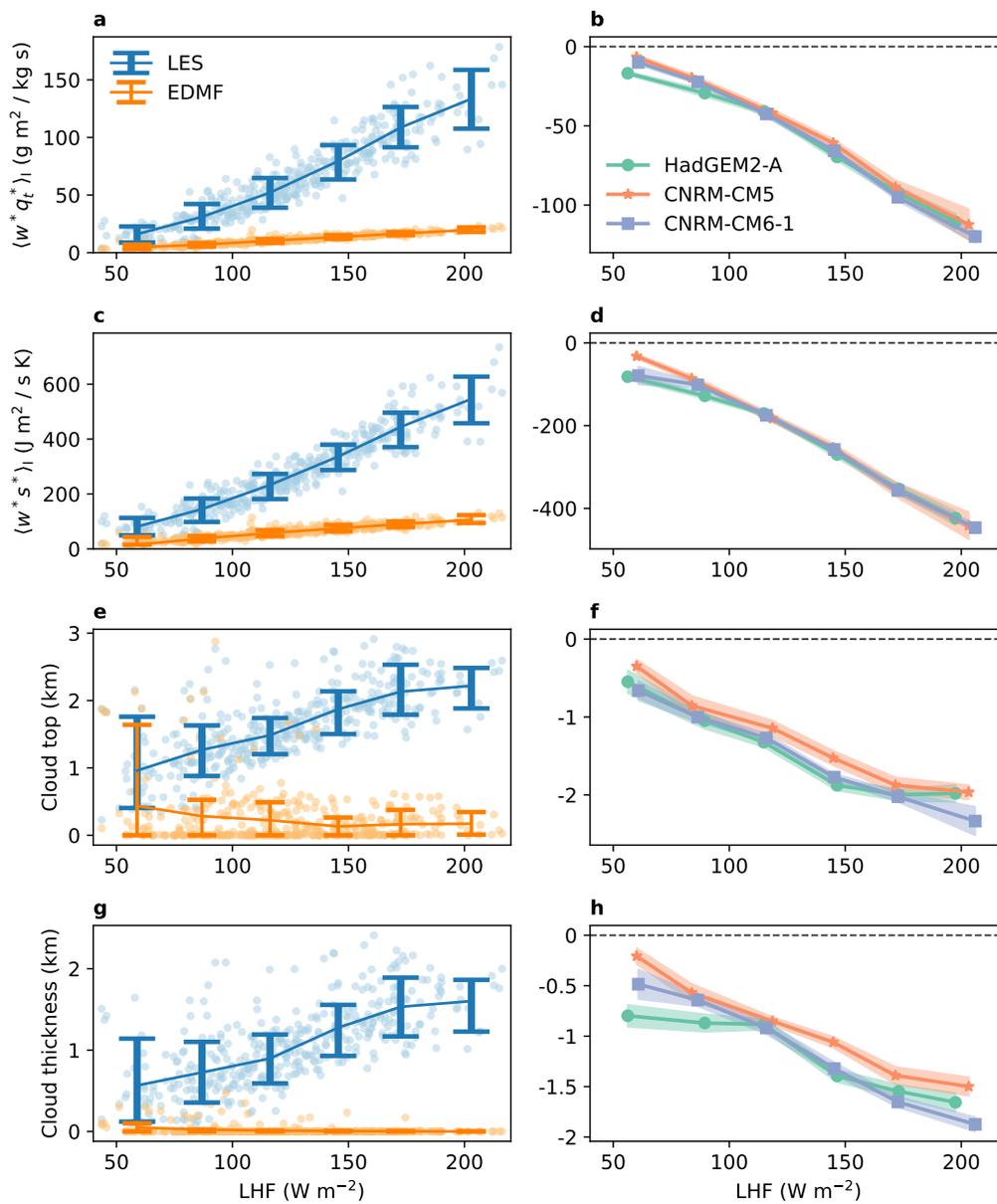


Figure 5.11: Same as Figure 5.2, but for a miscalibrated extended EDMF scheme with a random feature closure for entrainment fluxes, at ESM resolution.

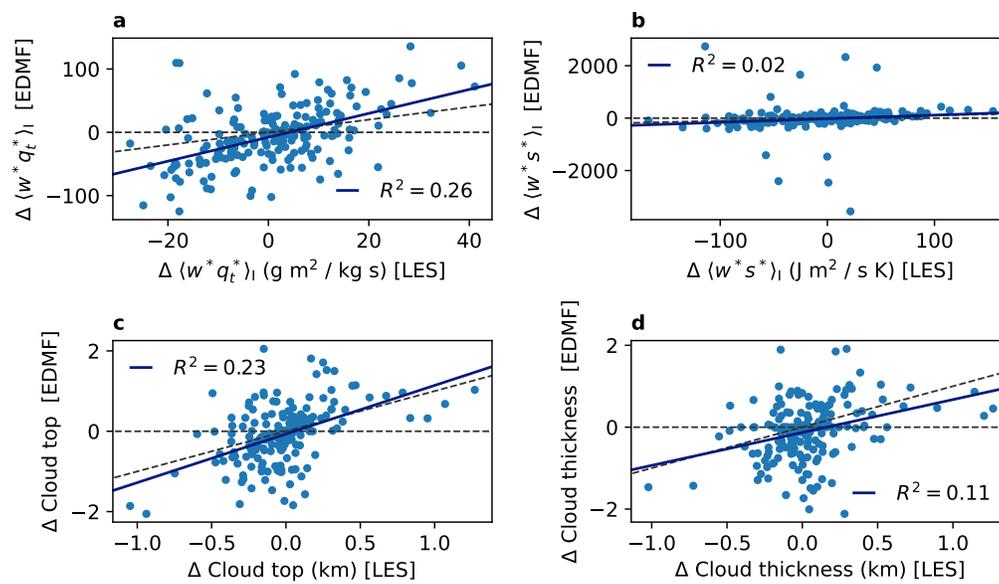


Figure 5.12: Same as Figure 5.9, but for a miscalibrated extended EDMF scheme at ESM resolution.

*Chapter 6*

## CONCLUSIONS

This dissertation addresses several challenges in the pursuit of an operational unified model of atmospheric turbulence and convection for climate prediction. The equations within the extended eddy-diffusivity mass-flux (EDMF) scheme are shown to be consistent with a systematic partition and filtering of the governing equations for an anelastic fluid. The assumptions necessary to yield the model are highlighted and compared with assumptions used in other turbulence and convection schemes, which provides valuable insight into the expected differences in the modeled dynamics.

The extended EDMF scheme requires closures for turbulent diffusion, microphysical processes, and mass and momentum exchange between convective regions and their environment. A generalized model of turbulent diffusion derived from a leading order balance between turbulent production and dissipation is shown to enable the accurate representation of dynamical regimes that continue haunting climate models, like stratocumulus-topped boundary layers and high-latitude stable boundary layers (Brient et al., 2019; H.-H. Lee et al., 2022; Konsta et al., 2022).

Together with closures for entrainment, detrainment, and momentum exchange between subdomains, the extended EDMF scheme provides a comprehensive representation of tropospheric mixing processes. The unified scheme is shown to remain skillful in atmospheric conditions ranging from the Arctic to the Amazon, all with a single set of  $\sim 20$  parameters. The increasing availability of data from high-resolution simulations and satellite observations represents a valuable opportunity to reduce parametric uncertainty even further and learn better closures from data.

These data, however, are typically not directly informative about the closures underlying parameterization schemes within climate models. This constraint either limits what can be learned from data to a differentiable correction of the model dynamics if supervised learning is employed (Brenowitz and Bretherton, 2018; Rasp et al., 2018; Bretherton et al., 2022), or requires the use of flexible optimization tools that allow indirect learning. This thesis demonstrates the feasibility of the second approach, by leveraging Kalman inversion processes (Iglesias et al., 2013; D. Z. Huang et al., 2022a; D. Z. Huang et al., 2022b) to learn from noisy and indirect data. This approach is shown to provide an efficient method to learn about the pa-

parameters within complex geophysical models, such as the extended EDMF scheme, even if the models are imperfect or if they become numerically unstable for certain parameter combinations.

Ensemble Kalman processes enable informing the extended EDMF scheme with data, and the rapid evaluation of different closure formulations in terms of their best possible online performance. The online performance of the extended EDMF scheme with the closures presented in Lopez-Gomez et al. (2020), Cohen et al. (2020) and J. He et al. (2021) is evaluated under realistic climatological conditions typical of the stratocumulus-to-cumulus transition in the eastern Pacific Ocean. The data-informed model, trained on present climate data from a single climate model, is shown to be skillful at forecasting low clouds and tropospheric mixing processes under previously unseen conditions, including large-scale conditions from a 4 K warmer climate. Some biases related to cloud condensate amount and horizontal distribution are still evident when the model is used at operational resolution. Indirect learning through ensemble Kalman processes can accelerate the discovery of better closures within the scheme to reduce these biases.

Although the skill shown by the scheme on a single-column setting far exceeds the performance of the forcing climate models<sup>1</sup> under similar large-scale conditions, the effective improvement in climate simulations of global models integrating the scheme remains to be shown. Global simulations are subject to additional constraints to prevent climate drift, such as matching the globally-averaged top-of-atmosphere outgoing longwave flux to the incoming shortwave flux (Mauritsen et al., 2012; Hourdin et al., 2017). This is often done by recalibrating parameters related to clouds, due to their strong modulation of the radiative properties of the atmosphere, which may potentially require the introduction of unphysical compensating mechanisms (Konsta et al., 2022). Other sources of bias, such as parameterized radiation (Pincus et al., 2003; Singer et al., 2021), will also need to be investigated in future work. Significant improvements in the skill of global weather models that have integrated other versions of the EDMF scheme to represent turbulence and shallow convection are empirical evidence that improved skill in single-column settings translates into better global performance (Köhler et al., 2011; Suselj et al., 2014; Hogan et al., 2014; Suselj et al., 2021). Taking the integration of the scheme one step further by representing all unresolved convective processes with a single unified model, including deep convection, is an exciting avenue of future research.

---

<sup>1</sup>Spanning Phases 5 and 6 of the Coupled Model Intercomparison Project.

