UNIVERSITY OF READING

Department of Meteorology

Improving Turbulent Representation of Shallow Cumulus Convection in the Grey-Zone.

ALANNA POWER

A thesis submitted for the degree of Doctor of Philosophy.

October 2024

Declaration:

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Alanna Power

Acknowledgements

Firstly, I would like to sincerely thank Bob, Peter, and George for guiding my PhD project over the last four years. All three of you have always gone above and beyond to help with any challenges which came up throughout this PhD, and I am so grateful to have had you as my supervisors. Thanks for sticking with me through my delayed timelines, my cobbled-together code that was going to run whether it liked it or not, and my unclear explanations that started in the middle and ended at the beginning. You all have been exceptionally helpful, kind, and knowledgeable (not to mention very patient!) and I couldn't have completed this PhD without you, for as we all know "In LES nobody can hear you scream - because it's anelastic". Thank you to Todd, who was involved in this PhD for the first two years and helped me navigate the complex world of the Jasmin supercomputer. I would like to thank my monitoring committee, Andy and Omduth, for all your advice and suggestions over the last 4 years; it has been incredibly useful to get an outsider's perspective on the project. Thank you to my examiners, Steef and Omduth, for taking the time to read this thesis, and for the interesting discussion that we had about this work during the viva. Thanks to Wendy and Imogen for being so helpful and friendly whenever I asked for help with everything and anything that popped up during this PhD.

I owe a huge thank you to my parents for their never-ending support, and for predicting from a young age that I would be "the eternal student" - you were right. To Aidan, I can't thank you enough for everything you do. Thanks for keeping me sane, despite having your own PhD woes, and for allowing me to use your office as my own whenever I visit Edinburgh. I'm forever grateful to Ahmad, Moe, Marimel, and Jiheun, who made sure I had a roof over my head and food in my belly during the final stages of the write-up. I could not have gotten through the last few months of this PhD without each of you, and I am so lucky to have you all - you basically carried me to the finish line. Your kindness and generosity are going to be hard to repay, but I will try! To Neri, the Brighton Road housemate who escaped Reading, thanks for always coming back for walks, chats, and sibling-like bickering. Thank you to my housemate Charlie, and of course the honorary housemate Natalie, for sharing this PhD journey, however chaotic, both in the office and at home - I don't think our sentence structures will ever be fully coherent again. A big thanks to Brian, Izzy, Blair, Hannah, Laura, Beth, Meryl, Gabriel, and James for making sure that working on this PhD was much more fun than it had any right to be. To all the student visitors we had the luck of hosting in 2U06, thanks for sharing your stories, coding tips, homemade food, UNIX hacks, and drinking game traditions with us. We probably would have had a slightly more productive, and infinitely more boring office without you all! A huge thank you to everyone in the Met Department, who provided support, feedback, debugging skills, as well as lots of fun times too. Lastly, thank you to the cats, Schrodinger and Peach, who kept me company during the 3am writing sprees - you were probably more of a hindrance than a help.

Abstract

Shallow cumulus clouds are precursors to deep convection, which can lead to extreme weather events that are difficult to forecast. Therefore, to more accurately predict these events, the preceding environment must first be accurately forecast. Advances in computational power have enabled numerical weather prediction models to run with kilometre and sub-kilometer scale grids. Such models can only partially resolve the dominant turbulent structures within the flow, and therefore the unresolved turbulence still needs to be parametrized. Traditional parametrization schemes for shallow cumulus convection are not valid at these grid scales, as they rely on the assumption that all turbulent motions are entirely subgrid scale. This gives rise to the grey-zone regime, which occurs when the scales of the dominant turbulent motions are comparable to the grid length of the model. The aim of this work is to improve the model's ability to capture the effects, at coarse resolution, of the turbulent structures which transport heat, moisture, and momentum to cumulus clouds, thereby mitigating the impact of the grey-zone.

The primary objective of this research is to develop a parametrization that enables the extension of large eddy simulation (LES) abilities to coarser grid scales, traditionally categorised as grey-zone scales, while maintaining accuracy and low computational costs. The Met Office/NERC Cloud (MONC) LES model was used to produce high-resolution fields for three case studies: an idealised dry convective boundary layer case, the Barbados Oceanographic and Meteorological EXperiment (BOMEX) case, and the Atmospheric Radiation Measurement (ARM) case. The dynamic Smagorinsky equations were applied to the resulting fields to produce flow-dependent Smagorinsky parameters. These parameters define the mixing length in the model. By analysing their behaviour, significant variations in turbulent mixing have been identified between the mixed layer, cloud-free environment, and in-cloud regions of the cloud-topped boundary layer (CTBL). The findings reveal that the Smagorinsky parameters for momentum, heat, and moisture are significantly influenced by both the flow regime and filter scale. These dependencies are not accounted for in the standard model. A scale-adaptive relationship between height and the Smagorinsky parameters could then be derived for each variable.

A novel parametrization scheme has been developed using these relationships to serve as a grey-zone adaption, enabling the model to capture the key flow dependencies without incurring high computational expense. This aims to deliver the benefits of a dynamic Smagorinsky method but negates the need to compute the parameters at each grid point and time step. The MONC model was modified to include this parametrization in the subgrid scheme, and the resulting grey-zone simulations demonstrated substantial improvements, particularly in terms of cloud initiation time and cloud layer growth. Furthermore, the research underscores the importance of recognizing the variations in turbulent mixing lengths, as the dynamics of the turbulent flow in the

CTBL are inherently linked to the grid scale, stability, and flow regimes. This parametrization addresses the limitations of using fixed parameter values when modelling convective turbulence in a CTBL. This research presents a significant step forward in addressing the challenges posed by the grey-zone in modelling shallow cumulus convection. Valuable insights into the dynamics of turbulent mixing in the CTBL offer a promising framework for advancing the capabilities of LES models in the grey-zone regime.

Table of Contents

List of Abbreviations

ARM	Atmospheric Radiation Measurement
BOMEX	Barbados Oceanographic and Meteorological Experiment
CBL	Convective boundary layer
CC	Cloud core
CFE	Cloud-free environment (encompasses the ML and NC regions)
CL	Cloud layer
CTBL	Cloud-topped boundary layer
CU	Cloud updraft
DNS	Direct numerical simulation
EUREC^{4}A	Elucidating the Role of Clouds–Circulation Coupling in Climate
FT	Free troposphere
GCM	Global climate model
IC	In-cloud
ISR	Inertial subrange
LCL	Lifting condensation level
LES	Large eddy simulation
LFC	Level of free convection
LNB	Level of neutral buoyancy
ML	Mixed layer
MONC	Met Office/NERC cloud (LES model)
NC	Non-cloudy (area of the CL)
NWP	Numerical weather prediction
SGS	Subgrid scale
ShCu	Shallow cumulus
STS	Subtest scale
TKE	Turbulent kinetic energy

List of Symbols

z_i	Capping inversion height
z_{ML}	Height of the mixed layer top
z_{CB}	Cloud base height
z_{CT}	Cloud top height
θ	Potential temperature
θ_e	Equivalent potential temperature
θ_L	Liquid water potential temperature
θ_v	Virtual potential temperature
q_t	Total water content
q_v	Water vapour
Pr	Prandtl number
Sc	Schmidt number
ℓ_{mix}	Mixing length (for momentum)
ℓ_ψ	Mixing length for any scalar ψ
ℓ_d	Dissipation length scale
ν_{m}	Kinematic eddy viscosity
ν_ψ	Diffusivity of any scalar ψ
ν_{h}	Thermal diffusivity
C_s	Smagorinsky parameter for momentum
C_ψ	Smagorinsky parameter for any scalar ψ
C_θ	Smagorinsky parameter for heat
C_{θ_L}	Smagorinsky parameter for θ_L
C_{q_t}	Smagorinsky parameter for total moisture
R_{eff}	Effective resolution
Δ_{eff}	Effective grid scale
Δ	Grid spacing of the Model
$\overline{\Delta}$	Effective grid scale of data which has been filtered once
$\widehat{\overline{\Delta}}$	Effective grid scale of data which has been filtered twice
σ	Filter scale of the Gaussian filter
σ_{smag}	Filter scale of the Smagorinsky scheme
τ_{ij}	Stress tensor
S_{ij}	Rate of strain tensor
L_{ij}	Difference in stress between two scales
M_{ij}	Difference in strain between two scales
ϵ	dissipation
Ri	Richardson number
Ri_{c}	Critical Richardson number
$\mathcal{S}_{\text{w}}$	Power spectrum of the vertical velocity
$\mathcal{S}_e$	Power spectrum of the TKE
$\mathcal{Q}^*$	The constant surface buoyancy flux
w^*	The free-convection scaling velocity

1 Introduction

The objective of this project is to improve simulations of cloud-topped boundary layers (CTBL) at grey-zone resolutions. In the grey-zone regime, the model cannot fully resolve the turbulent eddies in the flow, causing the structures which transport momentum, heat, and moisture to be only partially resolved. The grey-zone was first defined by Wyngaard (2004) as occurring when the grid spacing of the model is of the same order as the length scales of the dominant coherent structures in the flow. However, the effects of the grey-zone regime begin when grid spacings are finer than the dominant scales of the flow. This is partly because dissipation of energy due to the model dynamics exacerbates the grey-zone issue, and partially because the scales of motion vary with location and can have very small scales, particularly at boundaries, inversions, and transitions. The overarching aim of this project is to impede excessive energy dissipation at coarse resolution and delay the onset of the grey-zone by extending the capabilities of the large eddy simulation (LES) method to larger grid spacings. A dynamic Smagorinsky method is applied to high-resolution fields of CTBL data via the Germano identity. This enables the computation of flow-dependent parameters, and is used to calculate fields of the Smagorinsky parameter for momentum, heat, and moisture in the CTBL. These parameters govern the dissipation of quantities in the simulation. It is hypothesised that, by analysing the dynamically computed parameter fields, relationships between the flow and the Smagorinsky parameter can be found and parametrized, thus allowing for the development of a model which more accurately dissipates energy at grey-zone resolutions.

This study uses the Met Office/NERC Cloud (MONC) LES model to produce high-resolution output fields of the CTBL. MONC uses the Smagorinsky subgrid scheme, which prevents a buildup of energy at small scales by imposing an eddy viscosity on the flow. In the standard Smagorinsky model this viscosity is a function of the Smagorinsky parameter, which is set to a constant. In order to investigate if the viscosity being prescribed by the model is too large, and thus compounding the energy dissipation problem in the grey-zone, flow and scale-dependent parameters are calculated. Dynamic methods (outlined in Section 3) are used to calculate these flow and scale-dependent parameters for momentum, heat, and moisture individually, unlike the standard model which fixes all scalars to the same value parameter. The standard model also forces the diffusion of scalar parameters to depend on their momentum counterpart. The Smagorinsky parameter for scalars is determined using the equivalent momentum parameter and a Prandtl number, which is used as the constant of proportionality (see Equation 3.28). One aim of this project is to investigate if allowing parameters to be scalar-dependent is important, rather than relying on the current approach of prescribing them with fixed values and momentum dependencies. This project focuses on computing flow-dependent mixing lengths for not only momentum, as in most of the previous literature, but also for the heat and total moisture scalars. The output fields of each of these mixing lengths are analysed, with particular focus on values near and within clouds and thermals, as well as at cloud base, cloud top, and temperature inversions.

The dominant scales of the overturning circulation in a CTBL are determined by the mixed layer (ML) inversion height and cloud layer (CL) depth, with $\mathcal{O}(\lambda_{ML}) \approx \mathcal{O}(\lambda_{CL}) \approx 1\,\text{km}$. Three well-studied cases of convection within the cloud-topped boundary layer have been studied, an idealised dry convective boundary layer (CBL) case, the Barbados Oceanographic and Meteorological Experiment (BOMEX) case and the Atmospheric Radiation Measurement (ARM) case (based in the Southern Great Plains, Oklahoma). The BOMEX case is of a marine boundary layer (MBL), quasi-steady and driven by latent heat fluxes. In contrast, the ARM case is based on a diurnal cycle over arid, flat land, driven by sensible heat fluxes and demonstrating many transitions: from the shallow morning-time boundary layer, to a convective "dry" boundary layer, to one with the onset of clouds, moving to a more rapid cloud development stage, and finally reaching a quasi-steady like stage in the evening. These cases allow for the effects of time dependence, regime and stability dependence, filter scale dependence, and scalar dependence of the Smagorinsky parameters to be investigated. The Smagorinsky parameter is strongly influenced by features of the CBL such as stability changes (eg: at the mixed layer inversion height, cloud top height), and position relative to clouds and thermals. The cloudy regions are further partitioned into 3 subsets: in-cloud, cloud updraft, and cloud core (See Section 7 for definitions) to investigate the effect that motions within the cloud can have on the Smagorinsky parameters. The fields are filtered to various resolutions, with filter scales ranging from 2Δ to 128Δ being analysed (See Section 4.4 for filter scale definitions), to investigate how these features interact with the Smagorinsky parameter as resolution decreases toward grey-zone scales. The effects of the CTBL features and changes in grid-scale on the Smagorinsky parameter values are quantified, allowing relationships between the parameters and the flow/filter scale to be devised.

The relationships between the Smagorinsky parameters and (a) the flow regime, and (b) the filter-scale of a model can then be used to devise a new parametrization scheme. In theory this would allow the model to capture the effect that these features have on the flow, without the computational expense associated with dynamically calculating the parameters. The hope is that, with the new parametrization scheme, the model will be better able to adapt to the challenges posed by the grey-zone regime. To test this hypothesis, the MONC model is altered to include the new parametrization scheme. Multiple different configurations of the new parametrization scheme will be used to run MONC simulations at coarse, grey-zone resolutions. The outputs from the updated model can be compared to high-resolution LES fields from the standard model. Furthermore, MONC simulations at grey-zone scales are also used as a comparison to determine the effect that the new parametrization scheme has on CTBL simulations.

1.1 Aims of this Thesis

This project aims to answer the following research questions:

- Does the Smagorinsky scheme exhibit a specific filter shape that scales uniformly with grid spacing, and can its filter scale be derived? How does the Smagorinsky filter compare to the Gaussian filter? How can the Gaussian-filtered data be related back to a grid-based scale?
- Do the Smagorinsky parameters exhibit systematic patterns based on flow regimes, and if so, are there significant differences in the transport and diffusion characteristics of momentum, heat, and moisture within each of these regimes?
- What is the impact of filter scale on the various Smagorinsky parameters within the three flow regimes of interest?
- Can a scale-adaptive relationship be established between height and the Smagorinsky parameters within each flow regime? Does this empirical parametrization improve LES of CTBLs in the grey-zone?

1.2 Thesis Outline and Structure

Throughout this work, the dynamic Smagorinsky equations and the Germano method were employed to produce flow-dependent fields of mixing length parameters for CTBLs. This data has been analysed and it has been demonstrated that the mixing length scales are different between distinct parts of the flow. Furthermore, the value of these length scales have been found to depend on the variable in question, be it momentum, heat, or moisture. The length scale parameters exhibit a strong dependency on the filter scale; Parameter values decrease, indicating the model's diminishing ability to resolve the flow, as the filter scale increases and resolution coarsens. These dependencies have been amalgamated to derive a new parametrization for the Smagorinsky parameter in the subgrid scheme. This parametrization has been implemented in the MONC model, resulting in improvements in cloud initiation time, and overall evolution of the cloud layer.

The work in this thesis is structured as follows. This thesis begins with Chapter 2 presenting an overview of previous literature which is important for this work. This chapter provides foundational knowledge of the atmospheric boundary layer theory, particularly in relation to shallow cumulus convection. Key concepts are introduced relating to transport, diffusion, and the challenges posed by the grey-zone in turbulence modelling. Chapter 3 discusses the Smagorinsky subgrid scheme in detail, with the defining equations of this model being presented there.

Chapter 4 focuses on the case studies and model set-up, used throughout this investigation. The basics of the MONC model, the configurations for various test cases (e.g., the dry CBL, BOMEX, ARM), and data post-processing methods are detailed in this chapter. The method of filtering LES outputs is also explained, and the offline dynamic model which is used to compute Smagorinsky parameters is introduced. Chapter 5 explores whether the Smagorinsky scheme behaves as a filter, comparing its behaviour with Gaussian filters in spectral space. This chapter provides a detailed comparison between the two filters. Chapter 6 investigates how turbulent mixing differs between the distinct flow regimes within the CTBL. This chapter also examines the effects of partitioning the high-resolution fields of the CTBL into the mixed layer, cloud-free, and in cloud regions. Chapter 7 investigates the effect of filter scale on the Smagorinsky parameter values. It evaluates the impact of filter scale on cloud cover and discusses trends and responses of momentum and scalar parameters to these changes. Chapter 8 combines all the work from the previous chapters to derive a new mixing length parametrization. This parametrization allows the Smagorinsky parameter to account for systematic dependencies observed in previous chapters. The variations in Smagorinsky parameters across different regions are analysed, and experiments using different configurations of the model are used to test the new parameterization. Finally, Chapter 9 concludes by summarizing key findings in this thesis, discussing the broader implications of the research, addressing limitations, and offering directions for future work before closing with final remarks.

2 Background Theory and Literature Review

2.1 The Boundary Layer

The shallow cumulus (ShCu) boundary layer consists of a convective cloud field of cumulus cloud which forms on top of the CBL, a layer which is dominated by large convective updrafts (Houze, 2014). While precipitation can be common in tradewind ShCu convection, the cloud-topped boundary layers focused on in this study are limited to those with non-precipitating cumulus. The CBL is a well mixed layer where the circulation is driven by convective thermal updrafts that exist across the whole domain. These thermals originate at the surface and extend in the vertical, until reaching the top of the mixed layer (Garratt, 1994). This turbulent circulation is a very important transport process (Stull, 2009), with these thermals carrying momentum, moisture, heat, and other quantities throughout the depth of the CBL. The largest thermals are known as the dominant coherent structures of the flow, and these are the energy-containing eddies which scale to the depth of the CBL z_i.

CTBL commonly occur when both sensible and latent heat are provided to the system from the surface, giving rise to a conditionally unstable layer. In such cases air rises until meeting the lifting condensation level (LCL), at which point the ascending air parcel becomes saturated and the moisture within the parcel begins to condense and release latent heat. The high levels of condensation occurring at the cloud base are then balanced by turbulent mixing, which allows the heat generated to be transported up throughout the cloud (Siebesma and Cuijpers, 1995). Cumulus clouds are low level clouds, with Orlanski (1975) stating that their cloud base is usually located in the bottom 2 km of the atmosphere. ShCu are identified as individual clouds, with both their vertical and horizontal scales of the order of 1 km (Houze, 2014). The majority of these clouds stay this size for the entirety of their lifetime, provided they remain isolated from one another. Though shallow cumuli are small in scale, they have a substantial impact on the climate system due to their prevalence.