## BIBLIOGRAPHY

- Abkar, Mahdi and Parviz Moin (2017). “Large-eddy simulation of thermally stratified atmospheric boundary-layer flow using a minimum dissipation model”. In: *Boundary-Layer Meteorology* 165, pp. 405–419. ISSN: 1573-1472. DOI: 10.1007/s10546-017-0288-4.
- Adcroft, Alistair et al. (2019). “The GFDL global ocean and sea ice model OM4.0: Model description and simulation features”. In: *Journal of Advances in Modeling Earth Systems* 11 (10), pp. 3167–3211. DOI: 10.1029/2019MS001726.
- Anderson, Jeffrey L (2012). “Localization and sampling error correction in ensemble Kalman filter data assimilation”. In: *Monthly Weather Review* 140, pp. 2359–2371. DOI: 10.1175/MWR-D-11-00013.1.
- Anderson, Jeffrey L and Stephen L Anderson (1999). “A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts”. In: *Monthly Weather Review* 127 (12), pp. 2741–2758. DOI: 10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2.
- Angevine, Wayne M, Hongli Jiang, and Thorsten Mauritsen (2010). “Performance of an eddy diffusivity–mass flux scheme for shallow cumulus boundary layers”. In: *Monthly Weather Review* 138, pp. 2895–2912.
- Arakawa, Akio (2004). “The cumulus parameterization problem: past, present, and future”. In: *Journal of Climate* 17 (13). doi: 10.1175/1520-0442(2004)0172.0.CO;2, pp. 2493–2525. ISSN: 0894-8755. DOI: 10.1175/1520-0442(2004)017<2493:RATCPP>2.0.CO;2.
- Arakawa, Akio and Wayne H Schubert (1974). “Interaction of a cumulus cloud ensemble with the large-scale environment, part {I}”. In: *Journal of the Atmospheric Sciences* 31 (3), pp. 674–701. DOI: 10.1175/1520-0469(1974)031<0674:IOACCE>2.0.CO;2.
- Arrhenius, Svante (1896). “XXXI. On the influence of carbonic acid in the air upon the temperature of the ground”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 41.251, pp. 237–276.
- Asai, Tomio and Akira Kasahara (1967). “A theoretical study of the compensating downward motions associated with cumulus clouds”. In: *Journal of Atmospheric Sciences* 24 (5), pp. 487–496. DOI: 10.1175/1520-0469(1967)024<0487:ATSOTC>2.0.CO;2.
- Augier, Pierre, Jean-Marc Chomaz, and Paul Billant (2012). “Spectral analysis of the transition to turbulence from a dipole in stratified fluid”. In: *Journal of Fluid Mechanics* 713, pp. 86–108. DOI: 10.1017/jfm.2012.437.
- Barthélémy, Sébastien et al. (2021). *Super-resolution data assimilation*. DOI: 10.48550/arxiv.2109.08017.

- Beare, Robert J. et al. (2006). “An intercomparison of large-eddy simulations of the stable boundary layer”. In: *Boundary-Layer Meteorology* 118, pp. 247–272. DOI: 10.1007/s10546-004-2820-6.
- Beljaars, Anton C M (1995). “The parametrization of surface fluxes in large-scale models under free convection”. In: *Quarterly Journal of the Royal Meteorological Society* 121, pp. 255–270. DOI: 10.1002/qj.49712152203.
- Betts, A K (1986). “A new convective adjustment scheme. Part I: Observational and theoretical basis”. In: *Quarterly Journal of the Royal Meteorological Society* 112, pp. 677–691. ISSN: 0035-9009. DOI: 10.1002/qj.49711247307.
- Betts, A K and M J Miller (1986). “A new convective adjustment scheme. Part II: Single column tests using GATE wave, BOMEX, ATEX and arctic air-mass data sets”. In: *Quarterly Journal of the Royal Meteorological Society* 112, pp. 693–709. ISSN: 0035-9009. DOI: 10.1002/qj.49711247308.
- Beucler, Tom, Timothy Cronin, and Kerry Emanuel (2018). “A linear response framework for radiative-convective instability”. In: *Journal of Advances in Modeling Earth Systems* 10.8, pp. 1924–1951. DOI: 10.1029/2018MS001280.
- Beucler, Tom et al. (2021). “Enforcing analytic constraints in neural networks emulating physical systems”. In: *Physical Review Letters* 126, p. 98302. DOI: 10.1103/PhysRevLett.126.098302.
- Billant, Paul and Jean-Marc Chomaz (2001). “Self-similarity of strongly stratified inviscid flows”. In: *Physics of Fluids* 13, pp. 1645–1651. DOI: 10.1063/1.1369125.
- Bjerknes, J (1938). “Saturated-adiabatic ascent of air through dry-adiabatically descending environment”. In: *Quarterly Journal of the Royal Meteorological Society* 64, pp. 325–330.
- Blackadar, Alfred K (1962). “The vertical distribution of wind and turbulent exchange in a neutral atmosphere”. In: *Journal of Geophysical Research* 67, pp. 3095–3102. DOI: 10.1029/JZ067i008p03095.
- Bocquet, M. and P. Sakov (2013). “Joint state and parameter estimation with an iterative ensemble Kalman smoother”. In: *Nonlinear Processes in Geophysics* 20, pp. 803–818. ISSN: 1607-7946. DOI: 10.5194/npg-20-803-2013.
- Boltzmann, Ludwig (1884). “Ableitung des Stefan’schen gesetzes betreffend die abhängigkeit der wärmestrahlung von der temperatur aus der electromagnetischen lichttheorie”. In: *Annalen der Physik und Chemie* 22, pp. 291–294.
- Bonan, G. B., M. Williams, R. A. Fisher, and K. W. Oleson (2014). “Modeling stomatal conductance in the earth system: linking leaf water-use efficiency and water transport along the soil–plant–atmosphere continuum”. In: *Geoscientific Model Development* 7 (5), pp. 2193–2222. DOI: 10.5194/gmd-7-2193-2014.

- Bony, Sandrine and Jean-Louis Dufresne (2005). “Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models”. In: *Geophysical Research Letters* 32.20. DOI: 10.1029/2005GL023851.
- Boucher, Olivier et al. (2013). “Clouds and aerosols. Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change”. In: *Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA*, pp. 571–657.
- Bougeault, Ph. (1981). “Modeling the trade-wind cumulus boundary layer. Part I: Testing the ensemble cloud relations against numerical data”. In: *Journal of the Atmospheric Sciences* 38, pp. 2414–2428. DOI: 10.1175/1520-0469(1981)038<2414:MTWCB>2.0.CO;2.
- Brajard, Julien et al. (2021). “Combining data assimilation and machine learning to infer unresolved scale parametrization”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 379, p. 20200086. ISSN: 1364503X. DOI: 10.1098/rsta.2020.0086.
- Brenowitz, Noah D and Christopher S Bretherton (2018). “Prognostic validation of a neural network unified physics parameterization”. In: *Geophysical Research Letters* 45, pp. 6289–6298. ISSN: 0094-8276. DOI: 10.1029/2018GL078510.
- Brenowitz, Noah D et al. (2020). “Interpreting and stabilizing machine-learning parametrizations of convection”. In: *Journal of the Atmospheric Sciences* 77, pp. 4357–4375. DOI: 10.1175/JAS-D-20-0082.1.
- Bretherton, Christopher S (2015). “Insights into low-latitude cloud feedbacks from high-resolution models”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 373 (2054), p. 20140415. DOI: 10.1098/rsta.2014.0415.
- Bretherton, Christopher S, Peter N Blossey, and Christopher R Jones (2013). “Mechanisms of marine low cloud sensitivity to idealized climate perturbations: A single-LES exploration extending the CGILS cases”. In: *Journal of Advances in Modeling Earth Systems* 5 (2), pp. 316–337. ISSN: 1942-2466. DOI: 10.1002/jame.20019.
- Bretherton, Christopher S and Matthew C Wyant (1997). “Moisture transport, lower-tropospheric stability, and decoupling of cloud-topped boundary layers”. In: *Journal of the Atmospheric Sciences* 54 (1), pp. 148–167. DOI: 10.1175/1520-0469(1997)054<0148:MTL TSA>2.0.CO;2.
- Bretherton, Christopher S, Brian Henn, Anna Kwa, Noah D Brenowitz, Oliver Watt-Meyer, Jeremy McGibbon, W Andre Perkins, Spencer K Clark, and Lucas Harris (2022). “Correcting coarse-grid weather and climate models by machine learning from global storm-resolving simulations”. In: *Journal of Advances in Modeling Earth Systems* 14, e2021MS002794. ISSN: 1942-2466. DOI: 10.1029/2021MS002794.

- Brient, Florent, Romain Roehrig, and Aurore Voltaire (2019). “Evaluating marine stratocumulus clouds in the CNRM-CM6-1 model using short-term hindcasts”. In: *Journal of Advances in Modeling Earth Systems* 11.1, pp. 127–148. DOI: 10.1029/2018MS001461.
- Brient, Florent and Tapio Schneider (2016). “Constraints on climate sensitivity from space-based measurements of low-cloud reflection”. In: *Journal of Climate* 29 (16), pp. 5821–5835. DOI: 10.1175/JCLI-D-15-0897.1.
- Bryan, George H and J Michael Fritsch (2002). “A benchmark simulation for moist nonhydrostatic numerical models”. In: *Monthly Weather Review* 130, pp. 2917–2928. DOI: 10.1175/1520-0493(2002)130<2917:ABSFMN>2.0.CO;2.
- Brynjarsdóttir, Jenný and Anthony O’Hagan (2014). “Learning about physical parameters: the importance of model discrepancy”. In: *Inverse Problems* 30, p. 114007. ISSN: 0266-5611. DOI: 10.1088/0266-5611/30/11/114007.
- Burgers, Gerrit, Peter Jan van Leeuwen, and Geir Evensen (1998). “Analysis scheme in the ensemble Kalman filter”. English. In: *Monthly Weather Review* 126, pp. 1719–1724. DOI: 10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.
- Businger, J A et al. (1971). “Flux-profile relationships in the atmospheric surface layer”. In: *Journal of the Atmospheric Sciences* 28, pp. 181–189. ISSN: 0022-4928. DOI: 10.1175/1520-0469(1971)028<0181:FPRITA>2.0.CO;2.
- Byun, Daewon W (1990). “On the analytical solutions of flux-profile relationships for the atmospheric surface layer”. In: *Journal of Applied Meteorology* 29, pp. 652–657. DOI: 10.1175/1520-0450(1990)029<0652:OTASOF>2.0.CO;2.
- Callendar, G S (1938). “The artificial production of carbon dioxide and its influence on temperature”. In: *Quarterly Journal of the Royal Meteorological Society* 64.275, pp. 223–240. DOI: 10.1002/qj.49706427503.
- Chada, Neil K, Andrew M Stuart, and Xin T Tong (2020). “Tikhonov regularization within ensemble Kalman inversion”. In: *SIAM Journal on Numerical Analysis* 58, pp. 1263–1294. ISSN: 0036-1429. DOI: 10.1137/19M1242331.
- Chakraborty, Nilanjan and R. S. Cant (2009). “Effects of Lewis number on scalar transport in turbulent premixed flames”. In: *Physics of Fluids* 21.3, p. 035110. DOI: 10.1063/1.3097007.
- Chen, Yan and Dean S Oliver (2012). “Ensemble randomized maximum likelihood method as an iterative ensemble smoother”. In: *Mathematical Geosciences* 44, pp. 1–26. ISSN: 1874-8953. DOI: 10.1007/s11004-011-9376-z.
- Chepfer, H et al. (2010). “The GCM-oriented CALIPSO cloud product (CALIPSO-GOCCP)”. In: *Journal of Geophysical Research: Atmospheres* 115 (D4). ISSN: 0148-0227. DOI: 10.1029/2009JD012251.