2.1.1 Marine Boundary Layer

Convective turbulence in the MBLs are mostly driven by latent heat fluxes, with fluxes of sensible heat being less influential. Often times MBLs are topped by a cloud layer, with a significant portion of the surface evaporation driving the moisture flux near the cloud base (Garratt, 1994). As the MBL top does not always coincide with a temperature inversion or cloud top, the MBL is defined as the part of the atmosphere which is directly coupled to the ocean's surface by turbulent transfer of quantities such as moisture, momentum, and heat (Garratt, 1994). Due to the large heat capacity of the ocean, MBLs do not experience significant changes during the diurnal cycle and so are often considered as being in quasi-steady equilibrium. This is notable as Siebesma and Cuijpers (1995) states that it is common for diagnostic cumulus parametrization schemes to assume steady-state conditions.

Due to the typical atmospheric conditions over oceans, fields of small cumulus clouds are a common feature above MBLs. Johnson (1976) states that, across the tropics, shallow cumuli are the most abundant cloud type, while in the sub-tropics Norris (1999) notes that ShCu convection underlies and sustains a substantial amount of the stratocumulus fields. As a consequence of their prevalence, ShCu clouds are of great importance to the global climate. These clouds have a notable impact on the Earth's radiative energy budget, both at the surface and at the top of the cloud layer (Neggers et al., 2003; Bretherton et al., 2004). CTBLs are thought to have a large influence over the exchange of latent and sensible heat at the the ocean–atmosphere boundary (Neggers et al., 2003), and trade wind cumuli are known to supply this heat energy to the Hadley circulation, thus intensifying the large scale global circulation (Siebesma and Cuijpers, 1995). ShCu clouds also have an effect on key characteristics of the boundary layer. The boundary layer depth, temperature, relative humidity, and winds are altered by this convection, as air is vented from the mixed layer to the free troposphere (Bretherton et al., 2004). ShCu convection thereby has a large influence over the vertical transport of momentum, heat, and moisture (Neggers et al., 2003).

2.1.2 Shallow Cumulus Convection

CTBL occur in areas where both sensible and latent heat are input at the surface, as these conditions give rise to a conditionally unstable layer. ShCu boundary layers (BL) consist of a CL, which is comprised of a field of non-precipitating cumulus clouds that have formed on top of the ML (Houze, 2014). The circulation in the BL is an important transport process as the thermals carry momentum, moisture, heat, and other quantities throughout the depth of the CBL (Stull, 2009). The top of the mixed layer, z_{ML}, is defined as the height at which turbulence decouples from the surface forcings. The case studies in this work show z_{ML} aligning with both the cloud base height z_{cb}, and the temperature inversion level z_i. While these three levels are often close, they do not always correspond to the same height levels. The height levels which are relevant to the cases analysed in this work are discussed further in Section 6.2.1.

The cloud base in CBLs can be used as a proxy for the lifting condensation level (LCL). The LCL is the point at which moisture in ascending air parcels begins to condense out and release latent heat resulting in the formation of clouds. The height level at which this latent heat release causes the parcel to be warmer than its surrounding environment, and therefore positively buoyant, is known as the level of free convection (LFC). ShCu clouds usually form on top of thermals as they supply heat, moisture, and energy to the growing cloud. Bretherton et al. (2004) states that turbulent mixing processes in both the mixed layer and the near-cloud environment can have a strong effect on cloud formation and growth. The large amount of condensation occurring at the cloud base results in high levels of mixing, as the latent heat release drives the circulation within the cloud, resulting in the transport of buoyant energy throughout the cloud (Siebesma and Cuijpers, 1995). The main path taken by the positively buoyant ascending air forms the cloud core region. Cloud cores often influence the surrounding air and drag it upwards alongside the main channel of ascending air. This causes an area of positive vertical velocity, despite not being positively buoyant, known as the cloud updraft. These upward motions continue rising until the moisture has condensed out or been detrained into the environment and the parcel becomes the same temperature as the surrounding non-cloudy environment. The height at which this occurs is known as the level of neutral buoyancy (LNB), and it can be used as a proxy for the average cloud top height, z_{CT}. Some rising air parcels will overshoot the LNB, before becoming negatively buoyant and sinking. This overshooting serves to grow the cloud layer by raising the average cloud top height, while also entraining warm, dry air from the overlying air into the CL, resulting in high levels of mixing at the cloud top.

Neggers et al. (2003) discuss how sensitivity studies have shown ShCu convection to have a significant impact on the resolved climate in global climate models (GCM). Despite the pronounced role of ShCu in global weather and climate, as well as their prevalence, ShCu convection is not well represented in numerical models (Siebesma and Cuijpers, 1995). The underlying assumptions made for convection parametrizations in NWP models no longer hold when models move to higher resolutions and into grey-zone regimes. This is because the energy-containing convective structures become the same scale as the model grid, and as a result, these motions are under-resolved and grid-scale dependent, rendering the closure methods used in conventional parametrization schemes invalid (Sakradzija et al., 2016). Despite these known issues, turbulence closures for conventional NWP schemes continue to be used for models in the grey-zone due to the lack of alternative parametrization schemes. Neggers et al. (2003) comments that problems still arise from using cumulus convection schemes in GCMs when representing aspects of ShCu topped boundary layers. A schematic of the stages of the life cycle of shallow cumuli from Cuijpers (1994) is presented in Figure 2.1. This schematic represents the growth of cumulus clouds in a MBL over time, however the lifecycle of clouds over land during the diurnal cycle is similar.

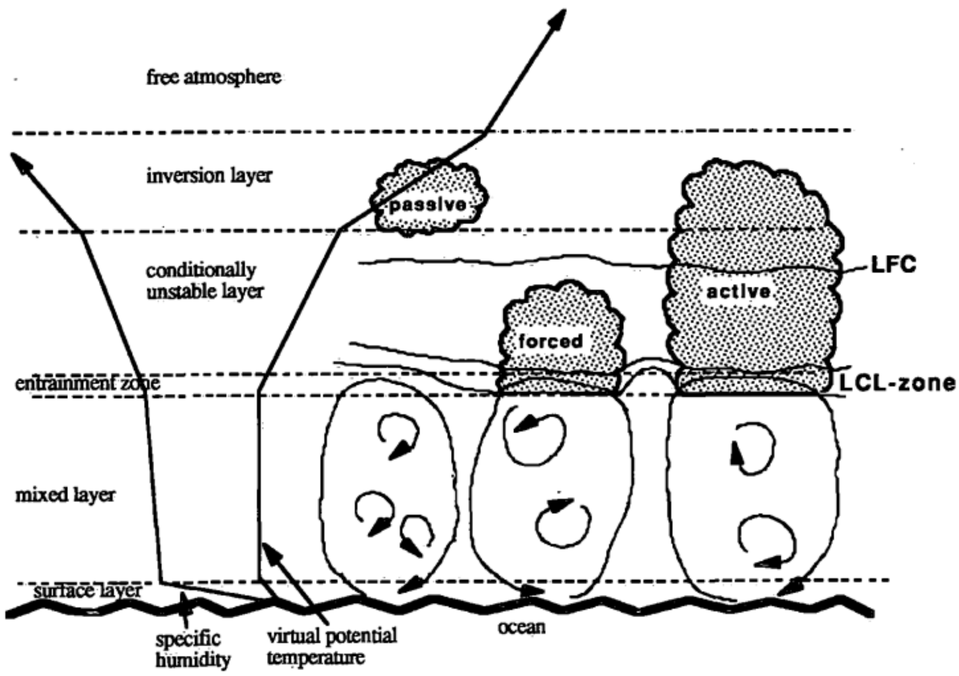

Figure 2.1: *Lifecycle of ShCu clouds in a MBL from Cuijpers (1994).*

2.1.3 Climate Impacts of Shallow Cumulus Clouds

ShCu clouds typically form over MBLs, with this study focusing on non-precipitating ShCu. These clouds play an influential role in the Earth's climate system due to their spatial extent across the globe. The radiative properties of ShCu, particularly their capacity to reflect incoming shortwave radiation, causing a net cooling effect, mean that they have a large influence on the surface energy budget. However, ShCu clouds remain a major source of uncertainty in climate projections, largely due to the difficulty in representing their formation, evolution, and radiative impact within GCMs. Work by Bony and Dufresne (2005) demonstrated that variations in low-cloud cover across climate models account for a significant portion of the spread in equilibrium climate sensitivity, highlighting the importance of identifying and quantifying the feedbacks associated with ShCu. This was one of the motivations for the EUREC^{4}A campaign: Elucidating the Role of Clouds–Circulation Coupling in Climate. Bony et al. (2017) states that the aim of this field campaign was to quantify the physical properties of trade-wind ShCu, such as the water content and cloud fraction as a function of the large-scale environment. If these properties can be correctly accounted for in models, the hope is that it will improve the interaction between ShCu and the large-scale system in GCMs.

Studies have increasingly focused on the dynamical coupling between ShCu and their surrounding environment. Nuijens et al. (2015) examined the sensitivities of trade-wind cumulus to large-scale processes, comparing model data to observations from satellite and aircraft data.

This study concluded that many models were unable to capture the processes which affect the cloud fraction and lifetime of low-level clouds, thereby impairing the model's ability to predict future climate scenarios. Other work has been carried out to investigate the interaction between ShCu convection and climate variability. Brient and Schneider (2016) investigated cloud feedback sensitivity to sea surface temperature anomalies and found that the vertical velocity of the shallow convection, along with the relative humidity, influences the shortwave cloud radiative effects. These are all mechanisms which are sensitive to climate change, and thus understanding their influence and interaction on ShCu is important if it is to be captured correctly in GCMs.

In summary, the role of ShCu in the climate system remains a focal point in climate science due to the sensitivity of ShCu to environmental changes, along with ShCu's role in influencing the Earth's radiation budget. As the climate system continues to warm, reducing uncertainties associated with ShCu convection is essential for narrowing projections of future climate change. While there have been advances in the understanding of cloud–circulation coupling and feedbacks in recent years, persistent biases in the cloud representation within models highlights the need for further research, improved high-resolution modelling, and new parametrization schemes.

2.2 Transport and Diffusion in the Convective Boundary Layer

In atmospheric modelling, turbulent transport of heat and moisture plays a crucial role in representing processes within the atmospheric boundary layer. LES models often impose similar mechanisms of transport for heat and other scalars, such as moisture and pollutants, within the turbulent flow despite differences in how they interact with the system. In such models, potential temperature (θ) and water vapour (q_v) are recognised as active scalars as they interact with the modelled flow field by affecting the density field. Meanwhile, other scalars such as pollutants are passive scalars as they do not have a direct feedback on the simulated flow. The assumption that all of these scalars are being transported in the same way leads to the Schmidt number, which defines the ratio of momentum diffusivity to mass diffusivity, being assumed identical to the Prandtl number for heat in atmospheric models (Li, 2019). The Prandtl number is similar to the Schmidt number in that it describes the ratio of momentum diffusivity to heat diffusivity. Both of these quantities are discussed further in Section 2.2.2 below.

2.2.1 Transport of Heat

In terms of heat transport and diffusion, different forms of potential temperature can be employed to better capture various processes. The potential temperature θ, relates to the dry atmosphere, while virtual potential temperature (θ_v) recognises density differences between water vapour and air, and is therefore a useful proxy for buoyancy. However both these thermodynamic variables are not cloud conserved. The liquid water potential temperature (θ_L), introduced by Betts (1973), accounts for the presence of liquid water in non-precipitating clouds. The equivalent potential temperature (θ_e) represents the total energy, combining sensible heat, latent heat, and moisture. The θ_L and θ_e quantities are conserved within clouds. The equations for these variables are grouped according to their conservation properties in Table 2.1.

Non-conserved Heat Variables	Conserved Heat Variables
Potential temperature: $\theta = T\left(\frac{p}{p_0}\right)^{-\frac{R_d}{c_{pd}}}$	Equivalent potential temperature: $\theta_e \approx \left(T\frac{L_v}{c_{pd}}r_v\right)\left(\frac{p}{p_0}\right)^{-\frac{R_d}{c_{pd}}}$
Virtual potential temperature: $\theta_v \approx \theta\,(1 - 0.61r_v - r_L)$	Liquid water potential temperature: $\theta_L \approx \theta - \left(\frac{L_v}{C_{pd}}\right) r_L$

Table 2.1: *Formulas to calculate the various heat variables, grouped according to their conservation properties.*

In Table 2.1, T is the temperature at the height level of interest, p is the pressure at this level, while p_0 is the reference pressure, usually $p_0 = 1000\,$hPa. The gas constant for dry air is given by $R_d = 287.04$, while the specific heat capacity for dry air is $c_{pd} \approx 1003.5$. L_v is the latent heat of vaporization, r_v is the mixing ratio of water vapour in the air, while r_L is the mixing ratio of liquid water in the air.

Efstathiou et al. (2024) emphasized the role of heat transport in clouds, particularly during the shallow cloud stage. This study shows significant positive fluxes in the cloud layer, indicating strong counter-gradient heat transport. In their simulations, LES data with a grid spacing of Δx $= 50\,$m suggests that turbulent Prandtl numbers tend to zero in clouds. This highlights the failure of the Smagorinsky model as it is unable to account for upscale transport, and instead resorts to shutting off heat dissipation. Similar findings by Shi et al. (2018) show that the subgrid scale (SGS) fluxes of heat and moisture behave differently within cloud layers, necessitating scalar-specific treatment for accurate representation of turbulent transport.

2.2.2 Prandtl and Schmidt Numbers

The Prandtl and Schmidt numbers vary within individual flows, suggesting that constant values used in traditional models do not adequately capture the nuances of scalar transport. This is particularly important when considering SGS scalar diffusivities. A study by Shi et al. (2018)

shows that dynamic models which compute independent SGS eddy diffusivities for scalar mixing, perform better compared to models that rely on constant Prandtl and Schmidt numbers. Specifically, this work indicates that moisture transport differs from heat transport, especially near the cloud base, where the eddy diffusivity for heat shows a sharp transition, with large values below the cloud and suppressed mixing within the cloud layer. However, the eddy diffusivity for moisture does not show this tendency to reduce mixing in-cloud. The majority of the CL exhibits values comparable to the ML, except near the cloud top, where a reduction in moisture mixing is observed.

The Prandtl number is an intrinsic property of the fluid. It describes the ratio of kinematic eddy viscosity ν_m to heat diffusivity ν_θ, essentially determining the degree of similarity between the transport of heat and momentum (Li, 2019):

$$\text{Pr} = \frac{\nu_m}{\nu_\theta} \tag{2.1}$$

A small Prandtl number value of less than one means that the heat diffusivity dominates the flow, with large thermal structures transporting heat and smaller structures transporting momentum. A Prandtl number of one suggests that the structures transporting heat are of the same size as those transporting momentum, and there is identical diffusion between these two quantities. Meanwhile, Prandtl number values larger than one indicate that momentum diffusivity is the dominant behaviour within the flow. For scalars other than heat, this ratio can be made more general by using the Prandtl number's counterpart, the Schmidt number Sc. The Schmidt number is the ratio of ν_m, the momentum diffusivity, to ν_ψ, where ν_ψ is the mass diffusivity of a given scalar ψ in the flow.

$$\text{Sc}_\psi = \frac{\nu_m}{\nu_\psi} \tag{2.2}$$

The Schmidt number for any scalar Sc_ψ is a property of both the fluid and the scalar being diffused, and it is of the order one in the atmospheric boundary layer (Gualtieri et al., 2017).

2.3 Modelling the Cloud-Topped Boundary Layer

Neggers et al. (2003) discuss how sensitivity studies have shown ShCu convection to have a significant impact on the resolved climate in GCMs. Despite this, along with the prevalence of ShCu clouds and their pronounced role in global weather and climate, ShCu convection is not well represented in numerical models (Siebesma and Cuijpers, 1995). The underlying assumptions made for convection parametrizations in numerical weather prediction (NWP) models no longer hold as models move to higher resolutions and into grey-zone regimes. This is because the energy containing convective structures that emerge on the model grid are under-resolved and grid-scale dependent, rendering the closure methods used in conventional parametrization schemes invalid (Sakradzija et al., 2016). Neggers et al. (2003) comments that problems still arise from using cumulus convection schemes in GCMs when representing aspects of ShCu topped boundary layers.

The grey-zone poses a significant problem when modelling ShCu convection. Coarse resolution LES struggle to accurately resolve cloud size and initiation time. Analysis of LES outputs from CBL simulations show that the length scales of the resolved flow vary widely in the grey-zone regime (this is further discussed in Chapters 6 and 7). This is notable as the mixing length for momentum and scalars is often set to a constant in LES models, meaning that the model does not allow for variations in these length scales. The Smagorinsky scheme is a common subgrid scheme used in LES models (further details given in Section 3). This scheme assumes that scalars such as heat and moisture are being transported and dissipated in the same way within the simulation, and so are set to the same mixing length value. This is an important feature, as the dissipation of energy is controlled by the viscosity imposed on the flow by the subgrid scheme, and this viscosity is a function of mixing length. The mixing length for momentum $\ell_{\rm mix}$, is assumed proportional to the grid spacing. Meanwhile l_ψ, the mixing length for a scalar ψ, is a simple function of $\ell_{\rm mix}$ and the Prandtl number for that scalar Pr_ψ. The Prandtl number is a ratio between kinematic eddy viscosity $\nu_{\rm m}$ to scalar diffusivity ν_ψ, and it is often set to a constant (typically using $\mathrm{Pr}_{air} \approx 0.7$) in models.

Many LES models rely on the assumption that energy production is balanced by dissipation. As a result, these models assume that turbulence is isotropic at the scales where the viscosity is causing the dissipation of energy to occur. It is therefore necessary for the model to resolve turbulent motions down to scales which are well within the inertial subrange (ISR). However, in the grey-zone regime, the dominant coherent structures are approximately the same scale as the grid spacing. This means the dominant coherent structures, which are the energy-containing eddies, are the smallest scale eddies being resolved/partially resolved. It is important to note that in a CBL these dominant structures are driven by buoyancy and thus are not isotropic. As a result, in grey-zone regimes, the simulation does not resolve down to the scales small enough for turbulence to be considered isotropic, meaning that the dominant coherent structures experience unrealistic levels of energy dissipation. Therefore the current subgrid dissipation schemes are not a sufficient treatment for turbulence in the grey-zone.