- Chure, Griffin, Rachel A Banks, Avi I Flamholz, Nicholas S Sarai, Mason Kamb, Ignacio Lopez-Gomez, Yinon Bar-On, Ron Milo, and Rob Phillips (2022). “Anthroponumbers.org: A quantitative database of human impacts on Planet Earth”. In: *Patterns* 3 (9), p. 100552. ISSN: 2666-3899. DOI: 10.1016/j.patter.2022.100552.
- Clark, Spencer K. et al. (2022). “Correcting a 200 km resolution climate model in multiple climates by machine learning from 25 km resolution simulations”. In: *Journal of Advances in Modeling Earth Systems* 14.9. e2022MS003219 2022MS003219, e2022MS003219. DOI: 10.1029/2022MS003219.
- Cleary, Emmet et al. (2021). “Calibrate, emulate, sample”. In: *Journal of Computational Physics* 424, p. 109716. ISSN: 0021-9991. DOI: 10.1016/j.jcp.2020.109716.
- Cohen, Yair, Ignacio Lopez-Gomez, Anna Jaruga, Jia He, Colleen M Kaul, and Tapio Schneider (2020). “Unified entrainment and detrainment closures for extended eddy-diffusivity mass-flux schemes”. In: *Journal of Advances in Modeling Earth Systems* 12, e2020MS002162. DOI: 10.1029/2020MS002162.
- Cohn, Stephen E (1997). “An introduction to estimation theory”. In: *Journal of the Meteorological Society of Japan. Ser. II* 75, pp. 257–288. DOI: 10.2151/jmsj1965.75.1B\_257.
- Cole, J. N. S. et al. (2005). “Global consequences of interactions between clouds and radiation at scales unresolved by global climate models”. In: *Geophysical Research Letters* 32.6. DOI: 10.1029/2004GL020945.
- Collins, William D et al. (2004). “Description of the NCAR community atmosphere model (CAM 3.0)”. In: *NCAR Technical Note NCAR/TN-464+ STR* 226.
- Cotter, S L et al. (2013). “MCMC methods for functions: Modifying old algorithms to make them faster”. In: *Statistical Science* 28, pp. 424–446. DOI: 10.1214/13-STS421.
- Cottrill, Andrew et al. (2013). “Seasonal forecasting in the Pacific using the coupled model POAMA-2”. In: *Weather and Forecasting* 28 (3), pp. 668–680. DOI: 10.1175/WAF-D-12-00072.1.
- Couvreur, F, F Hourdin, and C Rio (2010). “Resolved versus parametrized boundary-layer plumes. Part I: A parametrization-oriented conditional sampling in large-eddy simulations”. In: *Boundary-Layer Meteorology* 134.3, pp. 441–458. ISSN: 1573-1472. DOI: 10.1007/s10546-009-9456-5.
- Curry, J. A. et al. (1997). “Evolution of the cloudy boundary layer during the autumnal freezing of the Beaufort Sea”. In: *Journal of Geophysical Research: Atmospheres* 102, pp. 13851–13860. DOI: 10.1029/96JD03089.

- Dal Gesso, S and Roel A J Neggers (2018). “Can we use single-column models for understanding the boundary layer cloud-climate feedback?” In: *Journal of Advances in Modeling Earth Systems* 10, pp. 245–261. ISSN: 1942-2466. DOI: 10.1002/2017MS001113.
- Danabasoglu, G et al. (2020). “The Community Earth System Model Version 2 (CESM2)”. In: *Journal of Advances in Modeling Earth Systems* 12 (2), e2019MS001916. ISSN: 1942-2466. DOI: 10.1029/2019MS001916.
- Davini, Paolo et al. (2017). “Coherent structures in large-eddy simulations of a nonprecipitating stratocumulus-topped boundary layer”. In: *Journal of the Atmospheric Sciences* 74, pp. 4117–4137. DOI: 10.1175/JAS-D-17-0050.1.
- de Roode, Stephan R, Peter G Duynkerke, and A Pier Siebesma (2000). “Analogies between mass-flux and Reynolds-averaged equations”. In: *Journal of the Atmospheric Sciences* 57, pp. 1585–1598. DOI: 10.1175/1520-0469(2000)057<1585:ABMFAR>2.0.CO;2.
- de Roode, Stephan R et al. (2012). “Parameterization of the vertical velocity equation for shallow cumulus clouds”. In: *Monthly Weather Review* 140, pp. 2424–2436. ISSN: 0027-0644. DOI: 10.1175/MWR-D-11-00277.1.
- de Rooy, Wim C. et al. (2013). “Entrainment and detrainment in cumulus convection: An overview”. In: *Quarterly Journal of the Royal Meteorological Society* 139, pp. 1–19. DOI: 10.1002/qj.1959.
- Deardorff, James W. (1966). “The counter-gradient heat flux in the lower atmosphere and in the laboratory”. In: *Journal of the Atmospheric Sciences* 23, pp. 503–506. DOI: 10.1175/1520-0469(1966)023<0503:TCGHFI>2.0.CO;2.
- (1976). “Usefulness of liquid-water potential temperature in a shallow-cloud model”. In: *Journal of Applied Meteorology* 15, pp. 98–102. DOI: 10.1175/1520-0450(1976)015<0098:UOLWPT>2.0.CO;2.
- Dunbar, Oliver R A et al. (2021). “Calibration and uncertainty quantification of convective parameters in an idealized GCM”. In: *Journal of Advances in Modeling Earth Systems* 13, e2020MS002454. ISSN: 1942-2466. DOI: 10.1029/2020MS002454.
- Dunbar, Oliver R A et al. (2022). “Ensemble-based experimental design for targeting data acquisition to inform climate models”. In: *Journal of Advances in Modeling Earth Systems* 14 (9), e2022MS002997. ISSN: 1942-2466. DOI: 10.1029/2022MS002997.
- Emanuel, Kerry A (1991). “A scheme for representing cumulus convection in large-scale models”. In: *Journal of Atmospheric Sciences* 48 (21), pp. 2313–2329. DOI: 10.1175/1520-0469(1991)048<2313:ASFRCC>2.0.CO;2.

- Emanuel, Kerry A, J David Neelin, and Christopher S Bretherton (July 1994). “On large-scale circulations in convecting atmospheres”. In: *Quarterly Journal of the Royal Meteorological Society* 120 (519). <https://doi.org/10.1002/qj.49712051902>, pp. 1111–1143. ISSN: 0035-9009. DOI: 10.1002/qj.49712051902.
- Emanuel, Kerry A and David J Raymond (1993). “The representation of cumulus convection in numerical models”. In: *Meteorological Monographs* 24, pp. 1–246. DOI: 10.1175/0065-9401-24.46.1.
- Emerick, Alexandre A and Albert C Reynolds (2013). “Ensemble smoother with multiple data assimilation”. In: *Computers & Geosciences* 55, pp. 3–15. ISSN: 0098-3004. DOI: 10.1016/j.cageo.2012.03.011.
- Evensen, Geir (1994). “Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics”. In: *Journal of Geophysical Research: Oceans* 99, pp. 10143–10162. ISSN: 0148-0227. DOI: 10.1029/94JC00572.
- Eyring, Veronika et al. (2016). “Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization”. In: *Geoscientific Model Development* 9 (5), pp. 1937–1958. ISSN: 1991-9603. DOI: 10.5194/gmd-9-1937-2016.
- Favre, AJ (1965). *The equations of compressible turbulent gases*. Tech. rep. Université d’Aix-Marseille, Institut de Mécanique Statistique de la Turbulence.
- Fourier, Jean-Baptiste Joseph (1827). “Mémoire sur les températures du globe terrestre et des espaces planétaires”. In: *Mémoires de l’Académie Royale des Sciences de l’Institut de France* 7, pp. 570–604.
- Fureby, C and G Tabor (1997). “Mathematical and physical constraints on large-eddy simulations”. In: *Theoretical and Computational Fluid Dynamics* 9.2, pp. 85–102. DOI: 10.1007/s001620050034.
- Garbuno-Inigo, Alfredo et al. (2020). “Interacting Langevin diffusions: Gradient structure and ensemble Kalman sampler”. In: *SIAM Journal on Applied Dynamical Systems* 19, pp. 412–441. DOI: 10.1137/19M1251655.
- Gentine, P et al. (2018). “Could machine learning break the convection parameterization deadlock?” In: *Geophysical Research Letters* 45 (11), pp. 5742–5751. ISSN: 0094-8276. DOI: 10.1029/2018GL078202.
- Gettelman, A and Steven C Sherwood (2016). “Processes responsible for cloud feedback”. In: *Current Climate Change Reports* 2 (4), pp. 179–189. ISSN: 2198-6061. DOI: 10.1007/s40641-016-0052-8.
- Giordani, Hervé, Romain Bourdallé-Badie, and Gurvan Madec (2020). “An eddy-diffusivity mass-flux parameterization for modeling oceanic convection”. In: *Journal of Advances in Modeling Earth Systems* 12 (9), e2020MS002078. ISSN: 1942-2466. DOI: 10.1029/2020MS002078.

- Goddard, Lisa, Walter Baethgen, Ben Kirtman, and Gerald Meehl (2009). “The urgent need for improved climate models and predictions”. In: *Eos, Transactions American Geophysical Union* 90 (39), p. 343. ISSN: 0096-3941. DOI: 10.1029/2009E0390004.
- Grachev, Andrey A et al. (2013). “The critical Richardson number and limits of applicability of local similarity theory in the stable boundary layer”. In: *Boundary-Layer Meteorology* 147, pp. 51–82. DOI: 10.1007/s10546-012-9771-0.
- Gregory, David (2001). “Estimation of entrainment rate in simple models of convective clouds”. In: *Quarterly Journal of the Royal Meteorological Society* 127, pp. 53–72.
- Gregory, David and P R Rowntree (1990). “A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure”. In: *Monthly Weather Review* 118 (7), pp. 1483–1506. DOI: 10.1175/1520-0493(1990)118<1483:AMFCSW>2.0.CO;2.
- Grell, Georg A. and Dezső Dévényi (2002). “A generalized approach to parameterizing convection combining ensemble and data assimilation techniques”. In: *Geophysical Research Letters* 29.14, pp. 38-1-38-4. DOI: 10.1029/2002GL015311.
- Grenier, Hervé and Christopher S Bretherton (2001). “A moist PBL parameterization for large-scale models and its application to subtropical cloud-topped marine boundary layers”. In: *Monthly Weather Review* 129.3, pp. 357–377. DOI: 10.1175/1520-0493(2001)129<0357:AMPPFL>2.0.CO;2.
- Grisogono, B. (2010). “Generalizing ‘z-less’ mixing length for stable boundary layers”. In: *Quarterly Journal of the Royal Meteorological Society* 136, pp. 213–221. DOI: 10.1002/qj.529.
- Guillaumin, Arthur P and Laure Zanna (2021). “Stochastic-deep learning parameterization of ocean momentum forcing”. In: *Journal of Advances in Modeling Earth Systems* 13, e2021MS002534. ISSN: 1942-2466. DOI: 10.1029/2021MS002534.
- Haarsma, Reindert J. et al. (2016). “High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6”. In: *Geoscientific Model Development* 9 (11), pp. 4185–4208. ISSN: 1991-9603. DOI: 10.5194/gmd-9-4185-2016.
- Han, Jongil and Christopher S Bretherton (2019). “TKE-based moist eddy-diffusivity mass-Flux (EDMF) parameterization for vertical turbulent mixing”. In: *Weather and Forecasting*. DOI: 10.1175/WAF-D-18-0146.1.
- Han, Jongil et al. (2017). “Updates in the NCEP GFS cumulus convection schemes with scale and aerosol awareness”. In: *Weather and Forecasting* 32 (5), pp. 2005–2017. DOI: 10.1175/WAF-D-17-0046.1.
- Hansen, Per Christian (1990). “Truncated singular value decomposition solutions to discrete ill-posed problems with ill-determined numerical rank”. In: *SIAM Journal on Scientific and Statistical Computing* 11, pp. 503–518. ISSN: 0196-5204. DOI: 10.1137/0911028.

- Hansen, Per Christian (1998). *Rank-deficient and discrete ill-posed problems*. Society for Industrial and Applied Mathematics. ISBN: 978-0-89871-403-6. DOI: 10.1137/1.9780898719697.
- Hartmann, Dennis L, Maureen E Ockert-Bell, and Marc L Michelsen (1992). “The effect of cloud type on Earth’s energy balance: global analysis”. In: *Journal of Climate* 5 (11), pp. 1281–1304. DOI: 10.1175/1520-0442(1992)005<1281:TEOCTO>2.0.CO;2.
- He, Jia, Yair Cohen, Ignacio Lopez-Gomez, Anna Jaruga, and Tapio Schneider (2021). “An improved perturbation pressure closure for eddy-diffusivity mass-flux schemes”. In: *Earth and Space Science Open Archive*, p. 37. DOI: 10.1002/essoar.10505084.2.
- He, Y., N. A. McFarlane, and A. H. Monahan (2019). “A new TKE-based parameterization of atmospheric turbulence in the Canadian global and regional climate models”. In: *Journal of Advances in Modeling Earth Systems* 11.
- Hogan, Timothy F et al. (2014). “The Navy global environmental model”. In: *Oceanography* 27 (3), pp. 116–125. ISSN: 10428275, 2377617X. URL: <http://www.jstor.org/stable/24862194>.
- Hope, Chris (Nov. 2015). “The \$10 trillion value of better information about the transient climate response”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 373 (2054), p. 20140429. DOI: 10.1098/rsta.2014.0429.
- Houghton, Henry G and Harrison E Cramer (1951). “A theory of entrainment in convective currents”. In: *Journal of Atmospheric Sciences* 8 (2), pp. 95–102. DOI: 10.1175/1520-0469(1951)008<0095:ATOEIC>2.0.CO;2.
- Hourdin, Frédéric et al. (2017). “The art and science of climate model tuning”. In: *Bulletin of the American Meteorological Society* 98 (3), pp. 589–602. DOI: 10.1175/BAMS-D-15-00135.1.
- Houtekamer, P L and Herschel L Mitchell (1998). “Data assimilation using an ensemble Kalman filter technique”. English. In: *Monthly Weather Review* 126, pp. 796–811. DOI: 10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.
- (2001). “A sequential ensemble Kalman filter for atmospheric data assimilation”. In: *Monthly Weather Review* 129, pp. 123–137. DOI: 10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2.
- Huang, Daniel Zhengyu, Tapio Schneider, and Andrew M Stuart (2022a). “Iterated Kalman methodology for inverse problems”. In: *Journal of Computational Physics*, p. 111262. ISSN: 0021-9991. DOI: 10.1016/j.jcp.2022.111262.
- Huang, Daniel Zhengyu, Jiaoyang Huang, Sebastian Reich, and Andrew M. Stuart (2022b). *Efficient derivative-free Bayesian inference for large-scale inverse problems*. DOI: 10.48550/arxiv.2204.04386.

- Hwong, Y.-L., S C Sherwood, and D Fuchs (Apr. 2022). “Can we use 1D models to predict 3D model response to forcing in an idealized framework?” In: *Journal of Advances in Modeling Earth Systems* 14.4, e2021MS002785. ISSN: 1942-2466. DOI: <https://doi.org/10.1029/2021MS002785>.
- Hyndman, Rob J and Yanan Fan (Nov. 1996). “Sample quantiles in statistical packages”. In: *The American Statistician* 50 (4), pp. 361–365. ISSN: 0003-1305. DOI: [10.1080/00031305.1996.10473566](https://doi.org/10.1080/00031305.1996.10473566).
- Iglesias, Marco A (2016). “A regularizing iterative ensemble Kalman method for PDE-constrained inverse problems”. In: *Inverse Problems* 32, p. 025002. ISSN: 0266-5611. DOI: [10.1088/0266-5611/32/2/025002](https://doi.org/10.1088/0266-5611/32/2/025002).
- Iglesias, Marco A, Kody J H Law, and Andrew M Stuart (2013). “Ensemble Kalman methods for inverse problems”. In: *Inverse Problems* 29 (4), p. 045001. ISSN: 0266-5611. DOI: [10.1088/0266-5611/29/4/045001](https://doi.org/10.1088/0266-5611/29/4/045001).
- Ingram, William and Andrew C Bushell (2021). “Sensitivity of climate feedbacks to vertical resolution in a general circulation model”. In: *Geophysical Research Letters* 48 (12), e2020GL092268. ISSN: 0094-8276. DOI: [10.1029/2020GL092268](https://doi.org/10.1029/2020GL092268).
- Janjić, Zaviša I (1994). “The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes”. In: *Monthly Weather Review* 122 (5), pp. 927–945. DOI: [10.1175/1520-0493\(1994\)122<0927:TSMECM>2.0.CO;2](https://doi.org/10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2).
- Kain, John S and J Michael Fritsch (1990). “A one-dimensional entraining/detraining plume model and its application in convective parameterization”. In: *Journal of Atmospheric Sciences* 47 (23), pp. 2784–2802. DOI: [10.1175/1520-0469\(1990\)047<2784:AODEPM>2.0.CO;2](https://doi.org/10.1175/1520-0469(1990)047<2784:AODEPM>2.0.CO;2).
- Kaipio, Jari and Erkki Somersalo (2006). *Statistical and computational inverse problems*. Vol. 160. Springer Science & Business Media.
- Kajikawa, Yoshiyuki et al. (2016). “Resolution dependence of deep convections in a global simulation from over 10-kilometer to sub-kilometer grid spacing”. In: *Progress in Earth and Planetary Science* 3, p. 16. DOI: [10.1186/s40645-016-0094-5](https://doi.org/10.1186/s40645-016-0094-5).
- Katul, Gabriel G. et al. (2013). “Mean scalar concentration profile in a sheared and thermally stratified atmospheric surface layer”. In: *Physical Review E* 87, p. 023004. DOI: [10.1103/PhysRevE.87.023004](https://doi.org/10.1103/PhysRevE.87.023004).
- Kennedy, Marc and Anthony O’Hagan (2001). “Bayesian calibration of computer models”. In: *Journal of the Royal Statistical Society Series B* 63, pp. 425–464. DOI: [10.1111/1467-9868.00294](https://doi.org/10.1111/1467-9868.00294).
- Keskar, Nitish Shirish et al. (2016). *On large-batch training for deep learning: Generalization gap and sharp minima*. DOI: [10.48550/arXiv.1609.04836](https://doi.org/10.48550/arXiv.1609.04836).

- Kessler, E (1995). “On the continuity and distribution of water substance in atmospheric circulations”. In: *Atmospheric Research* 38 (1), pp. 109–145. ISSN: 0169-8095. DOI: 10.1016/0169-8095(94)00090-Z.
- Kim, Dongchul, Chien Wang, Annica M L Ekman, Mary C Barth, and Phil J Rasch (2008). “Distribution and direct radiative forcing of carbonaceous and sulfate aerosols in an interactive size-resolving aerosol–climate model”. In: *Journal of Geophysical Research: Atmospheres* 113 (D16). DOI: 10.1029/2007JD009756.
- Kirchhoff, G (1860). “Über das verhältniss zwischen dem emissionsvermögen und dem absorptionsvermögen der körper für wärme und licht”. In: *Annalen der Physik* 185.2, pp. 275–301. DOI: 10.1002/andp.18601850205.
- Klinger, Carolin et al. (2017). “Effects of 3-D thermal radiation on the development of a shallow cumulus cloud field”. In: *Atmospheric Chemistry and Physics* 17 (8), pp. 5477–5500. ISSN: 1680-7324. DOI: 10.5194/acp-17-5477-2017.
- Köhler, M (2005). “Improved prediction of boundary layer clouds”. In: (104), pp. 18–22. DOI: 10.21957/812mkwz370. URL: <https://www.ecmwf.int/node/18106>.
- Köhler, M, M Ahlgrimm, and A Beljaars (2011). “Unified treatment of dry convective and stratocumulus-topped boundary layers in the ECMWF model”. In: *Quarterly Journal of the Royal Meteorological Society* 137 (654), pp. 43–57. ISSN: 0035-9009. DOI: doi.org/10.1002/qj.713.
- Konsta, Dimitra et al. (2022). “Low-level marine tropical clouds in six CMIP6 models are too few, too bright but also too compact and too homogeneous”. In: *Geophysical Research Letters* 49 (11), e2021GL097593. ISSN: 0094-8276. DOI: 10.1029/2021GL097593.
- Kovachki, Nikola B and Andrew M Stuart (2019). “Ensemble Kalman inversion: a derivative-free technique for machine learning tasks”. In: *Inverse Problems* 35, p. 095005. ISSN: 0266-5611. DOI: 10.1088/1361-6420/ab1c3a.
- Kuo, H L (1974). “Further studies of the parameterization of the influence of cumulus convection on large-scale flow”. In: *Journal of Atmospheric Sciences* 31 (5), pp. 1232–1240. DOI: 10.1175/1520-0469(1974)031<1232:FSOTPO>2.0.CO;2.
- Kurowski, Marcin J and Joao Teixeira (2018). “A scale-adaptive turbulent kinetic energy closure for the dry convective boundary layer”. In: *Journal of the Atmospheric Sciences* 75, pp. 675–690. DOI: 10.1175/JAS-D-16-0296.1.
- Kurowski, Marcin J et al. (2019). “Towards unifying the planetary boundary layer and shallow convection in CAM5 with the eddy-diffusivity/mass-flux approach”. In: *Atmosphere* 10. ISSN: 2073-4433. DOI: 10.3390/atmos10090484.