In LES models that use the Smagorinsky SGS, the extent of turbulent mixing of momentum is determined by the Smagorinsky parameter C_s and the model grid spacing Δ. As a result the standard Smagorinsky models, which set C_s to a constant value, are less able to accurately resolve turbulent fields in the presence of boundaries or transitioning flows. Both of these flow regimes are prominent features of the boundary layer. The dynamic Smagorinsky model allows simulations to be run with C_s being dynamically computed at each point based on the flow. Dynamic models show better agreement with direct numerical simulation (DNS) data, however, this method requires substantial computational power.

The standard Smagorinsky model assumes that scalars such as moisture and heat are transported in the same way, while also assuming that scalar parameters are determined by the Smagorinsky parameter for momentum, C_s. This assumption is based on the premise that it is the same thermals which are transporting these quantities. In the MONC model, scalar mixing is defined as a function of momentum mixing and stability functions, which encompass the Prandtl number (See Section 3.1.3 for equations).

2.3.1 Large Eddy Simulations

LES models are often used in the development and testing of parametrisation schemes for numerical weather prediction and climate models. They can produce high resolution fields of the turbulent motions, such as those associated with ShCu convection. Dairay et al. (2017) states that the LES method is based on introducing a low-pass filtering to define the large-scale flow, while the residual part of the flow is referred to as the SGS flow. LES uses grid spacings on the order of tens to hundreds of metres, allowing boundary layer turbulence to be resolved to scales well within the ISR. At these small scales, the flow is assumed to be in equilibrium, with energy production in balance with dissipation (Germano et al., 1991).

LES models impose a filter on the modelled fields, with the filter scale being set by the grid-spacing Δ of the model. Motions in the flow that are larger than $\mathcal{O}(\Delta)$ are well resolved, while motions with scales smaller than $\mathcal{O}(\Delta)$ are unresolved and referred to as being SGS. As such, the term subfilter is more appropriate for these unresolved motions however, the term subgrid is used in line with the current convention. The filtering of the flow by the grid, along with other model dynamics gives rise to the term "effective resolution" of the model R_{eff}, with values usually ranging between $6\Delta \leq R_{\text{eff}} \leq 10\Delta$. The effective resolution is model specific and defined the length scale at which eddies of the same scale or larger can be well resolved. Any motions smaller than the scale of R_{eff} will either not be resolved well, or at all, and therefore their effect on the flow must be accounted for by a subgrid scheme.

LES models are often used to resolve high resolution fields of the turbulent motions associated with ShCu convection. LES uses grid spacings on the order of tens to hundreds of metres to resolve the Navier Stokes equations. This allows boundary layer turbulence to be resolved to scales well within the ISR. The effect that small-scale motions have on the flow can therefore be captured through the large-scale variable terms in the model using certain laws of statistics (Emanuel, 1994). The transfer of energy from large to small scales is known as the turbulent energy cascade. This can be seen in environmental flows when a large eddy disturbs the nearby fluid and gives rise to smaller eddies, which propagate the same effect to even smaller scales until molecular viscosity dominates and the kinetic energy is dissipated to thermal energy. The scale at which molecular viscosity has an effect is known as the Kolmogorov scale.

2.4 The Grey-Zone

The grey-zone regime occurs when a model's grid spacing Δ is of the same scale as the largest eddies in the flow. For a CBL, this regime occurs when Δ is approximately equal to the height of the capping inversion z_i. This is because z_i scales with the largest thermal structures in the CBL. Previous findings outlined in Honnert et al. (2020) estimate that the transition to the grey-zone regime for dry CBL simulations begins at grid spacings of approximately $\Delta = 200$ m. At these grid scales, the dominant coherent structures cannot be fully resolved on the grid. Furthermore, any partially resolved eddies will experience excessive dissipation of energy by the subgrid scheme.

The assumptions behind the NWP parametrizations for convective turbulence no longer hold in grey-zone regimes because the grid-scale dominant turbulent structures are already partially resolved. The subgrid schemes used in LES models also encounter problems in the grey-zone as the model is unable to resolve down to small enough scales, and so in the grey-zone, energy is dissipated from the dominant structures. Further to this, LES assumes that the grid spacing lies within the ISR. In the ISR the flow is isotropic and energy cascades downscale. A flow which is not isotropic cannot be within the ISR and therefore the LES closure assumptions no longer hold. When turbulence is isotropic, it is assumed that energy production balances dissipation, but this is not the case for the dominant thermal structures in the CBL as they are buoyancy-driven motions. Therefore, LES models will encounter grey-zone issues once the grid spacing is of the same scale as these coherent structures. Wyngaard (2004) defined the grey-zone as the regime which occurs when the length scale of the dominant turbulent structures, l_{turb}, is approximately equal to the grid spacing of the model Δ, so that:

$$\frac{l_{\text{turb}}}{\Delta} \approx 1 \tag{2.3}$$

The numerical dissipation of a model (or the dissipation imposed by a filter) strongly influences the amount of turbulent kinetic energy (TKE) being resolved. The wavelength after which the numerical dissipation dominates over the local downscale flux of energy is defined as the dissipation length scale l_d. Beare (2014) states that for simulations in the grey-zone, numerical dissipation sources from both the advection scheme and the subgrid model are likely to be significant, and discusses the existence of a similarity law between l_d and the resolved TKE in the grey-zone. A new definition of the grey-zone is thus determined, involving the inversion height z_i and dissipation length scale, and it is given by the following ratio:

$$\frac{z_i}{l_{\mathrm{d}}} < 0.7 \tag{2.4}$$

It should be noted that the grey-zone is both a spatial and temporal problem. For example a spatial grey-zone exists near boundaries where the length scale of turbulence decreases with the proximity to the boundary, and turbulent motions inevitably become smaller than the grid spacing. Meanwhile the diurnal cycle encounters a temporal grey-zone. In the early morning, the flow transitions from a very shallow shear-driven night time boundary layer, to a deeper buoyancy-driven CBL. The small scale turbulence of the night time layer would be subgrid scale, and their transition to convective driven thermals would be in the grey-zone regime.

The Smagorinsky SGS used in LES models can also serve to amplify problems caused by the grey-zone regime, as the scheme dissipates energy from the turbulent structures which already lack energy due to the grey-zone effects. This is because LES models assume that the grid spacing lies within the ISR, where the flow is isotropic and the energy cascade is downscale. Grid spacings which are considered to be in the grey-zone are not within the scales of the ISR, and therefore the LES closure assumptions no longer hold. Therefore, the LES model encounters grey-zone issues once the grid spacing is of the same scale as these coherent structures. This aligns with Wyngaard (2004) definition of the grey-zone occurring when the length scale of the dominant turbulent structures, l_{turb}, is approximately equal to the grid spacing of the model Δ.

3 The Smagorinsky Subgrid Scheme

When LES is used to model the motions within a fluid flow, the data field is discretized with respect to the grid spacing Δ. The flow cannot be well resolved unless Δ lies within the ISR scales. However, the eddies with scales close to the Δ scale cannot be fully resolved, resulting in a turbulent cascade of energy that no longer follows the slope defined by Kolmogorov (-5/3) at wavelengths near the grid scale. The energy at these scales must still be dissipated to prevent an energy build-up at the smallest resolved scales. This is where a subgrid scheme is required. Subgrid schemes represent the interaction of the resolved flow with the small, unresolved scales. This ensures adequate dissipation of energy from the resolved scales to the unresolved scales, with the primary goal being to obtain the correct statistics of the energy-containing scales of motion.

3.1 Model Overview

The Smagorinsky-Lilly Subgrid Scheme, commonly referred to as the Smagorinsky Scheme, is a subgrid model used to represent the interaction between the resolved and unresolved scales in the flow of an LES. This scheme is derived from work by Smagorinsky (1963) and altered by Lilly (1962) to account for the effects of buoyancy. This subgrid scheme imitates the effect of molecular viscosity by dissipating energy from the smallest scales of resolved motions in the flow, and can be thought of as being similar to a DNS of flow but with a lower Reynolds number. It imposes an eddy viscosity on the flow, which is dependent on a prescribed mixing length and the stress at that given point in the flow. The eddy viscosity predominantly affects the smallest resolvable scales, preventing a build-up of energy at small scales which would otherwise occur due to the turbulent cascade of energy. The eddy viscosity thereby prevents the simulation from becoming overly energetic. The dissipation of energy at the small scales also ensures that the statistics for the dominant scales of motion are correct.

The Smagorinsky subgrid scheme assumes that energy is being resolved down to very small scales in the flow, so that a clear ISR exists in the energy spectrum. Energy in the ISR is in equilibrium, meaning that the dissipation of energy out of the system must balance with the flux of energy into the system (Germano et al., 1991). Therefore the dissipation is dependent on the energy containing eddies and the model can calculate dissipation using the rate of strain in the large-scale flow (this is explained further in Section 3.1.2). In the model, dissipation accounts for the downscale energy transfer that would occur between the resolved scales and unresolved scales due to the turbulent energy cascade. In other words, this dissipation aims to allow the power speactra to maintain the same slope in the ISR down as far as the partition between resolved and unresolved motions by preventing a build-up of energy that would otherwise occur at the smallest resolvable scales.

In environmental flows, the smallest scale possible where turbulent mixing can still occur before viscosity dominates is referred to as the Kolmogorov scale. Scales smaller than this see the kinematic viscosity dominate, and as a result, the turbulent kinetic energy is dissipated to thermal energy. A similar scale, referred to as the Batchelor scale, exists to describe the smallest length scales of concentration fluctuations of a passive scalar before molecular diffusion dominates. The Smagorinsky-Lilly scheme dissipates energy from resolved motions with wavelengths of $\lambda \approx \mathcal{O}(\overline{\Delta})$ through the use of the eddy viscosity, ν_{m}. The viscosity ν_{m} is dependent on a scale-adaptive mixing length ℓ_{mix}, calculated using the grid spacing Δ and the Smagorinsky parameter C_s. The C_s parameter, when multiplied by the model grid spacing Δ, regulates the ratio between resolved and subgrid scale (SGS) mixing.

3.1.1 The Defining Equations of the Smagorinsky Model

The Smagorinsky scheme is an eddy viscosity based subfilter-stress model which dissipates excess energy at small scales using subgrid stress tensors. This model requires energy to be resolved down to scales well within the ISR where turbulence is isotropic. The kinematic deviatoric stress τ_{ij}^d and scalar flux $h_{\psi,j}$ (for any scalar ψ) are determined by the following equations, and this turbulence closure scheme is discussed further in 3.1.2.

$$\tau_{ij}^d = -2\nu_{\mathrm{m}} S_{ij} \tag{3.1}$$

$$h_{\psi,j} = -\nu_{\mathrm{h}} \frac{\partial \psi}{\partial x_j} \tag{3.2}$$

where ν_{m} is the eddy viscosity, ν_{h} is the thermal diffusivity, S_{ij} is the rate of strain tensor (defined below), and the indices denote the use of Einstein summation notation.

$$S_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \tag{3.3}$$

$$|S| = \sqrt{2 S_{ij} S_{ij}} \tag{3.4}$$

The eddy viscosity is calculated using a flow-dependent rate of strain tensor and a scale-dependent mixing length. Due to the differentiation operator in the rate of strain tensor, the eddy viscosity can focus energy dissipation at the small scales, thus preventing a build-up at that end of the spectrum without substantially affecting the energy at larger scales. The kinematic viscosity ν_{m} and scalar diffusivity ν_{h} are calculated locally via the following equations:

$$\nu_{\mathrm{m}} = \ell_{\mathrm{mix}}^2 |S| \tag{3.5}$$

$$\nu_{\mathrm{h}} = \ell_{\psi}^2 |S| \tag{3.6}$$

with ℓ_{mix}, the mixing length for momentum, determined by a Smagorinsky Coefficient C_s and the grid spacing Δ of the model:

$$\ell_{\mathrm{mix}} = C_s \Delta \tag{3.7}$$

The mixing length for scalars is similar to that of momentum, but it has another dependency on the Prandtl number Pr, which describes the ratio of momentum diffusivity to thermal diffusivity (further details about Pr are given in Section 3.2.2).

$$\ell_{\psi} = \frac{C_s \Delta}{\sqrt{\mathrm{Pr}}} \tag{3.8}$$

The Smagorinsky scheme is scale-aware, as the mixing lengths depend on the grid spacing of the model Δ. When using isotropic grids, the choice for Δ in Equations 3.7 & 3.8 is trivial. However, when running simulations on anisotropic grids, the choice of Δ for these equations is less clear. In conventional models the Smagorinsky scheme imposes an isotropic 3-D filter on the simulated flow, however, the analysis method in this study permits the use of a 2D filter which is applied to the data at each horizontal level (further details given in Sections 4.4 & 4.5). Therefore, as this work focuses on deriving the coherency of the horizontal filter, the horizontal grid spacing $\Delta_{\mathrm{x,y}}$ is used in Equations 3.7 & 3.8. The horizontal grid has been chosen as the primary focus when filtering the flow since modifications in this direction also impact the vertical motions in the flow. Information in the vertical is lost when filtering is applied in the horizontal direction; however, as long as the dominant structures can be resolved on the horizontal grid, these structures can still be represented in the vertical.

In this study, the vertical grid spacing is held constant, as is common in work detailed in previous literature. Considering the vertical scale of key features in a CTBL - such as the surface layer, inversion layers, and the boundary layer as a whole - the coherency of these features degrades more rapidly under vertical filtering than under horizontal filtering with the same filter scale. Furthermore, at the coarsest resolutions investigated in this work, it is not possible to resolve the majority of the flow and therefore the vertical aspect does not have much influence on the resolved flow as a whole. As a result the focus is placed on the horizontal resolutions throughout this work.

3.1.2 Derivation of the Smagorinsky Scheme

The Navier-Stokes equations are used to describe the motion of fluids when modelling environmental flows. While the MONC model allows for flows to be simulated using the anelastic approximation, all three of the case studies investigated in this study use the Boussinesq approximation. Therefore a constant reference density is assumed for this derivation, resulting in an incompressible continuity equation. The model grid can be thought of as imposing a filter on the simulated flow, and quantities resolved on the grid (i.e. filtered quantities) are denoted by the overbar. The filtered continuity and momentum equations then read as follows:

$$\frac{\partial \overline{u_i}}{\partial x_i} = 0 \tag{3.9}$$

$$\frac{\partial \overline{u_i}}{\partial t} + \overline{u_j}\frac{\partial \overline{u_i}}{\partial x_j} = \nu_{\mathrm{m}} \frac{\partial}{\partial x_j}\left(\frac{\partial \overline{u_i}}{\partial x_j} + \frac{\partial \overline{u_j}}{\partial x_i}\right) + \frac{\partial \tau_{ij}^d}{\partial x_j} - \frac{1}{\rho}\frac{\partial \overline{p^*}}{\partial x_i} \tag{3.10}$$

Where u_i describes the velocity component in the i^{th} direction. τ_{ij}^d is the deviatoric stress tensor, defined by $\tau_{ij}^d = R_{ij} - \frac{1}{3}\delta_{ij}R_{ii}$, where R_{ij} is analogous to the Reynolds stress tensor, given by $R_{ij} = \overline{u_i u_j} - \overline{u_i}\,\overline{u_j}$. The variable $\overline{p^*}$ is the filtered pressure, including the isotropic residual stress: $\overline{p^*} = \bar{p} + \frac{1}{3}\rho R_{ij}$.

Following a similar method to Deardorff (1980), these equations make up the governing equations of the model. Taking the divergence of the momentum equation (3.10) gives a Poisson equation, which is used along with the continuity equation (3.9) to solve for the pressure field p (see Moeng (1984) for further details). However, solving for the velocity field **u** requires a closure scheme for the residual SGS stress tensor. This is the point at which the Smagorinsky subgrid scheme is applied. The Smagorinsky scheme can be derived from the conservation equation for τ_{ij}. Note that e is half the trace of the stress tensor τ_{ij}, while the deviatoric stress tensor τ_{ij}^d has the trace removed:

$$e = \frac{1}{2}\left(\overline{u_i u_i} - \overline{u_i}\,\overline{u_i}\right) = \frac{1}{2}u_i'^2 \tag{3.11}$$

$$\tau_{ij}^d = \overline{u_i u_j} - \frac{1}{3}\delta_{ij}\overline{u_k^2} \tag{3.12}$$

$$\frac{\partial \bar{e}}{\partial t} + \overline{u_i \frac{\partial e}{\partial x_i}} = -\tau_{ij}^{d} \left(\frac{\partial \overline{u_i}}{\partial x_j} + \frac{\partial \overline{u_j}}{\partial x_i} \right) + \frac{g}{\theta} \overline{w'\theta'} - \frac{\partial}{\partial x_i} \left(\overline{u_i e} \right) - \frac{1}{\rho} \frac{\partial}{\partial x_i} \left(\overline{u_i p} \right) - \epsilon_f \tag{3.13}$$

The left hand side (LHS) of this equation accounts for the storage and advection of the SGS TKE. On the right hand side (RHS) of the equation, term 1 is the mechanical shear production, term 2 is the buoyancy, term 3 is the turbulent transport, term 4 is the pressure transport, and term 5 is the dissipation in the filtered fields.

The Smagorinsky subgrid scheme is only applicable for simulations run at sufficiently high resolution such that the model can resolve down to scales small enough to be well within the ISR. It is assumed that the flow is in a steady state, therefore the LHS of Equation 3.13 is zero as advection is balanced by the storage. Applying the Smagorinsky scheme at high resolution allows for the assumption of local isotropy, meaning the turbulent transport and pressure transport terms can be disregarded as there is no divergence at these small scales. The buoyancy term can be neglected due to the assumption that energy dissipation only occurs in the ISR where turbulence is isotropic. It should be noted however that, if buoyancy is included in the derivation, the Smagorinsky model will also include the stability functions (discussed further in Section 3.1.3). This allows the model to account for variations in mixing, and therefore dissipation, within the flow due to gradients in potential temperature. When the filter scale is within the ISR scales, the filtered velocity field then accounts for the majority of the TKE in the flow (Pope, 2000), meaning that dissipation results from shear forces only, for length scales in the ISR. These assumptions allow all the terms to be disregarded in Equation 3.13 apart from the mechanical shear production and dissipation terms. The resulting equation now reads:

$$0 = -\tau_{ij}^{d} \left(\frac{\partial \overline{u_i}}{\partial x_j} + \frac{\partial \overline{u_j}}{\partial x_i} \right) + \epsilon_f \tag{3.14}$$

This equation can be rewritten by first applying the first-order closure scheme for momentum as detailed in Equation 3.1 and recalling the definition of the rate of strain tensor in Equation 3.3, yielding the following expression for the diffusion:

$$\epsilon_f = -2\nu_{\mathrm{m}} S_{ij}^2 \tag{3.15}$$

This closure scheme can be derived from the Reynolds averaged momentum equation detailed in Equation 3.10. A similar logic as was applied to the divergence of the momentum equation (Equation 3.13) is also applied here to neglect terms:

$$\cancel{\frac{\partial \overline{u_i}}{\partial t}} + \cancel{\overline{u_j}\frac{\partial \overline{u_i}}{\partial x_j}} = \nu_{\mathrm{m}}\frac{\partial}{\partial x_j}\left(\frac{\partial \overline{u_i}}{\partial x_j} + \frac{\partial \overline{u_j}}{\partial x_i}\right) + \frac{\partial \tau_{ij}^d}{\partial x_j} - \cancel{\frac{1}{\rho}\frac{\partial \overline{p^*}}{\partial x_i}} \tag{3.16}$$

$$\frac{\partial}{\partial x_j}\left[\nu_{\mathrm{m}}\left(\frac{\partial \overline{u_i}}{\partial x_j} + \frac{\partial \overline{u_j}}{\partial x_i}\right) + \tau_{ij}^d\right] = 0 \tag{3.17}$$

This aligns with Prandtl (1925), where it is argued that momentum is transported in the direction of the velocity gradient, with energy moving from eddies with higher velocity to those with lower velocity. Furthermore, the rate of strain tensor characterises the velocity gradient in each direction. Therefore, when the Boussinesq approximation is used, the stress is proportional to the velocity gradient, with the kinetic eddy viscosity ν_{m} denoting the constant of proportionality. This gives the closure of momentum equation detailed in Equation 3.1, and similar arguments can be used to derive the closure for the scalar transport equation.

$$\implies \tau_{ij}^d = -\nu_{\mathrm{m}}\left(2S_{ij}\right) \tag{3.18}$$

Returning now to deriving the expression for diffusion in the Smagorinsky model. Following Kolmogorov's arguments, it is assumed that when the flow is resolved down to scales well within the ISR, the dissipation resulting from the filter removes resolved energy e from the system at a dissipation timescale of t_d.

$$\epsilon_f = -\frac{2e}{t_d} \tag{3.19}$$

Note that there is a factor of 2 in Equation 3.19 results from the definition of energy in the system (see Equation 3.11). The Smagorinsky model relies on the assumption that that model resolution is sufficiently high such that dissipation only occurs in the ISR. Within the ISR the turbulent eddies are isotropic, allowing for their trajectory to be characterised by a length scale alone due to their uniform shape. Prandtl (1925) termed this length scale "the mixing length" ℓ_{mix}, describing it as be the distance travelled by a fluid parcel before its momentum is affected by the environment. Furthermore, in the ISR, the energy spectrum has a universal scaling law and as a result t_d depends only on ℓ_{mix} and the turbulent velocity scale, $\mathrm{u_t}$, of the eddies:

$$t_d = \frac{\ell_{\mathrm{mix}}}{\mathrm{u}_{\mathrm{t}}'} \tag{3.20}$$

Furthermore, at high resolution the model is assumed to dissipate energy from isotropic eddies only, and an additional assumption of incompressibility allows the the subgrid kinetic energy e, as defined in Equation 3.11, to be characterised by a velocity scale $\mathrm{u_t}$:

$$e = \frac{1}{2}\mathrm{u}_\mathrm{t}^2 \tag{3.21}$$

The scaling arguments outlined in Equations 3.19, 3.20, & 3.21 are now applied to Equation 3.15:

$$\epsilon_f \equiv -\frac{\mathrm{u}_\mathrm{t}^3}{\ell_\mathrm{mix}} = -2\nu_\mathrm{m} S_{ij}^2 \tag{3.22}$$

Therefore the eddy viscosity ν is involved in controlling the strength of diffusion ϵ, with Prandtl (1925) stating that both these quantities have the same dimensions, being products of a length and velocity:

$$\nu_\mathrm{m} = \mathrm{u}_\mathrm{t}\ell_\mathrm{mix} \tag{3.23}$$

Equation 3.23 can be rearranged to find an expression for $\mathrm{u_t}$, which can then be substituted into Equation 3.22, yielding a formula for the viscosity:

$$\nu_\mathrm{m} = \ell_\mathrm{mix}^2\sqrt{2S_{ij}S_{ij}} = \ell_\mathrm{mix}^2|S| \tag{3.24}$$

Note that Equation 3.4 was used in deriving Equation 3.24, which enables the viscosity to be defined in terms of the characteristic scales of the SGS eddies.

3.1.3 The Stability Functions

Regardless of model resolution, grey-zone challenges remain unavoidable, especially near boundaries and temperature inversions. Mason and Brown (1999) notes that the buoyancy dependence of the subgrid model is not significant within the flow interior, where the ISR of turbulence is likely well-resolved. However, the effects of buoyancy are no longer negligible as the surface is approached, and this poses an issue to the subgrid scheme. In these regions, stability variations result in eddies which are no longer isotropic, thereby violating one of the key underlying assumptions of the Smagorinsky model. The stability functions can be introduced to the Smagorinsky scheme through the viscosity equations as a means of accounting for these buoyancy effects:

$$\nu_{\mathrm{m}} = \ell_{\mathrm{mix}}^2 |S| f_{\mathrm{m}}(\mathrm{Ri}) \tag{3.25}$$

$$\nu_{\mathrm{h}} = \ell_{\psi}^2 |S| f_{\mathrm{h}}(\mathrm{Ri}) \tag{3.26}$$

where f_{m} and f_{h}, which are functions of the Richardson number Ri, denote the stability functions for momentum and scalars respectively. The Richardson number describes the ratio of buoyancy forces to shear forces, and is used as an indication of the stability conditions throughout the flow. The locally calculated flux Richardson number is calculated pointwise and defined as follows:

$$\mathrm{Ri} = \frac{\mathrm{g}}{\theta_0} \frac{\partial \theta}{\partial z} \tag{3.27}$$

These stability functions allow the Smagorinsky scheme to account for the effects of buoyancy when calculating the eddy viscosity. The formulation of the stability functions vary based on the stability of the flow, with a critical Richardson number, $\mathrm{Ri_c} = 0.25$, being used to determine the point between stable and unstable flow regimes. When $\mathrm{Ri} < 0$, the flow is unstable and convective turbulence dominates. The onset of turbulence begins when the flow experiences $0 \leq \mathrm{Ri} < \mathrm{Ri_c}$. Regions of the flow where $\mathrm{Ri} \geq \mathrm{Ri_c}$ indicate a laminar flow where buoyancy has little influence and shear dominates. The stability functions in the MONC model take the standard form, as detailed in Hill et al. (2018) and Gray et al. (2004). The equations defining these functions are given in Table 3.1 as follows:

Stability Function	$\mathrm{Ri} < 0$	$0 \leq \mathrm{Ri} < \mathrm{Ri_c}$	$\mathrm{Ri} \geq \mathrm{Ri_c}$
f_{m}	$(1 - 16\mathrm{Ri})^{\frac{1}{2}}$	$(1 - \frac{\mathrm{Ri}}{\mathrm{Ri_c}})^4$	0
f_{h}	$(1 - 40\mathrm{Ri})^{\frac{1}{2}}$	$(1 - \frac{\mathrm{Ri}}{\mathrm{Ri_c}})^4(1 - 1.2\mathrm{Ri})$	0

Table 3.1: *The stability functions for momentum and scalars, depending on the Richardson number.*

3.2 The Smagorinsky Parameter and Mixing Lengths

The Smagorinsky parameter is involved in determining the mixing, and therefore dissipation, of energy in the model. It is a determining factor in the proportion of SGS motions compared to those which are resolved. Other factors in this partition include the strain rate, along with the model's grid and advection scheme. The Smagorinsky parameter value varies between different flow regimes. In most models the Smagorinsky parameter for momentum C_s is set explicitly, usually to a constant value. However, there is no equivalent parameter for scalars in the majority of Smagorinsky models. Instead, the Smagorinsky parameter for scalars is assumed proportional to that of momentum, with the ratio fixed using the Prandtl number. In such models, scalars are also assumed to behave the same, and so their transport and diffusion are not scalar specific.

3.2.1 The Momentum Parameter

The Smagorinsky parameter can be understood as the ratio between the length scale of the Smagorinsky model and a measure of the numerical resolution, in this case the grid spacing itself. This parameter partially controls the level of mixing and dissipation of momentum and scalars within the flow. An optimal value of C_s has been derived using Kolmogorov theory, and this is discussed in the following section. However in practice, when using the Smagorinsky scheme in LES models, there is no one universal constant C_s value which can allow for fields to be accurately resolved across all different turbulent regimes. In terms of shallow cumulus simulations, this becomes a problem as the convective thermals in the interior of the flow have different turbulent properties to the shear-dominant near-surface flow. However, it is generally agreed that values for the Smagorinsky parameter which range between $C_s = 0.1$ (Deardorff, 1971) to $C_s = 0.23$ (Lilly, 1966) tend to produce realistic results.

Increasing the C_s value serves to increase the transfer of energy from the resolved scale to the SGS. Therefore, setting the C_s value too large could imply inefficient use of computational resources. This is because higher C_s values dissipate more kinetic energy and dampen the resolved structures. Conversely, lower C_s values do allow for faster spin-up, but if the C_s value is too low the model fields may become noisy and discretization errors may occur. This noise in the resolved field can also inhibit the development of coherent structures (Kealy et al., 2019; Efstathiou and Beare, 2015). For this reason, it is important to have a realistic value for C_s in the model.

Kolmogorov's similarity hypothesis implies that there is a universal constant value which the Smagorinsky parameter tends to when LES models are able to resolve eddies down to scales well within the ISR. This argument leads to the theoretical optimal value of the Smagorinsky Parameter, which describes the ratio between the mixing length (as in Equation 3.7) and the filter scale corresponding to the Smagorinsky scheme. Lilly (1966) determined a Smagorinsky parameter value of $C_s = 0.185$ using the known values of Kolmogorov's constant in the ISR. This "optimal" C_s value is universal and does not exhibit regime dependence. However, in practice, it is difficult to find a constant value of C_s that works in each model under every regime and flow type. This is because it is impossible to resolve the dominant scales of motion in each region of the modelled flow using an LES model, due to both transitions in the flow regime, and the presence of boundaries and temperature inversions, where the eddy scales tend toward zero. This may indicate that regime dependence is a result of the grey-zone. Furthermore, it is not just excessive dissipation due to the grey-zone that prevents a constant value of C_s from working in each case. The Smagorinsky scheme is not the only cause of energy dissipation in an LES model - the model may also be dissipative through its advection scheme and dynamics. Therefore, any "practical" C_s values calculated using model data, rather than Lilly's theoretical, are assumed to be model-dependent. As a result, Mason and Callen (1986) states that, in this context, C_s can be thought of as a measure of numerical accuracy.

In models using the standard Smagorinsky scheme, the coefficient C_s must be set to a constant value. As a result, numerous studies have sought to determine the best practical value for C_s in their specific model. While Lilly (1966) derived the optimal C_s theoretically, this study found that in practice, it is necessary to increase the C_s value. This alteration was required in order to account for the velocity differences across single grid intervals when calculating finite differences, and it was concluded that $C_s = 0.23$ is a more reasonable value to use in LES models. (Deardorff, 1970) highlighted the importance of adapting "practical" C_s values based on the flow regime, and derived suitable values using empirical fits between Smagorinsky data and DNS data. McMillan et al. (1980) found that in a homogeneous flow, C_s values decrease as strain rate increases. Experiments carried out by Mason (1994) showed that, in a neutral boundary layer, reducing the value of the Smagorinsky parameter allows the resolution of smaller scale structures. This study found setting $C_s = 0.2$ gave the most accurate results, and any further reductions to the C_s parameter value caused notable finite-difference errors. Shear-driven flows however, commonly have lower C_s values usually around 0.1 to 0.13 (Piomelli et al., 1988; Deardorff, 1970, 1971), while convective flows tend to need higher C_s values of between 0.17 and 0.23 to give reasonable results (Lilly, 1966; Mason and Callen, 1986).

Choosing the correct C_s for the CTBL is difficult as it exhibits different regimes, depending on position relative to the key features in the flow: shear-driven flows occur at boundaries and inversions, while the rest of the flow in a CTBL is strongly buoyancy-driven, especially within thermals and clouds. Within the cloud layer there is a substantial difference between the in-cloud regions, which are extremely turbulent, and the non-cloudy surrounding environment which is conditionally stable. The optimal value of C_s also varies widely during transitioning flows, such as the transition experienced as the nocturnal boundary layer evolves into a CBL during the morning and early afternoon. This is because, during the diurnal cycle, the BL progresses from a shallow shear-driven layer to a deeper convective-driven layer. This suggests the need for the C_s parameter to be able to adapt to the flow.

Recently, with the improvements in computational resources, there has been an increased focus on adapting the C_s parameter to either the flow itself, or to key features within the flow. Kealy et al. (2019) developed a functional relationship between the C_s parameter and the mixed layer capping inversion height. This allows the model to adapt the diffusion according to the growth of the dominant eddy scales as the CBL transitions from a stable early morning BL to the daytime CBL. This method, when implemented in an LES model at grey-zone scales, produced a similar spin-up time and a TKE field that closely matched the idealised TKE fields. These idealised fields were produced by filtering high resolution reference data to the same scale as the grey-zone data. However, this method only allowed C_s to vary in the vertical. Germano et al. (1991) derived a dynamic model, which was improved upon by Lilly (1992) (further details are discussed in Section 3.3) to account for variations in the flow across the entire domain. Experiments conducted by Efstathiou and Plant (2019) found that a scale-dependent Lagrangian averaged model with a dynamically calculated C_s produced more realistic levels of turbulence earlier in the simulation than the standard Smagorinsky model at grey-zone resolutions. This improvement was attributed to the adaptability of the dynamic C_s parameter, with the dynamic model demonstrating less sensitivity to grey-zone resolutions than the standard model.

3.2.2 Scalar Mixing: The Prandtl Number and Schmidt Number

As previously discussed, the C_s parameter is used to set the mixing length for momentum, which plays a role in determining the turbulent transfer of kinetic energy across the scales. In the standard Smagorinsky model this parameter not only controls kinetic energy dissipation, but also determines the diffusivity of scalars such as heat and moisture. This is because the turbulent transfer of heat is assumed to be proportional to that of momentum, with the relationship between the two set by a factor known as the Prandtl number, Pr (Mason and Brown, 1999). The parameter which governs the dissipation of heat in LES models, referred to as C_θ in this work, is determined by the following equation:

$$C_\theta^2 = \frac{C_s^2}{\mathrm{Pr}} \tag{3.28}$$

The governing equation for the Prandtl number, along with its properties in the CTBL are discussed in Section 2.2.2. In the neutral BL, de Roode et al. (2017) noted that while the Prandtl number is close to unity, as convective turbulence takes over, observations suggest the Prandtl number decreases, and the Prandtl number for dry air is $\mathrm{Pr} = 0.7$. Lower Prandtl numbers indicate that thermal diffusivity dominates over the diffusion of momentum.

The CTBL demonstrates different flow regimes, both throughout its layers and during its growth over time, causing the Prandtl number value to vary. Observational data shows that values for the turbulent Prandtl number $\mathrm{Pr_t}$ can vary widely across a given stability range, with Grachev et al. (2007) suggesting that a universal relationship between stability and $\mathrm{Pr_t}$ might not exist. However, when modelling the CTBL, previous work by Li (2016) and Efstathiou (2023) has demonstrated that Pr increases in the near-surface boundary region as the grid spacing is coarsened to grey-zone resolutions. Further to this, Efstathiou (2023) showed a significant increase in Pr in the CL at coarser resolutions. This is in agreement with Shi et al. (2018), who found that the mean profile of the Prandtl number is a relatively constant value of Pr ≈ 0.5 throughout the ML, before increasing to an average value of Pr ≈ 10 in the cloud layer. This dramatic increase in the Prandtl number indicates a reduction to near zero mixing of heat as the scale of thermal structures shrinks in comparison to the momentum structures.

For scalars other than heat, the calculation of the dissipation parameter can be made more general by using the Prandtl number's counterpart, the Schmidt number Sc. Equation 3.29 can then be rewritten in a form which expresses the Smagorinsky parameter for any scalar C_ψ in terms of the corresponding Schmidt number Sc_ψ:

$$C_\psi^2 = \frac{C_s^2}{\mathrm{Sc}_\psi} \tag{3.29}$$

Note that the formula for Sc_ψ is defined in Equation 2.2. Shi et al. (2018) found that the average Schmidt number for total moisture, Sc_{q_t}, throughout the depth of both the mixed layer and cloud layer was $\mathrm{Sc}_{q_t} \approx 0.5$, apart from a peak of $\mathrm{Sc}_{q_t} \approx 10$ at the top of the cloud layer. This is in contrast to the Pr values in the CL which were discussed previously. The dissimilarity between the turbulent mixing of heat and water vapour has been highlighted by Warhaft (1976), Goldberg et al. (2010), and Katul et al. (2016), particularly when modelling flows with inversions and unstable conditions, both of which are key features in the CTBL. Despite this, there are no operational formulations to account for scalar dissimilarities being used in LES models (Li, 2019).

The mixing parameters in the standard Smagorinsky scheme are not scalar specific, and each scalar is assumed to be dissipated at the same rate as that of heat. This assumption is based on the premise that the same thermals are transporting the scalars, and it is assumed that all quantities in a given thermal are being transported and diffused identically. As a result, these models rely solely on the Prandtl number to determine the dissipation rate for every scalar, rather than using scalar-specific Schmidt numbers. This forces the turbulent mixing of water to be the same as that of heat. In contrast to this, Shi et al. (2018) found that Sc_{q_t} changes independently from Pr, with mean profiles for Pr and Sc_{q_t} shown to be clearly distinct from each other within the cloud layer. When these differences between scalars are accounted for through the use of dynamic C_s, Pr_θ, and Sc_{q_t} parameters, a clear match between the dynamic simulation of the cloud field and field data is achieved (Shi et al., 2018). Furthermore, this study found that the standard Smagorinsky simulation produces an unrealistically thin cloud layer with lower values of cloud liquid water content than were observed in the field campaign. Therefore it seems important to account for the differences between water fluxes and heat fluxes when modelling CTBLs.