- Lappen, Cara-Lyn and David A Randall (2001). “Toward a unified parameterization of the boundary layer and moist convection. Part {I}: A new type of mass-flux model”. In: *Journal of the Atmospheric Sciences* 58 (15), pp. 2021–2036. DOI: 10.1175/1520-0469(2001)058<2021:TAUPOT>2.0.CO;2.
- Lee, Hsiang-He, Peter Bogenschutz, and Takanobu Yamaguchi (2022). “Resolving away stratocumulus biases in modern global climate models”. In: *Geophysical Research Letters* 49 (18), e2022GL099422. ISSN: 0094-8276. DOI: 10.1029/2022GL099422.
- Lee, Yoonsang (2021). *Sampling error correction in ensemble Kalman inversion*. DOI: 10.48550/arxiv.2105.11341. eprint: 2105.11341.
- Li, Dan (2019). “Turbulent Prandtl number in the atmospheric boundary layer - where are we now?” In: *Atmospheric Research* 216, pp. 86–105. DOI: <https://doi.org/10.1016/j.atmosres.2018.09.015>.
- Li, Dan, Scott T Salesky, and Tirtha Banerjee (2016). “Connections between the Ozmidov scale and mean velocity profile in stably stratified atmospheric surface layers”. In: *Journal of Fluid Mechanics* 797, R3. DOI: 10.1017/jfm.2016.311.
- Li, J et al. (2018). “Can regional climate modeling capture the observed changes in spatial organization of extreme storms at higher temperatures?” In: *Geophysical Research Letters* 45 (9), pp. 4475–4484. ISSN: 0094-8276. DOI: 10.1029/2018GL077716.
- Li, Mu et al. (2014). “Efficient mini-batch training for stochastic optimization”. In: ACM. ISBN: 9781450329569. DOI: 10.1145/2623330.2623612.
- Li, Qi, Yu Cheng, and Pierre Gentine (2021). “Connection between mass flux transport and eddy diffusivity in convective atmospheric boundary layers”. In: *Geophysical Research Letters* 48 (8), e2020GL092073. ISSN: 0094-8276. DOI: 10.1029/2020GL092073.
- Li, Zongyi et al. (2020). *Fourier neural operator for parametric partial differential equations*. DOI: 10.48550/arxiv.2010.08895.
- Lilly, D K (1968). “Models of cloud-topped mixed layers under a strong inversion”. In: *Quarterly Journal of the Royal Meteorological Society* 94, pp. 292–309. DOI: 10.1002/qj.49709440106.
- Lin, Jia-Lin, Taotao Qian, and Toshiaki Shinoda (2014). “Stratocumulus clouds in Southeastern Pacific simulated by eight CMIP5–CFMIP global climate models”. In: *Journal of Climate* 27 (8), pp. 3000–3022. DOI: 10.1175/JCLI-D-13-00376.1.
- Ling, Julia, Andrew Kurzawski, and Jeremy Templeton (2016). “Reynolds averaged turbulence modelling using deep neural networks with embedded invariance”. In: *Journal of Fluid Mechanics* 807, pp. 155–166. ISSN: 0022-1120. DOI: 10.1017/jfm.2016.615.

- Lopez-Gomez, Ignacio, Yair Cohen, Jia He, Anna Jaruga, and Tapio Schneider (2020). “A generalized mixing length closure for eddy-diffusivity mass-flux schemes of turbulence and convection”. In: *Journal of Advances in Modeling Earth Systems* 12, e2020MS002161. DOI: 10.1029/2020MS002161.
- Lopez-Gomez, Ignacio et al. (2022a). *Global extreme heat forecasting using neural weather models*. DOI: 10.48550/arxiv.2205.10972.
- Lopez-Gomez, Ignacio, Costa Christopoulos, Haakon Ludvig Langeland Ervik, Oliver R. A. Dunbar, Yair Cohen, and Tapio Schneider (2022b). “Training physics-based machine-learning parameterizations with gradient-free ensemble Kalman methods”. In: *Journal of Advances in Modeling Earth Systems* 14, e2022MS003105. DOI: 10.1029/2022MS003105.
- Lorenz, E. N. (1975). “Climatic predictability”. In: *The Physical Basis of Climate and Climate Modelling*. Vol. 16. GARP Publication Series. World Meteorological Organization, pp. 132–136.
- Madeleine, Jean-Baptiste et al. (2020). “Improved representation of clouds in the atmospheric component LMDZ6A of the IPSL-CM6A Earth system model”. In: *Journal of Advances in Modeling Earth Systems* 12 (10), e2020MS002046. ISSN: 1942-2466. DOI: 10.1029/2020MS002046.
- Manabe, Syukuro, Joseph Smagorinsky, and Robert F Strickler (1965). “Simulated climatology of a general circulation model with a hydrologic cycle”. English. In: *Monthly Weather Review* 93.12, pp. 769–798. DOI: 10.1175/1520-0493(1965)093<0769:SCOAGC>2.3.CO;2.
- Manabe, Syukuro and Robert F Strickler (1964). “Thermal equilibrium of the atmosphere with a convective adjustment”. English. In: *Journal of Atmospheric Sciences* 21.4, pp. 361–385. DOI: 10.1175/1520-0469(1964)021<0361:TEOTAW>2.0.CO;2.
- Manabe, Syukuro and Richard T Wetherald (1967). “Thermal equilibrium of the atmosphere with a given distribution of relative humidity”. English. In: *Journal of Atmospheric Sciences* 24.3, pp. 241–259. DOI: 10.1175/1520-0469(1967)024<0241:TEOTAW>2.0.CO;2.
- (1975). “The effects of doubling the CO<sub>2</sub> concentration on the climate of a general circulation model”. English. In: *Journal of Atmospheric Sciences* 32.1, pp. 3–15. DOI: 10.1175/1520-0469(1975)032<0003:TEODTC>2.0.CO;2.
- Marquet, Pascal (2011). “Definition of a moist entropy potential temperature: application to FIRE-I data flights”. In: *Quarterly Journal of the Royal Meteorological Society* 137.656, pp. 768–791. DOI: 10.1002/qj.787.
- Mauritsen, Thorsten et al. (2012). “Tuning the climate of a global model”. In: *Journal of Advances in Modeling Earth Systems* 4.3. ISSN: 1942-2466. DOI: 10.1029/2012MS000154.

- McIntyre, William A, Georgios A Efstathiou, and John Thuburn (2022). “A two-fluid single-column model of turbulent shallow convection, Part III: Results and parameter sensitivity”. In: *Quarterly Journal of the Royal Meteorological Society*. ISSN: 0035-9009.
- Mellor, George L (1977). “The Gaussian cloud model relations”. In: *Journal of the Atmospheric Sciences* 34, pp. 356–358. DOI: 10.1175/1520-0469(1977)034<0356:TGCMR>2.0.CO;2.
- Mellor, George L and Tetsuji Yamada (1974). “A hierarchy of turbulence closure models for planetary boundary layers”. In: *Journal of Atmospheric Sciences* 31.7, pp. 1791–1806. DOI: 10.1175/1520-0469(1974)031<1791:AHOTCM>2.0.CO;2.
- (1982). “Development of a turbulence closure model for geophysical fluid problems”. In: *Reviews of Geophysics* 20 (4), pp. 851–875. ISSN: 8755-1209. DOI: 10.1029/RG020i004p00851.
- Metropolis, Nicholas et al. (1953). “Equation of state calculations by fast computing machines”. In: *The Journal of Chemical Physics* 21, pp. 1087–1092. ISSN: 0021-9606. DOI: 10.1063/1.1699114.
- Meyer, David et al. (2022). “Machine learning emulation of 3D cloud radiative effects”. In: *Journal of Advances in Modeling Earth Systems* 14, e2021MS002550. ISSN: 1942-2466. DOI: 10.1029/2021MS002550.
- Mlawer, Eli J et al. (1997). “Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave”. In: *Journal of Geophysical Research: Atmospheres* 102, pp. 16663–16682. ISSN: 0148-0227. DOI: 10.1029/97JD00237.
- Möller, F (1963). “On the influence of changes in the CO<sub>2</sub> concentration in air on the radiation balance of the Earth’s surface and on the climate”. In: *Journal of Geophysical Research (1896-1977)* 68.13, pp. 3877–3886. DOI: 10.1029/JZ068i013p03877.
- Monin, A.S. and A.M. Obukhov (1954). “Basic laws of turbulent mixing in the surface layer of the atmosphere”. In: *Trudy Akademii Nauk SSSR Geofizicheskii Institut* 24, pp. 163–187.
- Moral, Pierre Del, Arnaud Doucet, and Ajay Jasra (2006). “Sequential Monte Carlo samplers”. In: *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 68, pp. 411–436. ISSN: 13697412, 14679868.
- Morzfeld, Matthias et al. (2018). “Feature-based data assimilation in geophysics”. In: *Nonlinear Processes in Geophysics* 25, pp. 355–374. ISSN: 1607-7946. DOI: 10.5194/npg-25-355-2018.
- Muñoz-Sabater, Joaquín et al. (2021). “ERA5-Land: a state-of-the-art global reanalysis dataset for land applications”. In: *Earth System Science Data* 13, pp. 4349–4383. ISSN: 1866-3516. DOI: 10.5194/essd-13-4349-2021.

- Nakanishi, Mikio (2001). “Improvement of The Mellor-Yamada turbulence closure model based on large-eddy simulation data”. In: *Boundary-Layer Meteorology* 99, pp. 349–378. DOI: 10.1023/A:1018915827400.
- National Academies of Sciences, Engineering, and Medicine (2018). *Thriving on our changing planet: A decadal strategy for Earth observation from space*. Washington, DC: The National Academies Press. ISBN: 978-0-309-46757-5. DOI: 10.17226/24938.
- Neggers, Roel A J (2009). “A dual mass flux framework for boundary layer convection. Part II: Clouds”. English. In: *Journal of the Atmospheric Sciences* 66, pp. 1489–1506. DOI: 10.1175/2008JAS2636.1.
- (2015). “Exploring bin-macrophysics models for moist convective transport and clouds”. In: *Journal of Advances in Modeling Earth Systems* 7 (4), pp. 2079–2104. ISSN: 1942-2466. DOI: 10.1002/2015MS000502.
- Neggers, Roel A J and P J Griewank (2022). “A decentralized approach for modeling organized convection based on thermal populations on microgrids”. In: *Journal of Advances in Modeling Earth Systems* (n/a), e2022MS003042. ISSN: 1942-2466. DOI: 10.1029/2022MS003042.
- Nelsen, Nicholas H and Andrew M Stuart (2021). “The random feature model for input-output maps between Banach spaces”. In: *SIAM Journal on Scientific Computing* 43, A3212–A3243. ISSN: 1064-8275. DOI: 10.1137/20M133957X.
- Nieuwstadt, F T M et al. (1993). “Large-eddy simulation of the convective boundary layer: A comparison of four computer codes”. In: ed. by Franz Durst et al. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 343–367. ISBN: 978-3-642-77674-8.
- Nitta, Tsuyoshi (1975). “Observational determination of cloud mass flux distributions”. In: *Journal of Atmospheric Sciences* 32 (1), pp. 73–91. DOI: 10.1175/1520-0469(1975)032<0073:ODOCMF>2.0.CO;2.
- O’Gorman, Paul A and John G Dwyer (2018). “Using machine learning to parameterize moist convection: Potential for modeling of climate, climate change, and extreme events”. In: *Journal of Advances in Modeling Earth Systems* 10 (10), pp. 2548–2563. ISSN: 1942-2466. DOI: 10.1029/2018MS001351.
- O’Gorman, Paul A and Tapio Schneider (2006). “Stochastic models for the kinematics of moisture transport and condensation in homogeneous turbulent flows”. In: *Journal of the Atmospheric Sciences* 63.11, pp. 2992–3005. ISSN: 0022-4928. DOI: 10.1175/JAS3794.1.
- O’Gorman, Paul A (2011). “The effective static stability experienced by eddies in a moist atmosphere”. In: *Journal of the Atmospheric Sciences* 68.1, pp. 75–90. ISSN: 0022-4928. DOI: 10.1175/2010JAS3537.1.