3.2.3 The Mixing Length

The mixing length can be thought of as the measure of the ability for mixing to occur in a flow (Stull, 2009). In physical space, it is analogous to the length of the mean free path along which a parcel can conserve its characteristics before beginning to mix into the surrounding environment. The mixing length hypothesis was first introduced by Prandtl (1925) as a method of describing the turbulent transfer of energy using an eddy viscosity. In the Smagorinsky scheme, the mixing length describes the level of dissipation that the resolved flow experiences. The basic mixing length scale ℓ_0 is calculated using the grid scale Δ and the Smagorinsky parameter for momentum C_s:

$$\ell_0 = C_s \Delta \tag{3.30}$$

In the CBL, the Smagorinsky parameter usually takes a value of $C_s \approx 0.2$. It should be noted that the scale at which motions become unresolved begins at much larger scales than the mixing length. Models can resolve motions with wavelengths larger than 6Δ to 8Δ, with the smallest wavelength possible to solve on the grid being 2Δ. Therefore in LES models, the mixing length does not represent the size of the unresolved eddies, but rather the viscous dissipation rate of the turbulent flow. As discussed previously, it is well known that turbulent mixing and energy dissipation are greatly reduced as boundaries are approached. The issue posed by the use of a constant Smagorinsky parameter in the surface layer where shear dominates can be overcome by accounting for the effect of the surface when calculating the mixing length ℓ_{mix} using the formula derived by Blackadar (1962):

$$\frac{1}{\ell_{\text{mix}}^2} = \frac{1}{\ell_0^2} + \frac{1}{[k(z + z_0)]^2} \tag{3.31}$$

where $k = 0.4$ is the von Karman constant, z is the height above the surface, and z_0 is the surface roughness length. This equation reduces the mixing length as the surface is approached, allowing for a smooth transition from the mixing length in the flow's interior to one which is a function of distance from the surface (Gray et al., 2004). This ensures that the effect of the near-surface shear, which reduces the turbulent mixing as discussed in Section 3.2.1, is accounted for in the model. In the standard model, where the Smagorinsky parameters are set to constants, a similar length scale for scalars ℓ_ψ is employed by the model to account for the mixing and dissipation of heat energy. The difference between fluxes of heat and momentum is accounted for when calculating the basic mixing length for scalars using the Prandtl number Pr, with C_θ from Equation 3.28 being substituted in for C_s into Equations 3.30 & 3.31. The same method can be used for other scalars by using their corresponding Schmidt numbers, though as previously mentioned most standard models are not scalar-specific and so, are assumed to have the same mixing length as that of heat.

3.3 The Standard Smagorinsky Model

There are different versions of the Smagorinsky subgrid scheme; the Smagorinsky-Lilly model (Lilly, 1966) is the scheme used as the "standard model" throughout this study. The standard Smagorinsky scheme does not allow the value of the Smagorinsky parameters to vary in the flow. Instead, the Smagorinsky parameter controlling the dissipation of momentum, C_s, is fixed to a constant value. In the MONC model, which uses the standard Smagorinsky scheme, the scalar fluxes are defined as functions of the momentum flux and stability functions (see Section 3.1.3). These scalar fluxes differ from the momentum fluxes only by a factor; for example, the Smagorinsky parameter for heat, C_θ, is governed by C_s and the Prandtl number Pr:

$$C_\theta^2 = \frac{C_s}{\text{Pr}} \tag{3.32}$$

This parameter sets the mixing length, and thus dissipation, for all scalars in the simulation. Therefore, if stability functions are not included in the standard model, the mixing lengths for momentum and scalars are fixed to constant values throughout the flow. While this can work well in the interior of well-resolved flows, this method breaks down in regions near boundaries and inversions. In these areas, the eddies become smaller and smaller as the boundary is approached, due to the shear that is acting upon the flow. This results in grey-zones forming in the areas affected by this shear-dominant flow. Adaptions such as Blackadar's formula (Equation 3.31) can be used here, as it forces the mixing length to tend towards zero in the near-surface region. While many models using the standard Smagorinsky scheme employ Blackadar's formula to overcome the issues posed by using a constant C_s value in the surface layer, similar adaptions are not taken for the free shear-layers found at the ML capping inversion z_{ML} or the CL top z_{CT}. Additionally, setting C_s to a constant value becomes significantly less effective as the model's resolution decreases and grey-zone effects begin to impact the entire flow.

While it is clear that the Smagorinsky parameter C_s controls the turbulent momentum flux, it also determines the heat fluxes and moisture fluxes in the standard Smagorinsky model. This is because scalar fluxes are defined as a function of momentum fluxes and some constant of proportionality in the standard scheme. In the case of thermal diffusivity, the Prandtl number Pr is used to set this ratio. Recall from Section 3.2.2 that the Prandtl number describes the ratio of the momentum diffusion to the thermal diffusion. The standard scheme sets the Prandtl number to a constant value of Pr $= 0.7$ throughout the flow, as this is the Prandtl number for thermal diffusion in air. This allows for the calculation of the Smagorinsky parameter for heat, C_θ. Furthermore, the standard Smagorinsky model assumes that heat and moisture are transported and diffused in the same way (Li, 2019). This essentially sets the Schmidt number Sc, which describes the ratio of momentum diffusivity to mass diffusivity for any scalar, to be equal to its counterpart the Prandtl number (Gualtieri et al., 2017). Therefore the scalar fluxes in the standard Smagorinsky model are not scalar-dependent, with the parameter for any scalar ψ being set to the heat parameter: $C_\psi = C_\theta$. This is despite the fact that heat is an active scalar that can alter the flow, whereas other scalars like moisture and pollution are passive scalars in this case.

In the standard Smagorinsky model, scale awareness is achieved by relating the mixing length to the model grid spacing using the Smagorinsky parameter. However, as the Smagorinsky parameter does not change value as the grid spacing changes, the standard model is not scale-adaptive. This can pose an issue when using LES in grey-zone regimes. The MONC model extends the classic Smagorinsky approach by having the mixing length also account for distance from the surface, as in Equation 3.31. This can be considered an adaptation for the near-surface grey-zone due to the small-scale eddies present in that region of the flow.

3.4 The Dynamic Smagorinsky Model

The standard Smagorinsky scheme performs well for high-resolution LES, but it begins to break down as grid spacing increases to grey-zone resolutions. However, adaptations to the Smagorinsky model can be made to tackle the issues posed by the grey-zone. At coarse resolutions which would typically be considered grey-zone scales, having fixed values for the Smagorinsky parameter and Prandtl number may exacerbate the already excessive dissipation of energy. To address the limitations of using a fixed viscosity coefficient, Germano et al. (1991) developed the dynamic Smagorinsky model. This model was further optimised by Lilly (1992) to use a least squares approach to minimise the error when calculating C_s. The dynamic method calculates a value for C_s at each point in the domain according to the stress-strain relationship between the smallest resolved scales in the flow. However, this method leads to an overdetermined system, hence the use of the least mean squared error, and this is discussed further in

Section 3.4.1. The dynamic model can also calculate dissipation coefficients for scalars without relying on the momentum flux or the Prandtl number. Flow-dependent, scalar-specific diffusivity parameters C_ψ can be calculated for each scalar ψ based on the scalar gradients in the field. This enables the scalar fluxes to be included in the model individually, rather than assuming that each scalar is transported and dissipated at the same rate as heat. The equations used for dynamically calculating flow-dependent Smagorinsky parameters are detailed in Section 3.4.1.

These dynamic parameters can adjust to fluctuations in eddy length scales, which vary as a result of position relative to boundaries, inversions, and/or clouds. The scale of eddies also varies in time as eddies grow in size, either during model spin-up (when the turbulence reaches a resolvable scale) or as the boundary layer transitions from a shallow layer to a deeper convective layer (Kealy et al., 2019). This spatial and temporal variation in the turbulence scale is a contributing factor to the grey-zone problem (Mason and Callen, 1986), and therefore the use of a dynamic model can help in overcoming these issues when modelling at coarse resolutions. The use of a flow-dependent Smagorinsky parameter enables pointwise variations in the flow to be accounted for when calculating the dissipation, yielding better results in grey-zone regimes than the standard Smagorinsky scheme (Efstathiou, 2023). Dynamic models show better agreement with DNS data and are better able to resolve flows in the grey-zone regime, however, they require substantial computational power. This high computational cost is due to the pointwise nature of the dynamic C_s calculations.

The dynamic model computes the difference in stress and strain between data at two different effective grid spacings. Effective grid spacing describes the grid scale that a filtered dataset corresponds to in physical space, given that a filter with a known filter scale has been applied to the dataset in spectral space. The first scale used in the Germano dynamic model is the subgrid scale (SGS), which is calculated by applying a filter with an effective grid spacing of $\overline{\Delta}$ to the flow. Note that for many studies, this SGS filter is taken to be the filter imposed on a flow by the grid of the LES model. The second scale used in the Germano method is the subtest scale (STS), which was calculated by applying a filter with an effective grid spacing of $\widehat{\overline{\Delta}}$ to the SGS data. The power speactra of the SGS and STS data will deviate from the Kolmogorov -5⁄3 slope at different points, with the STS deviating at a larger wavelength than the SGS. By applying the convolution theorem, the differences between these two power spectra can be computed, and a “test window” is formed between the SGS and STS spectra, as illustrated in Figure 3.1.

The Germano approach to computing dynamic parameters assumes that C_s is scale invariant between the $\overline{\Delta}$ and $\widehat{\overline{\Delta}}$ scales. This requires the test window to be within the ISR. However, by definition, any model in the grey-zone regime will be unable to resolve down to scales in the ISR. To address the limitation of the traditional dynamic model, which requires scale invariance of C_s, Meneveau et al. (1996) investigated different methods to make the dynamic model scale-dependent. The method that proved most promising utilised a second STS filter to determine how the coefficient changes across scales, however, it required prior knowledge on the scaling tendencies of C_s. Porté-Agel et al. (2000) developed a scale-dependent dynamic model which also utilises a second STS filter to determine how the coefficient changes across scales. However, this method requires making relatively weak assumptions in order to solve the simultaneous equations needed to determine a scale-dependent C_s. To address this issue, Bou-Zeid et al. (2005) employed a second STS filter and used an iterative method to determine the Smagorinsky parameter, which differed from previous approaches. This method begins with an initial assumption of scale invariance, after which a scale-dependent C_s can be computed.

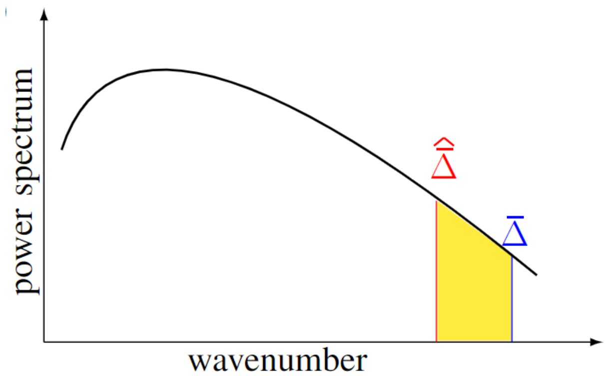

Figure 3.1: *The "test window" formed between data which has been filtered from a DNS base dataset to data with a filter scale of $\overline{\Delta}$, and filtered a second time to a coarser dataset with filter scale $\widehat{\Delta}$. For simplicity the diagram shows the filtering that would result from a spectral wavenumber cutoff filter.*

3.4.1 The Dynamic Smagorinsky Equations

The standard Smagorinsky model performs well for LES regimes, but begins to break down if grid spacing increases to grey-zone resolutions. However, adaptations to the Smagorinsky model can be made to tackle the issues posed by the grey-zone. One such adaptation is allowing the values of C_s and C_θ to vary based on the flow. Such a model is known as the dynamic Smagorinsky model. Another adjustment that can be made is allowing for the diffusivities of heat and moisture to differ from each other, rather than having a single fixed scalar diffusivity which depends on the Smagorinsky parameter for momentum and a prescribed Prandtl number. Flow-dependent scalar diffusivities can be achieved by calculating the C_θ and C_{q_t} parameters based on their respective scalar gradients in the field.

Following the method outlined by Lilly (1992), the first step is to compute the differences between stress and strain at two different filter scales, denoted by the overbar and hat symbols.

The difference between these scales forms the test window, as in Figure 3.1. The equations that define the Dynamic Smagorinsky model are then as follows:

$$\bar{S}_{ij} = \frac{1}{2}\left(\frac{\partial \overline{u_i}}{\partial x_j} + \frac{\partial \overline{u_j}}{\partial x_i}\right) \tag{3.33}$$

$$|\bar{S}| = \sqrt{2\bar{S}_{kl}\bar{S}_{kl}} \tag{3.34}$$

$$L_{ij} = \widehat{\overline{u_i}}\widehat{\overline{u_j}} - \widehat{\overline{u_i}\overline{u_j}} \tag{3.35}$$

$$M_{ij} = \widehat{\bar{\Delta}}^2|\widehat{\bar{S}}|\widehat{\bar{S}_{ij}} - \bar{\Delta}^2\widehat{|\bar{S}|\bar{S}_{ij}} \tag{3.36}$$

where the overbar denotes the grid scale, and the hat symbol denotes the coarser test filter scale. The L_{ij} tensor evaluates the difference in stress between the STS and SGS, while the M_{ij} tensor is analogous to the difference in rates of strain between the STS and SGS. Following the reasoning outlined in Lilly (1992), the differences in stress and strain within the test window lead to the following equation:

$$L_{ij} - \delta_{ij}L_{kk} = 2C_s^2 M_{ij} \tag{3.37}$$

As previously stated, this is an overdetermined system, and as a result, the least squares approach is employed to minimise the error $\mathbb{E}$ in order to determine an equation for C_s:

$$\mathbb{E}^2 = (L_{ij} - \delta_{ij}L_{kk} - 2C_s^2 M_{ij})^2 \tag{3.38}$$

Setting $\frac{\partial \mathbb{E}}{\partial (C_s^2)} = 0$ to find the minimum:

$$\frac{\partial \mathbb{E}}{\partial (C_s^2)} = 2L_{ij}M_{ij} - 4C_s M_{ij}^2 = 0 \tag{3.39}$$

The Smagorinsky parameter C_s can then be evaluated as:

$$C_s^2 = \frac{1}{2}\left(\frac{L_{ij}M_{ij}}{M_{kl}^2}\right) \tag{3.40}$$

$$C_s = \sqrt{\frac{1}{2}\left(\frac{\max(0, L_{ij}M_{ij})}{M_{kl}^2}\right)} \tag{3.41}$$

The dynamic Smagorinsky model can be altered further to allow for flow-dependent scalar fluxes, for any scalar ψ:

$$H_j = \widehat{\bar{u}}_j \widehat{\bar{\psi}} - \widehat{\bar{u}_j \bar{\psi}} \tag{3.42}$$

$$R_j = \widehat{\bar{\Delta}}^2 |\widehat{\bar{S}}| \widehat{\frac{\partial \bar{\psi}}{\partial x_j}} - \bar{\Delta}^2 \widehat{|\bar{S}| \frac{\partial \bar{\psi}}{\partial x_j}} \tag{3.43}$$

$$C_\psi^2 = \frac{1}{2} \left(\frac{H_j R_j}{R_k^2} \right) \tag{3.44}$$

$$C_\psi = \sqrt{\frac{1}{2} \left(\frac{\max(0, H_j R_j)}{R_k^2} \right)} \tag{3.45}$$

Here the angle brackets denote the planar average for each height level in the domain. Note that the “clipping” of negative C_s^2 and C_ψ^2 values by setting them to zero in Equations 3.41 and 3.45 is necessary as it removes the unrealistic negative values of viscosity. These negative values are unusable in the Smagorinsky scheme as they would cause the model to crash. They are removed from the averaging calculation and replaced with a zero, indicating a total lack of mixing at that specific point, rather than the unrealistic “negative mixing” previously indicated by the negative C_s and C_ψ values.