- Ogura, Yoshimitsu and Han-Ru Cho (1973). “Diagnostic determination of cumulus cloud populations from observed large-scale variables”. In: *Journal of Atmospheric Sciences* 30 (7), pp. 1276–1286. DOI: 10.1175/1520-0469(1973)030<1276:DDOCCP>2.0.CO;2.
- Ooyama, Katsuyuki (1971). “A theory on parameterization of cumulus convection”. In: *Journal of the Meteorological Society of Japan. Ser. II* 49A, pp. 744–756. DOI: 10.2151/jmsj1965.49A.0\_744.
- Optis, Michael, Adam Monahan, and Fred C Bosveld (2016). “Limitations and breakdown of Monin–Obukhov similarity theory for wind profile extrapolation under stable stratification”. In: *Wind Energy* 19, pp. 1053–1072. DOI: 10.1002/we.1883.
- Park, Sungsu (2014). “A unified convection scheme (UNICON). Part I: Formulation”. In: *Journal of the Atmospheric Sciences* 71, pp. 3902–3930. DOI: 10.1175/JAS-D-13-0233.1.
- Pathak, Jaideep et al. (2022). *FourCastNet: A global data-driven high-resolution weather model using adaptive Fourier neural operators*. DOI: 10.48550/arxiv.2202.11214.
- Pauluis, Olivier (2008). “Thermodynamic consistency of the anelastic approximation for a moist atmosphere”. In: *Journal of the Atmospheric Sciences* 65.8, pp. 2719–2729. DOI: 10.1175/2007JAS2475.1.
- Pincus, Robert, Howard W Barker, and Jean-Jacques Morcrette (July 2003). “A fast, flexible, approximate technique for computing radiative transfer in inhomogeneous cloud fields”. In: *Journal of Geophysical Research: Atmospheres* 108 (D13). ISSN: 0148-0227. DOI: 10.1029/2002JD003322.
- Planck, Max (1901). “Über das gesetz der energieverteilung im normalspectrum”. In: *Annalen der Physik* 309, pp. 553–563.
- Plass, Gilbert N (1956). “Infrared radiation in the atmosphere”. In: *American Journal of Physics* 24.5, pp. 303–321. ISSN: 0002-9505. DOI: 10.1119/1.1934220.
- Pouillet, Claude Servais Mathias (1838). “Mémoire sur la chaleur solaire, sur les pouvoirs rayonnants et absorbants de l’air atmosphérique et sur la température de l’espace”. In: *Comptes Rendus des Séances de l’Académie des Sciences*.
- Prechelt, Lutz (1998). *Early stopping - but when?* Ed. by Genevieve B Orr and Klaus-Robert Müller. DOI: 10.1007/3-540-49430-8\_3.
- Pressel, Kyle G et al. (2015). “Large-eddy simulation in an anelastic framework with closed water and entropy balances”. In: *Journal of Advances in Modeling Earth Systems* 7, pp. 1425–1456. ISSN: 1942-2466. DOI: 10.1002/2015MS000496.
- Pressel, Kyle G et al. (2017). “Numerics and subgrid-scale modeling in large eddy simulations of stratocumulus clouds”. In: *Journal of Advances in Modeling Earth Systems* 9, pp. 1342–1365. DOI: 10.1002/2016MS000778.

- Randall, David A (2013). “Beyond deadlock”. In: *Geophysical Research Letters* 40, pp. 5970–5976. DOI: 10.1002/2013GL057998.
- Randall, David A and George J Huffman (1982). “Entrainment and detrainment in a simple cumulus cloud model”. In: *Journal of Atmospheric Sciences* 39 (12), pp. 2793–2806. DOI: 10.1175/1520-0469(1982)039<2793:EADIAS>2.0.CO;2.
- Rasp, Stephan, Michael S Pritchard, and Pierre Gentine (2018). “Deep learning to represent subgrid processes in climate models”. In: *Proceedings of the National Academy of Sciences* 115, pp. 9684–9689. DOI: 10.1073/pnas.1810286115.
- Rasp, Stephan and Nils Thuerey (2021). “Data-driven medium-range weather prediction with a Resnet pretrained on climate simulations: A new model for Weather-Bench”. In: *Journal of Advances in Modeling Earth Systems* 13, e2020MS002405. ISSN: 1942-2466. DOI: 10.1029/2020MS002405.
- Ravuri, Suman et al. (2021). “Skilful precipitation nowcasting using deep generative models of radar”. In: *Nature* 597, pp. 672–677. ISSN: 1476-4687. DOI: 10.1038/s41586-021-03854-z.
- Reichel, Lothar and Giuseppe Rodriguez (2013). “Old and new parameter choice rules for discrete ill-posed problems”. In: *Numerical Algorithms* 63, pp. 65–87. DOI: 10.1007/s11075-012-9612-8.
- Reichstein, Markus et al. (2019). “Deep learning and process understanding for data-driven Earth system science”. In: *Nature* 566, pp. 195–204. DOI: 10.1038/s41586-019-0912-1.
- Romps, David M (2008). “The dry-entropy budget of a moist atmosphere”. In: *Journal of the Atmospheric Sciences* 65, pp. 3779–3799. DOI: 10.1175/2008JAS2679.1.
- Romps, David M and Alexander B Charn (2015). “Sticky thermals: Evidence for a dominant balance between buoyancy and drag in cloud updrafts”. In: *Journal of the Atmospheric Sciences* 72.8, pp. 2890–2901. ISSN: 0022-4928. DOI: 10.1175/JAS-D-15-0042.1.
- Saint-Martin, D et al. (2021). “Tracking changes in climate sensitivity in CNRM climate models”. In: *Journal of Advances in Modeling Earth Systems* 13 (6), e2020MS002190. ISSN: 1942-2466. DOI: 10.1029/2020MS002190.
- Savre, Julien and Michael Herzog (2019). “A general description of entrainment in buoyant cloudy plumes including the effects of mixing-induced evaporation”. In: *Journal of the Atmospheric Sciences* 76 (2), pp. 479–496. DOI: 10.1175/JAS-D-17-0326.1.
- Schillings, Claudia and Andrew M Stuart (2017). “Analysis of the ensemble Kalman filter for inverse problems”. In: *SIAM Journal on Numerical Analysis* 55, pp. 1264–1290. ISSN: 0036-1429. DOI: 10.1137/16M105959X.

- Schmit, Timothy J et al. (2017). “A closer look at the ABI on the GOES-R series”. In: *Bulletin of the American Meteorological Society* 98, pp. 681–698. DOI: 10.1175/BAMS-D-15-00230.1.
- Schneider, Tapio and Stephen M Griffies (1999). “A conceptual framework for predictability studies”. In: *Journal of Climate* 12, pp. 3133–3155. DOI: 10.1175/1520-0442(1999)012<3133:ACFFPS>2.0.CO;2.
- Schneider, Tapio, Colleen M Kaul, and Kyle G Pressel (2019). “Possible climate transitions from breakup of stratocumulus decks under greenhouse warming”. In: *Nature Geoscience* 12.3, pp. 163–167. ISSN: 1752-0908. DOI: 10.1038/s41561-019-0310-1. URL: <https://doi.org/10.1038/s41561-019-0310-1>.
- Schneider, Tapio, Andrew M Stuart, and Jin-Long Wu (2021a). “Learning stochastic closures using ensemble Kalman inversion”. In: *Transactions of Mathematics and Its Applications* 5, ttab003. ISSN: 2398-4945. DOI: 10.1093/imatrm/ttab003.
- (2020). *Ensemble Kalman inversion for sparse learning of dynamical systems from time-averaged data*. DOI: 10.48550/arxiv.2007.06175.
- Schneider, Tapio, Joao Teixeira, Christopher S Bretherton, Florent Brient, Kyle G Pressel, Christoph Schär, and A Pier Siebesma (2017a). “Climate goals and computing the future of clouds”. In: *Nature Climate Change* 7 (3). DOI: 10.1038/nclimate3190.
- Schneider, Tapio, Shiwei Lan, Andrew Stuart, and Joao Teixeira (2017b). “Earth system modeling 2.0: A blueprint for models that learn from observations and targeted high-resolution simulations”. In: *Geophysical Research Letters* 44 (24), pp. 12, 312–396, 417. ISSN: 0094-8276. DOI: 10.1002/2017GL076101.
- Schneider, Tapio et al. (2021b). “Epidemic management and control through risk-dependent individual contact interventions”. In: DOI: 10.48550/arXiv.2109.10970.
- Schubert, Wayne H (1976). “Experiments with Lilly’s cloud-topped mixed layer model”. In: *Journal of the Atmospheric Sciences* 33, pp. 436–446. DOI: 10.1175/1520-0469(1976)033<0436:EWLCTM>2.0.CO;2.
- Schumann, Ulrich and Chin-Hoh Moeng (1991). “Plume budgets in clear and cloudy convective boundary layers”. In: *Journal of the Atmospheric Sciences* 48, pp. 1758–1770. DOI: 10.1175/1520-0469(1991)048<1758:PBICAC>2.0.CO;2.
- Seifert, Axel and Stephan Rasp (2020). “Potential and limitations of machine learning for modeling warm-rain cloud microphysical processes”. In: *Journal of Advances in Modeling Earth Systems* 12, e2020MS002301. ISSN: 1942-2466. DOI: 10.1029/2020MS002301.