3.4.2 A Scale-Dependent Dynamic Model

The method of calculating C_s and C_ψ in the previous section assumes that the Smagorinsky parameter is scale invariant. However, this requires Δ to lie well within the ISR and as such, the assumption of scale invariance does not hold in the grey-zone. To address this shortcoming of the traditional dynamic model, Porté-Agel et al. (2000) developed a scale-dependent model, which was further modified by Bou-Zeid et al. (2005). This model uses a second test filter to determine the extent of the change in the Smagorinsky parameter across two distinct scales. This method introduces the β parameter to quantify the degree of scale dependence demonstrated by the Smagorinsky parameter. A value of $\beta = 1$ denotes scale invariance, while values where $\beta \to 0$ correspond to severe scale dependence. The β parameter via the following equation:

$$\beta = \max \left(\frac{\langle C_{s4\Delta}^2 \rangle}{\langle C_{s2\Delta}^2 \rangle}, 0.125 \right) \tag{3.46}$$

where the angle brackets denote any averaging operation applied to the Smagorinsky parameter. Note that, as the C_s parameter is determined using the least square minimisation method, it is clear from Equation 3.39 that the numerator and denominator of the defining equation for C_s must be averaged individually:

$$\langle C_s^2 \rangle = \frac{1}{2}\left(\frac{\langle \max(0, L_{ij}M_{ij})\rangle}{\langle M_{kl}^2 \rangle}\right) \tag{3.47}$$

The β clipping at 0.125 is to avoid numerical instabilities which would otherwise occur at points where the $C_{s4\Delta} \longrightarrow 0$ when $C_{s2\Delta} \not\longrightarrow 0$. Following Bou-Zeid et al. (2005), the cut-off value of 0.125 is used as it corresponds to half the theoretical average minimum value. The mixing length, defined by $\ell_{\text{mix}} = C_s^2\Delta^2$, is typically at its smallest value as the surface is approached, and in this near-surface region the mixing length scales with distance from the boundary. Assuming an isotropic grid, the smallest eddies occur within the first grid space from the surface. Therefore, by rearranging the mixing length equation, the ratio in Equation 3.46 can be rewritten to find the minimum possible theoretical value.

$$\beta_{\text{min}} = \frac{\left(\frac{\ell_{\text{mix}}}{4\Delta}\right)^2}{\left(\frac{\ell_{\text{mix}}}{2\Delta}\right)^2} = \frac{1}{4} \tag{3.48}$$

This theoretical minimum value is then halved in order to ensure that the clipping limit is significantly less than the physically expected limiting behaviours (Bou-Zeid et al., 2005). Clipping the β parameter at this value ensures that scale dependence within the theoretical limits is accounted for, while also preventing the simulation from violating local viscous stability conditions. The β profile that results can then be used to calculate a scale dependent $C_{s\beta}$.

$$C_{s\Delta}^2 \equiv C_{s\beta}^2 = \frac{C_{s2\Delta}^2}{\beta} \tag{3.49}$$

In the tests of atmospheric boundary layer flow over homogeneous surfaces performed in Porté-Agel et al. (2000), the scale-dependent dynamic model was implemented using planar averaging, i.e., the averages required to enforce the Germano identity were evaluated over horizontal planes parallel to the ground. This was appropriate for the simple geometries envisioned in those tests, where horizontal planes correspond to directions of statistical homogeneity of the turbulence. This method can be extended further, allowing scale-dependent C_s and C_ψ to be calculated for any grid spacing Δ by modifying the calculation of the M_{ij} and R_{ij} tensors. M_{ij} becomes:

$$M_{ij} = \beta\widehat{\bar{\Delta}}^2|\widehat{\bar{S}}|\widehat{\overline{S_{ij}}} - \bar{\Delta}^2\widehat{|\bar{S}|\overline{S_{ij}}} \tag{3.50}$$

similarly R_j becomes

$$R_j = \beta \widehat{\bar{\Delta}}^2 |\widehat{\bar{S}}| \widehat{\frac{\partial \bar{\psi}}{\partial x_j}} - \bar{\Delta}^2 \widehat{|\bar{S}| \frac{\partial \bar{\psi}}{\partial x_j}} \tag{3.51}$$

3.4.3 Stability Dependence in the Dynamic Smagorinsky Model

Stability functions can also be included in the dynamic Smagorinsky equations to account for buoyancy effects in regions of the flow where eddies may not be fully resolved. This is due to the reasons previously discussed in Section 3.1.3. Note that, in this study, the following stability functions are not included when applying the dynamic equations to data. The stability-dependent dynamic Smagorinsky equations are presented here solely for completeness, in accordance with the methodology outlined by Efstathiou et al. (2018). The stability functions f_m and f_h, for heat and scalars respectively, remain the same as those detailed in Table 3.1. The stability dependency appears in the M_{ij} and R_j equations in the dynamic model:

$$M_{ij} = \widehat{\bar{\Delta}}^2 |\widehat{\bar{S}}| \widehat{\overline{S_{ij}}} \widehat{f}_\mathrm{m}(\mathrm{Ri}) - \bar{\Delta}^2 \widehat{|\bar{S}| \overline{S_{ij}} f_\mathrm{m}(\mathrm{Ri})} \tag{3.52}$$

and similarly R_j becomes

$$R_j = \widehat{\bar{\Delta}}^2 |\widehat{\bar{S}}| \widehat{\frac{\partial \bar{\psi}}{\partial x_j}} \widehat{f}_\mathrm{h}(\mathrm{Ri}) - \bar{\Delta}^2 \widehat{|\bar{S}| \frac{\partial \bar{\psi}}{\partial x_j} f_\mathrm{h}(\mathrm{Ri})} \tag{3.53}$$

These equations can also be made scale dependent through the use of the β parameter, similar to the Equations 3.50 and 3.51. This would enable the dynamic Smagorinsky model to be both scale-dependent and stability-dependent.

4 Case Studies and Model Set-up

4.1 The MONC Model

The LES model used to conduct this study is the Met Office/NERC Cloud Model (MONC), which is a rewrite of the Met Office Large Eddy Model (LEM). The LEM is an older model which is unable to scale beyond 512 cores, whereas the MONC model has been designed to run on HPC systems with high CPU core counts (Brown et al., 2014). MONC applies the quasi-Boussinesq anelastic approximation to the 3D Navier Stokes equations, employing a height-dependent hydrostatic reference state for temperature, pressure and density (Hill et al., 2018; Efstathiou, 2023). These equations are resolved on an Arakawa-C grid. An energy-conserving centred difference scheme (Piacsek and Williams, 1970) is used for the advection of momentum, while a positivity-preserving total variation diminishing (TVD) scheme is used for scalars (Leonard et al., 1993). The model grid acts as a low pass filter to the flow, partitioning it into resolved and unresolved (subgrid) scales. The energy transfer resulting from the turbulent cascade of energy from resolved to unresolved scales is governed by the Smagorinsky scheme, as described in Section 3. This energy transfer is controlled by the Smagorinsky parameter, which the standard MONC configuration fixes to a constant value of $C_s = 0.23$. The viscosity, which controls the dissipation of energy from resolved to SGS (recall Equation 3.15), is calculated using C_s, the rate of strain, and an additional stability function. This study uses the standard form of the stability functions, as detailed in Section 3.1.3, to determine the subgrid contribution to the flow. These functions assess whether turbulence should be occurring in a specific stability regime through the use of a critical Richardson number and, for scalars the Prandtl number is also used as an additional factor.

4.2 Background of Case Studies Used

The MONC model, with a standard Smagorinsky subgrid scheme, was used to produce high-resolution simulations of three different case studies: a dry CBL, the BOMEX case, and the ARM case. The dry CBL case was used as a simple initial test case to ensure that the MONC model was set up correctly and producing the expected results. The main focus when analysing results from this case was determining the sensitivity of the model to grid spacing. There is no vapour or liquid water content in the simulation and the boundary layer cap was maintained at an approximate height of 1 km through the use of a steep temperature inversion. There is a constant positive heat flux at the surface, giving rise to a strong convective thermal circulation. This produces an idealised case of the eddy overturning which occurs in the daytime convective boundary layer. Further details on this case study are given in Section 4.3.1.

The Barbados Oceanographic and Meteorological Experiment (BOMEX) is a standard case of marine shallow cumulus convection. Analysis from the initial field campaign is detailed in Holland and Rasmusson (1973), while the set up of the LES model for this case was designed by Siebesma and Cuijpers (1995). This study uses data from Phase 3 of the BOMEX case, focusing on a five day period of settled conditions from the 22nd to the 26th of June 1969. During this period, conditions were remarkably undisturbed, such that the system can be classified as being in a quasi-steady state. Siebesma et al. (2003) notes that the trade wind cumulus clouds that formed did not precipitate or interact with mesoscale circulations. A large-scale downward motion dominates during this time (Nitta and Esbensen, 1974), limiting the cloud layer to lie between the 500 m to 1500 m levels (Siebesma et al., 2003). During this period, 15 rawinsonde soundings were launched per day, as per the program detailed by Davidson (1968). This data was used to produce large-scale profiles of temperature, humidity, wind speed and direction across the BOMEX domain, as shown in Figure 4.1. Further details on this case study are given in Section 4.3.2.

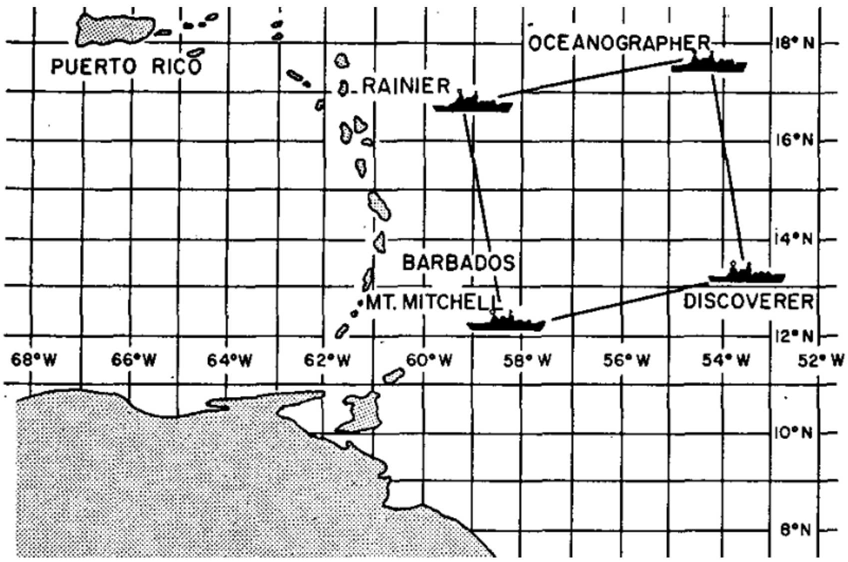

Figure 4.1: *The ship array during Phase 3 of the BOMEX Observation period (Nitta and Esbensen, 1974).*

The Atmospheric Radiation Measurement (ARM) case is an idealisation of measurements taken on the 21st June 1997 at the Southern Great Plains (SGP) site near Lamont, Oklahoma (Chlond et al., 2014). This is a well-studied diurnal cycle of shallow cumuli over arid, flat land driven by sensible heat fluxes. This case has been intensively studied by Browning et al. (1993) as part of the 6th Global Energy and Water Cycle Experiment (GEWEX) Cloud System study (GCSS), which tested numerous LES models' abilities to simulate the development of shallow cumulus over land. The model set-up for ARM used in this study was originally employed as a test bed for an LES intercomparison study by Brown et al. (2002). The ARM Program has been acknowledged by Ahlgrimm et al. (2016) as a case which has played a key role in the development of the parametrizations of subgrid processes for cloud, precipitation, and radiation in the ECMWF models.

The ARM case shows a field of non-precipitating cumulus clouds developing over an initially clear convective boundary layer on the day in question. During the diurnal cycle, this case transitions between five key stages: from the shallow morning-time boundary layer, to a dry convective boundary layer, to one with the onset of passive clouds, moving to a more rapid cloud development stage, and finally reaching a quasi-steady like stage in the evening (Efstathiou, 2023). The exact timings of these transitions, both for the local time and equivalent UTC (for ease of comparison to other literature based on the ARM case) are illustrated in Figure 4.2. Further details on this case study are given in Section 4.3.3.

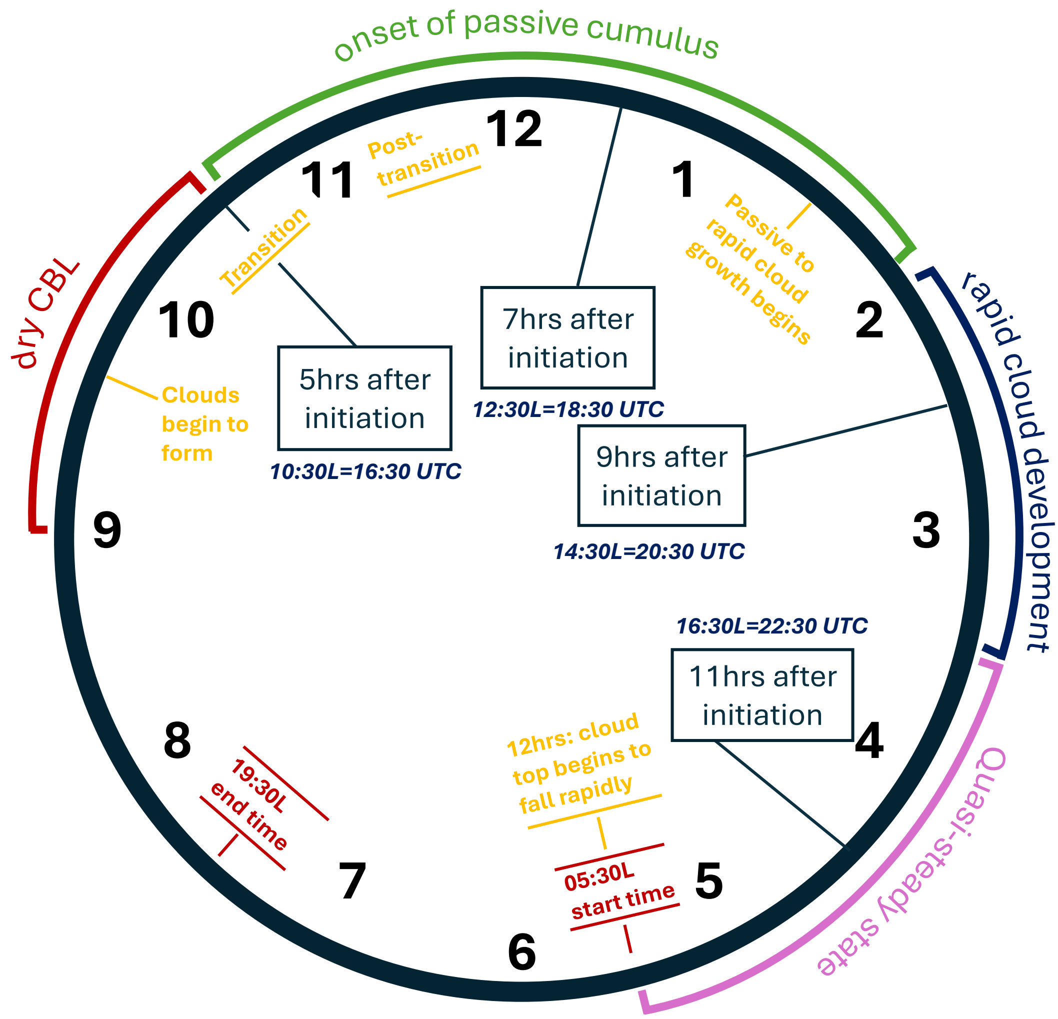

__Figure 4.2:__ A visualisation of the time stamps of interest during the diurnal cycle from the ARM case. The time samples used in the analysis are marked in black boxes. The local time (L) is used as convention throughout this thesis, however, other studies on the ARM case have used UTC, so this was included for ease of comparison.

4.3 LES Model Set-up

The MONC model, with a standard Smagorinsky subgrid scheme set-up, was used to produce LES for the three distinct cases. MONC employed an adaptive time step, governed by the Courant–Friedrichs–Lewy (CFL) criterion. The pressure field was computed using an iterative solver, which was subsequently incorporated into the momentum equation. The initial conditions specific to each case were then included in the model, resulting in the generation of LES for each scenario. The model set-up specific to each case is detailed in the following sections.

4.3.1 Set-Up for the Dry CBL

The dry CBL was initialised with no water content, and a mean wind speed of $u = 1\,\mathrm{m\,s^{-1}}$, aligned along the x axis. A constant positive heat flux of $241\,\mathrm{W\,m^{-2}}$ was input at the surface, resulting in a well-mixed BL. A constant potential temperature profile was initialised in the lower layer of the domain, while a strong temperature gradient formed a capping inversion in the upper layer. Details of the initial potential temperature profile are given in Table 4.1.

z (m)	θ (K)
0	300
950	300
1050	308
2000	310.85

Table 4.1: *Summary of initial profile for θ in the dry CBL case. Values at intermediate heights can be found by linear interpolation.*

This case was run at multiple different grid spacings, in order to ascertain the effect of resolution on the model's ability to produce accurate CBL fields. The high resolution LES of the dry CBL are produced on a 4.8 km × 4.8 km horizontal domain, with a vertical depth of 2 km. However, to ensure there were enough grid points to resolve motions in the flow, the very coarsest resolution simulations required the horizontal domain to be expanded. The exact details of the grid length and domain size for each simulation are given in Table 4.2.

$\Delta_{x,y}$ (m)	Horizontal Domain	Δ_z (m)	Vertical Depth
20	4.8 km × 4.8 km	20	2 km
25	4.8 km × 4.8 km	20	2 km
50	4.8 km × 4.8 km	20	2 km
100	4.8 km × 4.8 km	20	2 km
200	9.6 km × 9.6 km	20	2 km
400	9.6 km × 9.6 km	20	2 km
800	19.2 km × 19.2 km	20	2 km

Table 4.2: *The domain size and grid spacings used for each of the dry CBL simulations.*

The MONC model generated 4 hours of data for each resolution of the dry CBL simulation, with the data analysis consistently focusing on the 3 hour and 40 minute timestamp.

4.3.2 Set-Up for the BOMEX case

The MONC model was initialised using a set of conditions to mimic the BOMEX field observations. The resulting simulation is considered to be in quasi-steady state. The BOMEX simulation was run with a 16 km $\times$ 16 km horizontal domain using a grid spacing of $\Delta_{\mathrm{x}} = \Delta_{\mathrm{y}} = 20\,\mathrm{m}$ with periodic boundary conditions imposed on the lateral boundaries. In the vertical, the grid spacing is set to $\Delta_{\mathrm{z}} = 20\,\mathrm{m}$ and the domain extended to 3 km above the surface. The surface pressure p_0 was initialised to 1,000 hPa.

Initial conditions for both the potential temperature θ and water vapour q_v consist of a well-mixed layer extending from the surface to z = 520 m. Above this is the cloud layer, a conditionally unstable region initially set to lie between 520 m and 1500 m. The cloud layer is capped by a statically stable layer, defined by a temperature inversion extending from 1500 m to 2000 m. Above z = 2000 m the statically stable "free atmosphere" extends up to the domain lid at 3000 m. Unidirectional Easterly winds are initialised by setting u to negative values and v to zero throughout the column. In the lowest layer (0 m to 700 m), the wind speed is set to a constant value of $u = -8.75\,\mathrm{m\,s^{-1}}$, and above this, the winds decrease linearly with height to the geostrophic wind, with a value of $u = -4.61\,\mathrm{m\,s^{-1}}$ at z = 3000 m. The initial surface latent heat flux was set to 130.052 $\mathrm{W\,m^{-2}}$, while the sensible heat flux was initialised to 8.04 $\mathrm{W\,m^{-2}}$. These initial profiles are summarised in Table 4.3, where a dashed line indicates that, at that height, a value was not prescribed for that variable, but can instead be calculated by linearly interpolating between the variable values above and below this height.

Height (m)	θ (K)	q_{vapour} ($\mathrm{g\,kg^{-1}}$)	u ($\mathrm{m\,s^{-1}}$)	v ($\mathrm{m\,s^{-1}}$)
0	298.7	17.0	-8.75	0
520	298.7	16.3	-8.75	0
700	-	-	-8.75	0
1500	302.4	10.7	-	0
2000	308.2	4.2	-	0
3000	311.85	3.0	-4.61	0

Table 4.3: *Summary of initial variable profiles for the BOMEX case.*

The MONC model was run for 4 hours for the BOMEX case. However, as this case is in quasi-steady state, much of the data has the same statistical properties once model spin-up is complete. Data from the 4 hour time stamp was chosen to be the focus of this analysis.

4.3.3 Set-Up for the ARM Case

The MONC model was used to produce a 14 hour simulation of the ARM case, beginning shortly after sunrise at 05:30 local time (corresponding to 11:30 UTC), and ending at 19:30 local time. The domain was set to have a grid spacing of $\Delta_\mathrm{x} = \Delta_\mathrm{y} = 25\,\mathrm{m}$ in the horizontal and $\Delta_\mathrm{z} = 10\,\mathrm{m}$ grid spacing in the vertical. This simulation spans a horizontal domain of $19.2\,\mathrm{km} \times 19.2\,\mathrm{km}$ with a vertical depth of $4.4\,\mathrm{km}$. The model was initialised following the set-up detailed in Brown et al. (2002), based on the ARM data recorded at the Southern Great Plains (SGP) field site on the 21st June 1997.

The surface pressure was set to $p_0 = 970\,\mathrm{hPa}$ and random temperature perturbations were imposed at each grid point in the lowest 200 m to initiate turbulence. It should be noted that Brown et al. (2002) suggested some minor modifications to the initial temperature profile due to a mismatch in data between the central facility observations and the diagnosed large-scale forcings. This was believed to be due to the forcings representing averages across the entire $365\,\mathrm{km} \times 300\,\mathrm{km}$ SGP site, rather than being specific to the particular area under analysis. The potential temperature profiles were further modified by increasing the gradient in the inversion layer which overlies the cloud layer (z $\geq 2500\,\mathrm{m}$) to prevent clouds from nearing the domain top (Brown et al., 2002). Details of the initial profiles are given in Table 4.4 below.

z (m)	$\bar{\theta}$ (K)	$\overline{q_v}$ (g kg^{-1})	u, v (m s^{-1})
0	299	15.2	10, 0
50	301.5	15.17	10, 0
350	302.5	14.98	10, 0
650	303.53	14.8	10, 0
700	303.7	14.7	10, 0
1300	307.13	13.5	10, 0
2500	314	3	10, 0
5500	343.2	3	10, 0

Table 4.4: *Summary of initial mean profiles for each variable in the ARM case. At intermediate heights, linear interpolation is used to obtain values for each quantity. Note that the set up of the ARM case in MONC defines profiles up to $z = 5,500\,m$, however this simulation set the top of the domain to be $z = 4,400\,m$, therefore interpolation was used to find the initial profile values at this height.*

Prescribed time-dependent surface fluxes for sensible and latent heat, along with large-scale advective forcings, and radiative tendencies, drove the model to produce an idealised simulation of the ARM case. The boundary conditions, which are also based on observations, are provided to the model as simplified profiles (Chlond et al., 2014). A time series of the surface forcings is detailed in Table 4.5.

Local time	Sensible Heat Flux (W m^{-2})	Latent Heat Flux (W m^{-2})
05:30	-30	5
09:30	90	250
12:00	140	450
13:00	140	500
15:30	100	420
18:00	-10	180
20:00	-10	0

Table 4.5: *Time series of the surface heat fluxes for the ARM case. At intermediate heights, the value of the fluxes can be calculated using linear interpolation.*

The model output fields for the ARM case are analysed with a focus on four time stamps: 10:30L, 12:30L, 14:30L, 16:30L, as illustrated in Figure 4.2. Each of these times corresponds to a certain phase of interest from the diurnal cycle: the early morning dry CBL (DRY), the onset of passive clouds (ONS), the rapid cloud development stage (DEV), and the late afternoon quasi-steady state (QSS).

4.4 Data Post-Processing Procedure

The fields of LES data from the MONC model must undergo a series of post-processing procedures before they can be used to calculate fields of flow-dependent Smagorinsky parameters. Firstly, the reduction in resolved energy as a result of filtering by the Smagorinsky scheme must be accounted for before any other filters can be applied to the MONC output fields. This is to ensure that the correct filter scales are used when applying the Gaussian filter to the data. Furthermore, although the shape of the Smagorinsky filter is similar to that of a Gaussian filter, it is not an exact match. When calculating L_{ij} and M_{ij} using Equations 3.35 and 3.36, the difference between the two scales must only be a result of resolution, not filter shape. To ensure that the difference is solely due to the difference in filter scale rather than differences in filter shape, the raw MONC output data (with grid spacing Δ) must first be filtered using the Gaussian filter to produce the SGS data set. The resulting SGS data has an effective grid spacing of $\overline{\Delta}$. These Gaussian-filtered fields form the “first filter” dataset, and it is used as the base dataset for all dynamic Smagorinsky calculations. This base dataset is then filtered again, using a Gaussian filter, to an effective grid spacing of $\widehat{\overline{\Delta}}$, giving rise to the “second filter” dataset. The differences in fluxes (see Equations 3.35 and 3.42) and gradients (see Equations 3.36 and 3.43) between the SGS and the STS can then be calculated, and the dynamic Smagorinsky parameters computed.

4.4.1 The Gaussian Filter

Data from an LES with a grid spacing Δ can be filtered (smoothed) to a coarser resolution using a Gaussian filter applied to the data in Fourier space. The $\widehat{\ }$ symbol is used to indicate data which has been filtered. Filtered data has a corresponding "effective grid spacing", $\widehat{\Delta}$, which describes the resolution that the filtered data corresponds to in physical space. Data filtered to $\widehat{\Delta}$ is expected to be similar to a MONC data field with a grid length $\Delta = \widehat{\Delta}$, provided it is in the LES regime. Therefore $\widehat{\Delta}$ is a measure of the equivalent "grid spacing" of the filtered data. For MONC model runs, as the grid spacing Δ increases, the simulation will begin to encounter problems associated with the grey-zone and the model will be unable to resolve the correct amount of energy. In contrast, the statistical properties of the filtered data remain unaffected by the grey-zone, even at coarse resolutions, and can therefore be regarded as the "truth", against which lower resolution MONC simulations can be evaluated.

The filtered data in this study is calculated by applying a Gaussian filter to a high-resolution LES dataset. The MONC output data is transformed to spectral space by a Fourier transform operation. The Gaussian filter, $\mathcal{G}$, is then applied to this data in spectral space (denoted by ^{s}) and is defined as:

$$\mathcal{G}^s = \exp(-\frac{\sigma_{\mathrm{f}}^2 k^2}{2}) \tag{4.1}$$

Where k is the wavenumber, and σ_{f} is the filter scale. The σ_{f} variable sets the width of the Gaussian filter, and is related to the effective grid scale by:

$$\widehat{\Delta} = 2\sigma_{\mathrm{f}} \tag{4.2}$$

Note that this relation is derived and discussed further in Chapter 5. The Gaussian filter removes energy from the power spectrum $\mathrm{E_f}(k)$ of a variable f to give $\mathrm{E}_{\hat{\mathrm{f}}(k)}$, the power spectrum of the filtered variable $\widehat{f}$ in spectral space using the convolution theorem:

$$\mathrm{E}_{\hat{\mathrm{f}}(k)} = (\mathcal{G} * \mathrm{E_f})^s = \mathcal{G}^s \otimes \mathrm{E_f}^s = \exp(-\tfrac{1}{2}\sigma^2 k^2)\mathrm{E_f} \tag{4.3}$$

As energy density in the inertial subrange is defined by $\mathrm{E_f} = \alpha\epsilon^{2/3}k^{-5/3}$:

$$\mathrm{E}_{\hat{\mathrm{f}}(k)} = \exp(-\sigma^2 k^2)\alpha\epsilon^{2/3}k^{-5/3} \tag{4.4}$$

It can be seen that this filter has very little effect on low wavenumbers within the ISR, therefore allowing the Gaussian filter to limit the reduction of energy to the small scales, similar to that of the Smagorinsky scheme. Note that, if multiple Gaussian filters are applied to a data set, the Gaussian filter scales σ_i are combined, such that:

$$\sigma_{\mathrm{f}} = \sqrt{\sum_{i=0}^{n} \sigma_i^2} \quad \forall i \tag{4.5}$$

4.4.2 Filter-Scale imposed by Smagorinsky Scheme

As previously mentioned in Section 3, the Smagorinsky scheme imposes a filter on the data. This is because Smagorinsky acts to remove energy from the resolved flow, with a particular focus on the smallest scales, similar to the effect that viscous dissipation has on physical flows. Data from an idealised power spectrum, one with no deviation from the -5/3 Kolmogorov slope in the inertial sub-range, is taken and filtered by a Gaussian filter with filter-scale $\sigma = \frac{1}{2}\Delta$. Results presented in Chapter 5 show that the resulting power spectrum is very similar to the power spectrum produced from the output of a MONC model with a grid spacing of Δ. The exact reasoning behind this relation is discussed further in Section 5.4. This allows for the approximation of the filter scale corresponding to the Smagorinsky subgrid scheme, σ_{smag}.

$$\sigma_{\mathrm{smag}} = \frac{1}{2}\Delta \tag{4.6}$$

where Δ is the LES model grid spacing. The grid spacings and corresponding Smagorinsky filter scales for each case study are given in Table 4.8 below.

Case Study	$\Delta_{\mathrm{x,\,y}}$ (m)	Δ_{z} (m)	σ_{smag} (m)
Dry CBL	20	20	10
BOMEX	20	20	10
ARM	25	10	12.5

Table 4.6: *The grid spacings and corresponding Smagorinsky filter scale for each case study.*

4.4.3 Filtering the LES Output Fields (First Filter)

In order to remove any artefacts that the Smagorinsky filter shape might have imposed on the dataset, a Gaussian filter is applied to the raw MONC output fields. Adhering to convention, the data is filtered to coarser effective grid spacings ($\overline{\Delta}$) by multiplying the grid spacing (Δ) by powers of 2. Note that this spectral filter is only applied to the data at each level of the horizontal planes; data is not filtered in the vertical. The data is coarsened to $\overline{\Delta} = 2\Delta, 4\Delta, 8\Delta, 16\Delta, 32\Delta,$ and 64Δ, which encompasses effective grid scales ranging from typical LES regimes to grey-zone resolutions. To achieve the desired effective grid spacings while also taking the Smagorinsky filter into account, Equation 4.5 can be applied. This allows σ_{f1}, the filter scale for the "first filter" data, to be computed using the first Gaussian filter scale σ_1 and the approximate Smagorinsky filter scale σ_{smag}.

$$\sigma_{\mathrm{f1}} = \sqrt{\sigma_1^2 + \sigma_{\mathrm{smag}}^2} \tag{4.7}$$

Recall from Equation 4.2, for any given Gaussian filter f, with filter scale σ_{f}, the corresponding effective grid spacing Δ_{f} is given by $\Delta_{\mathrm{f}} = 2\sigma_{\mathrm{f}}$. In order to find what σ_1 to use when Gaussian filtering the data to the desired scales, recall that the first filter data requires an effective filter scale of $\Delta_{\mathrm{f}} = \overline{\Delta} \equiv \mathrm{n}\Delta$ where $\mathrm{n} \in \{2^1, 2^2, 2^3, 2^4, 2^5, 2^6\}$. Therefore, combining Equations 4.2 and 4.7 gives:

$$\overline{\Delta} = \mathrm{n}\Delta \equiv 2\sigma_{\mathrm{f1}} = 2\sqrt{\sigma_1^2 + \sigma_{\mathrm{smag}}^2}$$

and so the required Gaussian filter scale σ_1 can be computed:

$$\sigma_1 = \sqrt{\frac{\mathrm{n}^2}{4}\Delta^2 - \sigma_{\mathrm{smag}}^2} \quad \forall\, \mathrm{n} \in \{2^1, 2^2, 2^3, 2^4, 2^5, 2^6\} \tag{4.8}$$

The σ_1 values used to compute the first filter for the BOMEX and ARM cases are given in Table 4.8, along with the corresponding effective grid-scale $\overline{\Delta}$. This "first filter" data forms the base dataset from which the dynamic Smagorinsky parameters will be calculated. The base data has an effective grid spacing of $\overline{\Delta}$ in physical space, where the overbar denotes the first filter, and is no longer contaminated by the misalignment between the shape of the Gaussian and Smagorinsky filters. To compute the Smagorinsky parameters, the base data must be filtered a second time, which allows the differences between the two scales to be calculated following the Germano-Lilly method, as described in Section 3.4.1).

$\overline{\Delta}$ (m)	BOMEX σ_1 ($\Delta = 20\,\text{m}$)	ARM σ_1 ($\Delta = 25\,\text{m}$)
2Δ	17	22
4Δ	39	48
8Δ	79	99
16Δ	159	199
32Δ	320	400
64Δ	640	800

Table 4.7: *The effective grid spacings and corresponding Gaussian filter scale used to calculate the "first filter" base dataset for the BOMEX and ARM case studies.*

4.4.4 Filtering the Filtered Fields (Second Filter)

Following the arguments presented by Germano et al. (1991) and Lilly (1992), when data is resolved down to sufficiently small scales, the properties of turbulence scale with wavelength. Therefore, it is assumed that the appropriate amount of mixing, and thus dissipation, can be determined by looking at the difference between the base "first filtered" dataset (with an effective grid spacing $\overline{\Delta}$, which resolves down to the SGS) and coarser "second filtered" dataset (with an effective grid spacing $\widehat{\overline{\Delta}}$, which resolves down to the STS). The difference between the SGS and STS is scale selective, meaning the test filter window remains the same width for all scales, as depicted in Figure 4.3.

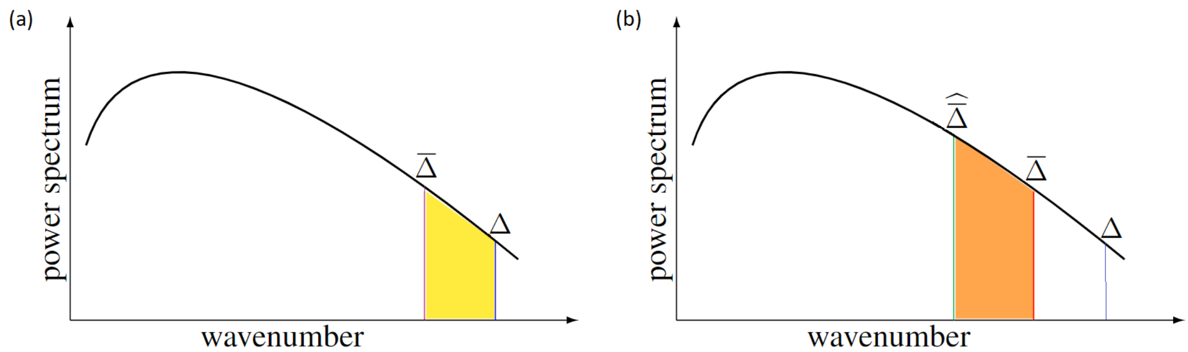

Figure 4.3: *When calculating the scale-invariant Smagorinsky parameter, the filtering test window maintains the same ratio between the base data set and the coarser filtered set on a log axis. The choice of base data depends on the desired scale for the filtered data. The scale of the test windows are shown for data filtered from (a) a base dataset from the raw LES data, with filter scale Δ, which has been filtered to a coarser dataset with filter scale $\overline{\Delta}$, and (b) a base dataset (from the filtered data) with a filter scale $\overline{\Delta}$, which has been filtered a second time to a coarser dataset with filter scale $\widehat{\overline{\Delta}}$.*

The "first filter" data is filtered a second time, with the "second filter" operation denoted by a hat. For the scale-invariant method it is filtered to twice the effective grid spacing of the "first filter" data (ie: four times the grid spacing of the raw MONC output data): $\widehat{\overline{\Delta}} = 2\overline{\Delta} = 4\Delta$. When computing scale-dependent parameters using the method defined by Bou-Zeid et al. (2005) (see Section 3.4.2) a coarser filter scale is applied to the base dataset and the data is filtered to an effective grid spacing of $\widehat{\overline{\Delta}} = 4\overline{\Delta} = 8\Delta$. Although this data is coarser, it has still only been filtered twice, and so is still considered the result of a "second filter". In more general terms, data resulting from a "second filter" requires an effective filter scale of $\Delta_{\mathrm{f}} = \widehat{\overline{\Delta}} \equiv \mathrm{m}_i\Delta$. For the scale-invariant method of calculating the Smagorinsky parameter $\mathrm{m}_1 = 2^2\Delta$, while $\mathrm{m}_2 = 2^3\Delta$ for the scale-dependent calculation of the Smagorinsky parameter. The correct σ values required to compute the "second filter" data can be determined by using Equations 4.2 and 4.7:

$$\widehat{\overline{\Delta}} = \mathrm{m}_i\Delta \equiv 2\sigma_{\mathrm{f2}} = 2\sqrt{\sigma_2^2 + \sigma_1^2 + \sigma_{\mathrm{smag}}^2} \tag{4.9}$$

From Equation 4.8, it is clear that $\sigma_1^2 + \sigma_{\mathrm{smag}}^2 = \frac{\mathrm{n}^2}{4}\Delta^2$ and so:

$$\frac{\mathrm{m}_i^2}{4}\Delta^2 = \frac{\mathrm{n}^2}{4}\Delta^2 + \sigma_2^2$$

The required Gaussian filter scale σ_2 can then be computed:

$$\sigma_2 = \frac{\Delta}{2}\sqrt{\mathrm{m}_i^2 - \mathrm{n}^2} \quad \text{for } \mathrm{m}_1 = 2^2 \text{ and } \mathrm{m}_2 = 2^3 \tag{4.10}$$

Note that this can be further simplified by recalling the relation between n and m_i. For $\widehat{\overline{\Delta}} = 4\Delta$ where $\mathrm{m}_i = \mathrm{m}_1 = 2\mathrm{n}$, and so:

$$\sigma_2 = \frac{\sqrt{3}}{4}\mathrm{n}\Delta \tag{4.11}$$

Meanwhile, for scale-dependent computations which require an even coarser scale, the effective grid scale $\widehat{\overline{\Delta}} = 8\Delta$ with $\mathrm{m}_i = \mathrm{m}_2 = 4\mathrm{n}$, and so:

$$\sigma_2 = \frac{\sqrt{15}}{4}\mathrm{n}\Delta \tag{4.12}$$

The σ_2 values used to compute the second filter data for the BOMEX and ARM cases are given in the following table, along with the corresponding effective grid-scale $\widehat{\overline{\Delta}}$:

Base dataset used: $\overline{\Delta}$ (m)	Filtered to: $\widehat{\overline{\Delta}}$ (m)	BOMEX ($\Delta = 20\,\mathrm{m}$)		ARM ($\Delta = 25\,\mathrm{m}$)	
		σ_1	σ_2	σ_1	σ_2
2Δ	4Δ	17	35	22	43
2Δ	8Δ	17	77	22	97
4Δ	8Δ	39	69	48	87
8Δ	16Δ	79	139	99	173
16Δ	32Δ	159	277	199	346
32Δ	64Δ	320	554	400	693
64Δ	128Δ	640	1109	800	1386

Table 4.8: *The effective grid spacings and corresponding Gaussian filter scale used to calculate the "second filter" coarse dataset for the BOMEX and ARM case studies. The second row of data in maroon shows the filter scales used to compute the "second filter" for the scale-dependent dynamic computations where* $\widehat{\overline{\Delta}} = 4\overline{\Delta}$. *This is in contrast to all the other "second filter" computations, which were for the scale-invariant method, and filtered data to* $\widehat{\overline{\Delta}} = 2\overline{\Delta}$.