- Shen, Zhaoyi, Akshay Sridhar, Zhihong Tan, Anna Jaruga, and Tapio Schneider (2022). “A library of large-eddy simulations forced by global climate models”. In: *Journal of Advances in Modeling Earth Systems*, e2021MS002631. ISSN: 1942-2466. DOI: 10.1029/2021MS002631.
- Sherwood, Steven C, Sandrine Bony, and Jean-Louis Dufresne (2014). “Spread in model climate sensitivity traced to atmospheric convective mixing”. In: *Nature* 505, p. 37. DOI: doi.org/10.1038/nature12829.
- Shin, Jihoon and Jong-Jin Baik (2022). “Parameterization of stochastically entraining convection using machine learning technique”. In: *Journal of Advances in Modeling Earth Systems* 14.5. e2021MS002817 2021MS002817, e2021MS002817. DOI: 10.1029/2021MS002817.
- Shiple, Daniel, Hilary Weller, Peter A Clark, and William A McIntyre (2022). “Two-fluid single-column modelling of Rayleigh–Bénard convection as a step towards multi-fluid modelling of atmospheric convection”. In: *Quarterly Journal of the Royal Meteorological Society* 148.742, pp. 351–377. ISSN: 0035-9009. DOI: 10.1002/qj.4209.
- Siebesma, A Pier (1996). “On the mass flux approach for atmospheric convection”. In: *Workshop on New Insights and Approaches to Convective Parametrization, 4-7 November 1996*. ECMWF. Shinfield Park, Reading: ECMWF, pp. 25–57.
- (1998). *Shallow cumulus convection*. Ed. by E J Plate et al. DOI: 10.1007/978-94-011-5058-3\_19.
- Siebesma, A Pier, Pedro M M Soares, and Joao Teixeira (2007). “A combined eddy-diffusivity mass-flux approach for the convective boundary layer”. In: *Journal of the Atmospheric Sciences* 64 (4), pp. 1230–1248. DOI: 10.1175/JAS3888.1.
- Siebesma, A Pier and Joao Teixeira (2000). “An advection-diffusion scheme for the convective boundary layer: description and 1d-results”. In: *Proceedings of the 14th AMS Symposium on Boundary Layers and Turbulence*, pp. 133–136.
- Siebesma, A Pier et al. (2003). “A large eddy simulation intercomparison study of shallow cumulus convection”. In: *Journal of the Atmospheric Sciences* 60, pp. 1201–1219. DOI: 10.1175/1520-0469(2003)60<1201:ALESIS>2.0.CO;2.
- Simpson, Joanne and Victor Wiggert (1969). “Models of precipitating cumulus towers”. In: *Monthly Weather Review* 97, pp. 471–489. DOI: 10.1175/1520-0493(1969)097<0471:MOPCT>2.3.CO;2.
- Singer, Clare E, Ignacio Lopez-Gomez, Xiyue Zhang, and Tapio Schneider (2021). “Top-of-atmosphere albedo bias from neglecting three-dimensional cloud radiative effects”. In: *Journal of the Atmospheric Sciences* 78 (12), pp. 4053–4069. DOI: 10.1175/JAS-D-21-0032.1.

- Skamarock, William C et al. (2019). “A description of the advanced research WRF model version 4”. In: *National Center for Atmospheric Research: Boulder, CO, USA* 145, p. 145.
- Smagorinsky, Joseph, Syukuro Manabe, and J Leith Holloway (1965). “Numerical results from a nine-level general circulation model of the atmosphere”. In: *Monthly Weather Review* 93 (12), pp. 727–768. DOI: 10.1175/1520-0493(1965)093<0727:NRFANL>2.3.CO;2.
- Smalley, M A et al. (2022). “Coupling warm rain With an eddy diffusivity/mass flux parameterization: 2. Sensitivities and comparison to observations”. In: *Journal of Advances in Modeling Earth Systems* 14 (8), e2021MS002729. ISSN: 1942-2466. DOI: 10.1029/2021MS002729.
- Soares, P M M et al. (2004). “An eddy-diffusivity/mass-flux parametrization for dry and shallow cumulus convection”. In: *Quarterly Journal of the Royal Meteorological Society* 130, pp. 3365–3383. DOI: 10.1256/qj.03.223.
- Sommeria, G. and James W. Deardorff (1977). “Subgrid-scale condensation in models of nonprecipitating clouds”. In: *Journal of the Atmospheric Sciences* 34, pp. 344–355. DOI: 10.1175/1520-0469(1977)034<0344:SSCIMO>2.0.CO;2.
- Sønderby, Casper Kaae et al. (2020). *MetNet: A neural weather model for precipitation forecasting*. DOI: 10.48550/arXiv.2003.12140.
- Sorbjan, Zbigniew and Andrey A Grachev (2010). “An evaluation of the flux–gradient relationship in the stable boundary layer”. In: *Boundary-Layer Meteorology* 135.3, pp. 385–405. DOI: 10.1007/s10546-010-9482-3.
- Souza, A N et al. (2020). “Uncertainty quantification of ocean parameterizations: Application to the K-profile-parameterization for penetrative convection”. In: *Journal of Advances in Modeling Earth Systems* 12 (12), e2020MS002108. ISSN: 1942-2466. DOI: 10.1029/2020MS002108.
- Spalart, Philippe R (1988). “Direct simulation of a turbulent boundary layer up to  $R\theta = 1410$ ”. In: *Journal of Fluid Mechanics* 187, pp. 61–98. DOI: 10.1017/S0022112088000345.
- Squires, P (1958). “Penetrative downdraughts in cumuli”. In: *Tellus* 10 (3), pp. 381–389. ISSN: 0040-2826. DOI: 10.1111/j.2153-3490.1958.tb02025.x.
- Squires, P and J. S. Turner (1962). “An entraining jet model for cumulo-nimbus updraughts”. In: *Tellus* 14 (4), pp. 422–434. ISSN: 00402826. DOI: 10.1111/j.2153-3490.1962.tb01355.x.
- Stefan, Josef (1879). “Über die beziehung zwischen der wärmestrahlung und der temperatur”. In: *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften, Mathematische- Naturwissenschaftliche Classe Abteilung II* 79, pp. 391–428.

- Stensrud, David J and J Michael Fritsch (1994). “Mesoscale convective systems in weakly forced large-scale environments. Part III: Numerical simulations and implications for operational forecasting”. English. In: *Monthly Weather Review* 122.9, pp. 2084–2104. DOI: 10.1175/1520-0493(1994)122<2084:MCSIWF>2.0.CO;2.
- Stephens, G L (1978). “Radiation profiles in extended water clouds. II: Parameterization schemes”. In: *Journal of Atmospheric Sciences* 35 (11), pp. 2123–2132. DOI: 10.1175/1520-0469(1978)035<2123:RPIEWC>2.0.CO;2.
- Stevens, Bjorn (2007). “On the growth of layers of nonprecipitating cumulus convection”. In: *Journal of the Atmospheric Sciences* 64.8, pp. 2916–2931. ISSN: 0022-4928. DOI: 10.1175/JAS3983.1.
- Stevens, Bjorn et al. (2003). “Dynamics and chemistry of marine stratocumulus - DYCOMS-II”. In: *Bulletin of the American Meteorological Society* 84, pp. 579–594. DOI: 10.1175/BAMS-84-5-579.
- Stevens, Bjorn et al. (2005). “Evaluation of large-eddy simulations via observations of nocturnal marine stratocumulus”. In: *Monthly Weather Review* 133, pp. 1443–1462. DOI: 10.1175/MWR2930.1.
- Stevens, Bjorn et al. (2013). “Atmospheric component of the MPI-M Earth System Model: ECHAM6”. In: *Journal of Advances in Modeling Earth Systems* 5.2, pp. 146–172. ISSN: 1942-2466. DOI: 10.1002/jame.20015.
- Suselj, Kay, Timothy F Hogan, and João Teixeira (2014). “Implementation of a stochastic eddy-diffusivity/mass-flux parameterization into the Navy global environmental model”. In: *Weather and Forecasting* 29 (6), pp. 1374–1390. DOI: 10.1175/WAF-D-14-00043.1.
- Suselj, Kay, Marcin J Kurowski, and Joao Teixeira (2019a). “A unified eddy-diffusivity/mass-flux approach for modeling atmospheric convection”. In: *Journal of the Atmospheric Sciences* 76 (8), pp. 2505–2537. DOI: 10.1175/JAS-D-18-0239.1.
- (2019b). “On the factors controlling the development of shallow convection in eddy-diffusivity/mass-flux models”. In: *Journal of the Atmospheric Sciences* 76 (2), pp. 433–456. DOI: 10.1175/JAS-D-18-0121.1.
- Suselj, Kay, Joao Teixeira, and Daniel Chung (2013). “A unified model for moist convective boundary layers based on a stochastic eddy-diffusivity/mass-flux parameterization”. In: *Journal of the Atmospheric Sciences* 70, pp. 1929–1953. DOI: 10.1175/JAS-D-12-0106.1.
- Suselj, Kay et al. (2020). “A new methodology for observation-based parameterization development”. In: *Monthly Weather Review* 148, pp. 4159–4184. DOI: 10.1175/MWR-D-20-0114.1.

- Suselj, Kay et al. (2021). “Improving the representation of subtropical boundary layer clouds in the NASA GEOS model with the eddy-diffusivity/mass-flux parameterization”. In: *Monthly Weather Review* 149, pp. 793–809. DOI: 10.1175/MWR-D-20-0183.1.
- Suselj, Kay et al. (2022). “Coupling warm rain with an eddy diffusivity/mass flux parameterization: 1. Model description and validation”. In: *Journal of Advances in Modeling Earth Systems* 14 (8), e2021MS002736. ISSN: 1942-2466. DOI: 10.1029/2021MS002736.
- Takaya, Yuhei et al. (2018). “Japan Meteorological Agency/Meteorological Research Institute-Coupled Prediction System version 2 (JMA/MRI-CPS2): atmosphere–land–ocean–sea ice coupled prediction system for operational seasonal forecasting”. In: *Climate Dynamics* 50 (3), pp. 751–765. ISSN: 1432-0894. DOI: 10.1007/s00382-017-3638-5.
- Tan, Zhihong, Colleen M Kaul, Kyle G Pressel, Yair Cohen, Tapio Schneider, and Joao Teixeira (2018). “An extended eddy-diffusivity mass-flux scheme for unified representation of subgrid-scale turbulence and convection”. In: *Journal of Advances in Modeling Earth Systems* 10, pp. 770–800. DOI: 10.1002/2017MS001162.
- Tarantola, Albert (2005). *Inverse problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics. ISBN: 978-0-89871-572-9. DOI: 10.1137/1.9780898717921.
- Teixeira, Joao et al. (2011). “Tropical and subtropical cloud transitions in weather and climate prediction models: The GCSS/WGNE Pacific cross-section intercomparison (GPCI)”. In: *Journal of Climate* 24 (20), pp. 5223–5256. DOI: 10.1175/2011JCLI3672.1.
- Thuburn, John, Georgios A Efstathiou, and Robert J Beare (2019). “A two-fluid single-column model of the dry, shear-free, convective boundary layer”. In: *Quarterly Journal of the Royal Meteorological Society* 145 (721), pp. 1535–1550. ISSN: 0035-9009. DOI: 10.1002/qj.3510.
- Thuburn, John, Georgios A Efstathiou, and William A McIntyre (2022a). “A two-fluid single-column model of turbulent shallow convection, Part I: Turbulence equations in the multi-fluid framework”. In: *Quarterly Journal of the Royal Meteorological Society*. ISSN: 0035-9009. DOI: 10.1002/qj.4366.
- (2022b). “A two-fluid single-column model of turbulent shallow convection, Part II: Single-column model formulation and numerics”. In: *Quarterly Journal of the Royal Meteorological Society*. ISSN: 0035-9009. DOI: 10.1002/qj.4361.
- Thuburn, John, Hilary Weller, Geoffrey K Vallis, Robert J Beare, and Michael Whitall (2018). “A framework for convection and boundary layer parameterization derived from conditional filtering”. In: *Journal of the Atmospheric Sciences* 75, pp. 965–981. DOI: 10.1175/JAS-D-17-0130.1.

- Tiedtke, M (1989). “A comprehensive mass flux scheme for cumulus parameterization in large-scale models”. English. In: *Monthly Weather Review* 117, pp. 1779–1800. DOI: 10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2.
- Titsias, Michalis K. (2016). “One-vs-each approximation to Softmax for scalable estimation of probabilities”. In: *Proceedings of the 30th International Conference on Neural Information Processing Systems*. NIPS’16. Barcelona, Spain: Curran Associates Inc., pp. 4168–4176. ISBN: 978-1-5108-3881-9.
- Tong, Xin T. and Matthias Morzfeld (2022). *Localization in ensemble Kalman inversion*. DOI: 10.48550/arXiv.2201.10821.
- Turner, J. S. (1962). “The ‘starting plume’ in neutral surroundings”. In: *Journal of Fluid Mechanics* 13 (3), pp. 356–368. ISSN: 0022-1120. DOI: 10.1017/S0022112062000762.
- Tyndall, John (1865). *Heat a mode of motion*. 2nd ed. London, UK: Longmans.
- Umlauf, L. and H. Burchard (2003). “A generic length-scale equation for geophysical turbulence models”. In: *Journal of Marine Research* 61, pp. 235–265. DOI: 10.1357/002224003322005087.
- van Leeuwen, Peter Jan (2015). “Representation errors and retrievals in linear and nonlinear data assimilation”. In: *Quarterly Journal of the Royal Meteorological Society* 141, pp. 1612–1623. ISSN: 0035-9009. DOI: 10.1002/qj.2464.
- Vassilicos, J Christos (2015). “Dissipation in turbulent flows”. In: *Annual Review of Fluid Mechanics* 47, pp. 95–114. DOI: 10.1146/annurev-fluid-010814-014637.
- Verstappen, Roel (2011). “When does eddy viscosity damp subfilter scales sufficiently?” In: *Journal of Scientific Computing* 49, p. 94. DOI: 10.1007/s10915-011-9504-4.
- Villalba-Pradas, Anahí and Francisco J. Tapiador (2022). “Empirical values and assumptions in the convection schemes of numerical models”. In: *Geoscientific Model Development* 15, pp. 3447–3518. ISSN: 1991-9603. DOI: 10.5194/gmd-15-3447-2022.
- Villefranque, Najda et al. (2021). “Process-based climate model development harnessing machine learning: III. The representation of cumulus geometry and their 3D radiative effects”. In: *Journal of Advances in Modeling Earth Systems* 13, e2020MS002423. ISSN: 1942-2466. DOI: 10.1029/2020MS002423.
- Waite, Michael L (2011). “Stratified turbulence at the buoyancy scale”. In: *Physics of Fluids* 23.6, p. 66602. DOI: 10.1063/1.3599699.
- Walters, David et al. (2019). “The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations”. In: *Geoscientific Model Development* 12.5, pp. 1909–1963. ISSN: 1991-9603. DOI: 10.5194/gmd-12-1909-2019.

- Wang, Xin et al. (2022). “Stable climate simulations using a realistic general circulation model with neural network parameterizations for atmospheric moist physics and radiation processes”. In: *Geoscientific Model Development* 15 (9), pp. 3923–3940. ISSN: 1991-9603. DOI: 10.5194/gmd-15-3923-2022.
- Webb, Mark J et al. (2015). “The diurnal cycle of marine cloud feedback in climate models”. In: *Climate Dynamics* 44 (5), pp. 1419–1436. ISSN: 1432-0894. DOI: 10.1007/s00382-014-2234-1.
- Weller, Hilary and William A McIntyre (2019). “Numerical solution of the conditionally averaged equations for representing net mass flux due to convection”. In: *Quarterly Journal of the Royal Meteorological Society* 145.721, pp. 1337–1353. ISSN: 0035-9009. DOI: 10.1002/qj.3490.
- Weyn, Jonathan A et al. (2021). “Sub-seasonal forecasting with a large ensemble of deep-learning weather prediction models”. In: *Journal of Advances in Modeling Earth Systems* 13, e2021MS002502. ISSN: 1942-2466. DOI: 10.1029/2021MS002502.
- Witek, Marcin L, Joao Teixeira, and Georgios Matheou (2011a). “An eddy diffusivity–mass flux approach to the vertical transport of turbulent kinetic energy in convective boundary layers”. In: *Journal of the Atmospheric Sciences* 68, pp. 2385–2394. ISSN: 0022-4928. DOI: 10.1175/jas-d-11-06.1.
- (2011b). “An integrated TKE-based eddy diffusivity/mass flux boundary layer closure for the dry convective boundary layer”. In: *Journal of the Atmospheric Sciences* 68, pp. 1526–1540. DOI: 10.1175/2011JAS3548.1.
- Wolfram, Stephen (1983). “Statistical mechanics of cellular automata”. In: *Reviews of Modern Physics* 55 (3), pp. 601–644. DOI: 10.1103/RevModPhys.55.601.
- Wood, R. et al. (Jan. 2011). “The VAMOS Ocean-Cloud-Atmosphere-Land Study Regional Experiment (VOCALS-REx): goals, platforms, and field operations”. In: *Atmospheric Chemistry and Physics* 11 (2), pp. 627–654. ISSN: 1680-7324. DOI: 10.5194/acp-11-627-2011.
- Wu, Elynn, Handa Yang, Jan Kleissl, Kay Suselj, Marcin J Kurowski, and Joao Teixeira (2020). “On the parameterization of convective downdrafts for marine stratocumulus clouds”. In: *Monthly Weather Review* 148, pp. 1931–1950. DOI: 10.1175/MWR-D-19-0292.1.
- Wu, Tongwen (2012). “A mass-flux cumulus parameterization scheme for large-scale models: description and test with observations”. In: *Climate Dynamics* 38.3, pp. 725–744. ISSN: 1432-0894. DOI: 10.1007/s00382-011-0995-3.
- Wu, Tongwen et al. (2019). “The Beijing Climate Center Climate System Model (BCC-CSM): the main progress from CMIP5 to CMIP6”. In: *Geoscientific Model Development* 12 (4), pp. 1573–1600. ISSN: 1991-9603. DOI: 10.5194/gmd-12-1573-2019.

- Xiao, H et al. (2016). “Quantifying and reducing model-form uncertainties in Reynolds-averaged Navier–Stokes simulations: A data-driven, physics-informed Bayesian approach”. In: *Journal of Computational Physics* 324, pp. 115–136. ISSN: 0021-9991. DOI: 10.1016/j.jcp.2016.07.038.
- Yanai, Michio, Steven Esbensen, and Jan-Hwa Chu (1973). “Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets”. In: *Journal of Atmospheric Sciences* 30 (4), pp. 611–627. DOI: 10.1175/1520-0469(1973)030<0611:DOBPOT>2.0.CO;2.
- Yuval, Janni, Paul A O’Gorman, and Chris N Hill (2021). “Use of neural networks for stable, accurate and physically consistent parameterization of sub-grid atmospheric processes with good performance at reduced precision”. In: *Geophysical Research Letters* 48 (6), e2020GL091363. ISSN: 0094-8276. DOI: 10.1029/2020GL091363.
- Yuval, Janni and Paul A O’Gorman (2020). “Stable machine-learning parameterization of subgrid processes for climate modeling at a range of resolutions”. In: *Nature Communications* 11 (1), p. 3295. ISSN: 2041-1723. DOI: 10.1038/s41467-020-17142-3.
- Zanna, Laure and Thomas Bolton (2020). “Data-driven equation discovery of ocean mesoscale closures”. In: *Geophysical Research Letters* 47, e2020GL088376. ISSN: 0094-8276. DOI: 10.1029/2020GL088376.
- Zelinka, Mark D, Chen Zhou, and Stephen A Klein (2016). “Insights from a refined decomposition of cloud feedbacks”. In: *Geophysical Research Letters* 43 (17), pp. 9259–9269. ISSN: 0094-8276. DOI: 10.1002/2016GL069917.
- Zelinka, Mark D et al. (2017). “Clearing clouds of uncertainty”. In: *Nature Climate Change* 7 (10), pp. 674–678. ISSN: 1758-6798. DOI: 10.1038/nclimate3402.
- Zelinka, Mark D et al. (2020). “Causes of higher climate sensitivity in CMIP6 models”. In: *Geophysical Research Letters* 47 (1), e2019GL085782. ISSN: 0094-8276. DOI: 10.1029/2019GL085782.
- Zhang, G J and Norman A McFarlane (1995). “Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian climate centre general circulation model”. In: *Atmosphere-Ocean* 33 (3), pp. 407–446. ISSN: 0705-5900. DOI: 10.1080/07055900.1995.9649539.
- Zhang, Huai-Min, JH Lawrimore, Boyin Huang, MJ Menne, Xungang Yin, A Sanchez-Lugo, BE Gleason, Russell Vose, Derek Arndt, JJ Rennie, et al. (2019). “Updated temperature data give a sharper view of climate trends”. In: *Eos* 100, p. 10. DOI: 10.1029/2019E0128229.
- Zhang, Minghua et al. (2012). “The CGILS experimental design to investigate low cloud feedbacks in general circulation models by using single-column and large-eddy simulation models”. In: *Journal of Advances in Modeling Earth Systems* 4 (4). <https://doi.org/10.1029/2012MS000182>. ISSN: 1942-2466. DOI: 10.1029/2012MS000182.

- Zhang, Minghua et al. (2013). “CGILS: Results from the first phase of an international project to understand the physical mechanisms of low cloud feedbacks in single column models”. In: *Journal of Advances in Modeling Earth Systems* 5, pp. 826–842. ISSN: 1942-2466. DOI: 10.1002/2013MS000246.
- Zhao, Ming (2014). “An investigation of the connections among convection, clouds, and climate sensitivity in a global climate model”. In: *Journal of Climate* 27 (5), pp. 1845–1862. DOI: 10.1175/JCLI-D-13-00145.1.
- Zhao, Wen Li et al. (2019). “Physics-constrained machine learning of evapotranspiration”. In: *Geophysical Research Letters* 46, pp. 14496–14507. ISSN: 0094-8276. DOI: 10.1029/2019GL085291.