4.5 The Offline Dynamic Model

These "first" and "second" filter data sets can then be used to calculate flow-dependent Smagorinsky parameters using equations from the dynamic model (see Section 3.4.1). As the equations are being applied during the post-processing procedure and not within the MONC model itself, this is referred to as the "Offline Dynamic Model". Note that in this study, the offline dynamic model does not include stability functions in the dynamic equations; however, the LES fields that the offline model uses as input data have been produced using a standard Smagorinsky scheme which included stability functions, as detailed in Section 3.1.3. The equations in Section 3.4.1 can also be included in the subgrid scheme of an LES model, meaning that calculations of the Smagorinsky parameters are performed by the model at each grid point for every time step. This method is referred to as the "Online Dynamic Smagorinsky Model". The online model requires substantial computational power. This study uses the offline model as an analysis tool, applying the dynamic Smagorinsky equations to velocity and scalar fields which have already been output from an LES model and filtered to the required resolutions. Offline analysis is used to diagnose relationships between the Smagorinsky parameters and key features of the flow, such as the ML capping inversion or clouds. Identified relationships will be parametrized and used to to run the MONC model at coarse resolutions in Chapter 8.

4.5.1 Computing Flow Dependent Smagorinsky Parameters Offline

By applying the dynamic equations to the LES data, fields of flow-dependent Smagorinsky parameters can be computed. However, the parameter value fields need additional processing before the results can be utilised. Firstly, as seen in other studies such as Shi et al. (2018) and Efstathiou (2023), the C_s^2 and C_ψ^2 parameters can be negative. Retaining these negative results will contaminate the averages, as a negative C_s or C_ψ yields a negative mixing length, which is inherently unrealistic and indicates that the assumptions of the Smagorinsky scheme have broken down locally. Recalling the equations for C_s and C_ψ (see Equations 3.41 and 3.45), it is evident that only the numerator can be negative. Therefore, these negative values are "clipped", meaning they are set to zero, to indicate a complete lack of subfilter mixing of the quantity at that specific point. Then the numerator and denominators can be horizontally averaged, as denoted by the angle brackets in the equations below, to produce profiles of C_s and C_ψ:

$$\langle C_s^2 \rangle = \frac{1}{2}\left(\frac{\langle \max(0, L_{ij}M_{ij}) \rangle}{\langle M_{kl}^2 \rangle}\right) \tag{4.13}$$

$$\langle C_\psi^2 \rangle = \frac{1}{2}\left(\frac{\langle \max(0, H_j R_j) \rangle}{\langle R_k^2 \rangle}\right) \tag{4.14}$$

The angle brackets denote the averaging operation, which can be conditional or across the entire domain, and it is performed after the indices of the tensors are contracted. The averaging operator used throughout this study is an average at each vertical level to produce profiles of the Smagorinsky parameter. The averaging operations have been applied to the numerator and denominator separately due to the use of the least squares approach when calculating the Smagorinsky parameters, as previously discussed in Section 3.4.2.

Analysing fields from only a single time stamp may introduce noise into the profiles, due to insufficient data for the averaging operation. To mitigate this, data from neighbouring time stamps can also be included in the averaging process. If the additional output time stamp is close to the original, particularly when the system is in or near a quasi-steady state, the additional data from the new output time stamp can be considered an extension of the original domain. As a result, the number of points being averaged over in both the numerator and denominator of the C_s or C_ψ calculation increases, leading to less noisy, and more reliable profiles.

This time-averaging was applied to both the BOMEX and ARM data. Both cases had instantaneous data output every 10 minutes. As the BOMEX case is in quasi-steady state, three time steps were used to produce the time-averaged profiles for C_s and C_ψ, for each scalar ψ. However as ARM is an evolving case, only two timesteps were used to produce the time-averaged profiles for the Smagorinsky parameters.

The flow can be partitioned into different regions, and for each region, the average values of the Smagorinsky parameter at each level can be calculated. One regions of interest is the non-cloudy environment, which is made up of the ML and the non-cloudy areas of the CL. The in-cloud averages at each height level are another region of interest, which can also be further sub divided into the cloud updraft and cloud core regions. In-cloud areas are defined as points where $q_{cl} \geq 10^{-7}\,\mathrm{kg\,kg^{-1}}$, the cloud updraft region is defined as points where $q_{cl} \geq 10^{-7}\,\mathrm{kg\,kg^{-1}}$ and $\mathrm{w}' \geq 0.5\,\mathrm{m\,s^{-1}}$, while the cloud core partition is defined as points where $q_{cl} \geq 10^{-7}\,\mathrm{kg\,kg^{-1}}$, $\mathrm{w}' \geq 0.5\,\mathrm{m\,s^{-1}}$, and $\theta_v' \geq 0\,\mathrm{K}$. To compute the profiles for each partition, Equations 4.13 and 4.14 are again applied, but only using $L_{ij}M_{ij}$, M_{ij}^2, H_jR_j, and R_j^2 values from points in the flow within these regions; i.e. points which fulfil the specific partition's criteria.

The conditional averages for a partition can also be computed; for example, the average value of the Smagorinsky parameter at each height level in-cloud, or the average value at each level of the ML. For this computation, the profile of a Smagorinsky parameter in a certain region r: $\langle C_s \rangle_r$ or $\langle C_\psi \rangle_r$, can be calculated by following the averaging procedure in Equations 4.13 and 4.14, ensuring the numerator and denominator are averaged separately, while only using points within the region in question. However, to avoid boundary effects, the lowest 10% of height levels and the highest 10% of height levels are excluded from this calculation. This gives the average parameter value throughout the entire depth of a distinct layer. This calculation can be performed for the parameters at each filter scale, and the overall trend of average C_s and C_ψ versus filter scale $\widehat{\widehat{\Delta}}$ can be determined.

While the C_ψ parameter is being used as the standard measurement of scalar dissipation throughout this work, the convention in the literature is to use a Prandtl number Pr for heat variables, and a Schmidt number Sc_ψ for any other scalar ψ. Multiple variables are considered, including conserved variables such as θ_L, θ_e, and q_t, along with other non-conserved variables such as θ and θ_v. In order to calculate the corresponding Prandtl and Schmidt numbers to enable direct comparison to previous literature, the following formulas were applied to the Smagorinsky parameters which had been calculated by the offline dynamic analysis:

$$\mathrm{Pr} = \frac{C_s^2}{C_\theta^2} \tag{4.15}$$

$$\mathrm{Sc}_\psi = \frac{C_s^2}{C_\psi^2} \tag{4.16}$$

Any averages that are required to produce profiles of the Prandtl or Schmidt numbers are computed by inputting the corresponding $\langle C_s \rangle$ and $\langle C_\psi \rangle$ average values, rather than applying the averaging operation to a field of Pr or Sc values.

5 Does Smagorinsky Behave as a Filter?

The Smagorinsky subgrid scheme is often regarded as a filtering operation, which is applied to LES data as it is being resolved (Mason and Brown, 1999). This is because the scheme dissipates energy from motions with length scales comparable to the grid scale. As a result, it has become common practice to use the terms “subgrid” and “subfilter” interchangeably, based on the assumption that they scale similarly.

Throughout this work, it is assumed that the Gaussian filter behaves similarly to the Smagorinsky scheme. Gaussian-filtered data is used as a proxy for the expected output of an LES model, which uses the Smagorinsky subgrid scheme, run with a coarser grid spacing. In cases where this scale falls within the grey-zone, the Gaussian-filtered data is considered the idealised outcome that the LES would have produced, had it been unaffected by grey-zone limitations. Therefore it is important to determine the corresponding “effective grid scale” ($\widehat{\bar{\Delta}}$) of the filtered data in the grid-based domain after applying a Gaussian filter with filter scale σ to the data in Fourier space. This enables both the comparison of LES output data with filtered data, and the calculation of dynamic Smagorinsky parameter fields using the method outlined in Section 3.4.1.

This chapter will examine whether the Smagorinsky scheme truly operates as a filter, and if so how its behaviour compares to that of a Gaussian filter, by addressing the following research questions:

- Does the Smagorinsky scheme demonstrate a specific filter shape in spectral space that maintains the same shape consistently for every grid spacing and case study? If so, does this filter shape scale uniformly with grid spacing?
- If the Smagorinsky scheme functions as a filter, how closely does its filter shape correspond to that of a Gaussian filter in spectral space?
- How can the Gaussian filtered data be related back to grid-based domain?

To address these questions, the data must first be transformed into Fourier space to evaluate the behaviour of its spectra. The work in this section builds upon work carried out by Moeng and Wyngaard (1988) to find the filter shape of the Smagorinsky subgrid scheme.

5.1 Method of Computing Dimensional Spectra

Two-dimensional spectra of the vertical wind velocity and the vertical kinematic eddy heat flux were computed at various levels of interest in the boundary layer using the method outlined in Durran et al. (2017). These spectra plots have a normalisation factor, $\mathbb{N}$, given by:

$$\mathbb{N} = \frac{\Delta_x \Delta_y \min(\Delta k_x, \Delta k_y)}{8\pi^2 \mathrm{n}_x \mathrm{n}_y}, \tag{5.1}$$

where $\Delta k_{x_i} = \dfrac{2\pi}{\Delta_{x_i}\, \mathrm{n}_{x_i}}$ for $x_i \in (x, y)$. The number of points in the x and y direction is given by n_x and n_y respectively, and the grid spacing in the x and y direction denoted by Δ_x and Δ_y. This method was employed to compute spectra using both LES output fields and filtered data, and the results of this are discussed in the following sections.

5.2 Shape of Smagorinsky Filter in Spectral Space

Spectra for the vertical velocity, $\mathcal{S}_\mathrm{w}$, have been computed using data from the mid-CBL height level of the dry CBL simulations at various grid spacings (Δ). The results are plotted in Figure 5.1, along with an idealised spectrum of what data from an unfiltered CTBL should look like, used as a reference.

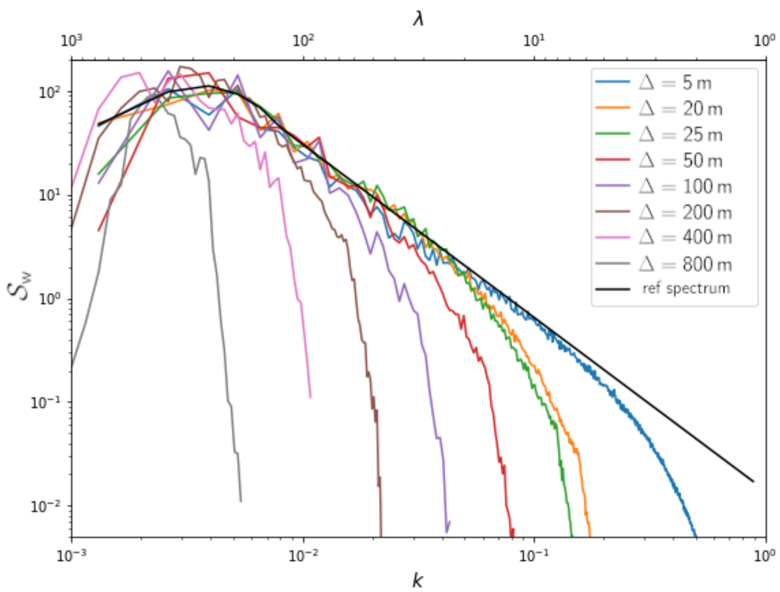

Figure 5.1: *Power spectrum of w from the middle of the CBL plotted against wavenumber k ($k = 2\pi/\lambda$), and an additional axis showing wave length λ. The "truth" spectrum is shown by the black line.*

The w spectra for the higher resolution LES runs show peaks at wavelength values just over $\lambda = 1000\,\mathrm{m}$, and a clearly defined inertial subrange is evident with the turbulent cascade of energy following the -5/3 Kolmogorov law. The slopes of the high-resolution spectra (when $\Delta \geq 100\,\mathrm{m}$) are not strongly impacted by the grey-zone. Therefore their ISR follows a -5/3 slope before deviating to a steeper slope as the model fails to resolve the energy accurately at the smallest scales in the simulation. This deviation defines the effective resolution of the model, R_{eff}. Any scales of motion within the flow which are smaller than R_{eff} cannot be fully resolved. Figure 5.1 demonstrates a steep decrease in the spectral slope after the $2\overline{\Delta}$ mark as energy cannot be resolved past this point. The 2Δ point demonstrates the Nyquist frequency, where the model cannot resolve motions using less than 2 grid spaces.

Meanwhile, the coarser resolution LES runs, where the model is in a grey-zone regime, show peaks occurring at higher wavelengths with no clear inertial subrange evident. In the case of the coarser resolution MONC simulations of the dry CBL, the dominant scales of motion begin to have the same scale as the R_{eff}. Therefore the model is unable to fully resolve these motions, resulting in the model being unable to resolve scales in ISR, instead the spectra show the steeper slope which is related to the Smagorinsky scheme and model dynamics immediately after the spectral peak. At the very coarse resolutions, such as $\Delta = 800\,\mathrm{m}$, it is obvious that the model is no longer able to fully resolve the energy-containing scales. This is evident from the decrease in the energy value reached at the peak, and the width of the peak itself. Therefore, at these grid spacings, the grey-zone has truly begun to impact the resolvable scales.

The spectra were then scaled by an idealised ISR slope following Kolmogorov's -5/3 law. This more clearly presents when each spectrum deviates from the idealised truth, i.e. at which wavelength each run begins to under-resolve energy. The result is presented in Figure 5.2 a, while Figure 5.2 b depicts the result while scaling the wavenumber by the grid spacing.

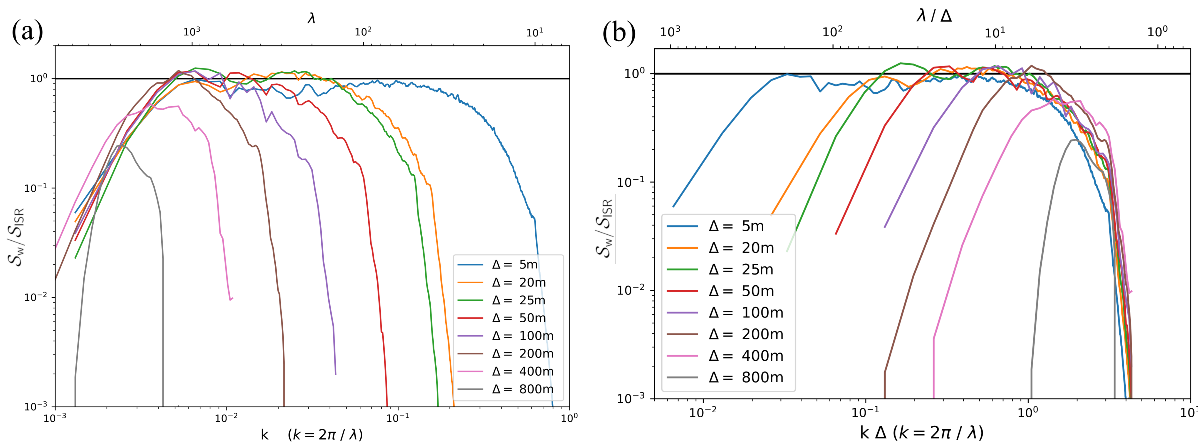

Figure 5.2: *Power spectra of w at the mid-CBL level from MONC simulations with different grid spacings Δ, scaled by the idealised -5/3 ISR slope plotted against (a) wavenumber k ($k = 2\pi/\lambda$), with an additional axis for wavelength λ, and (b) k scaled by Δ, with an additional axis for λ/Δ. The black line shows the total idealised energy, as determined by the ISR slope, to serve as a reference.*

Figure 5.2 a shows the ratio of resolved energy to the idealised energy amount plotted against wave number. This plot clearly shows when this ratio deviates from unity: ie. when the resolved energy in the model is less than the reference. Figure 5.2 a demonstrates that the MONC model has an effective resolution $R_{\text{eff}} \approx 8\overline{\Delta}$, where the spectral slope deviates from unity until reaching the $2\overline{\Delta}$ mark, after which it falls at a much steeper rate as energy cannot be resolved past this point. The slope between R_{eff} and 2Δ is a result of the model dynamics and the Smagorinsky scheme. The lengthscale corresponding to R_{eff} can be expressed as the dissipation length scale, which is discussed further in Section 5.3.3. Figure 5.2 b demonstrates that, by scaling the wavelength by the grid spacing relative to each spectrum, all the spectra collapse along the same slope after their corresponding R_{eff} point, which is observed occurring at $\lambda/\Delta \approx 8$, as expected. Furthermore Figure 5.2 b clearly shows the rapid decrease in energy when $\lambda/\Delta \leq 2$, corresponding to the Nyquist frequency. This plot explicitly illustrates that the standard Smagorinsky subgrid scheme has a universal filter shape, valid at every grid spacing, whether in the grey-zone regime or not.

$\mathcal{S}_{\text{w}}$, the power spectrum of w, can be related to w'^2 variance values via the following equation:

$$\overline{\text{w}'^2} = \int_{-\infty}^{\infty} \mathcal{S}_{\text{w}}\, \mathrm{d}k = 2\int_{0}^{\infty} \mathcal{S}_{\text{w}}\, \mathrm{d}k \tag{5.2}$$

$$\implies \overline{\text{w}'^2} \approx 2\int_{k_{\text{min}}}^{k_{\text{max}}} \mathcal{S}_{\text{w}}\, \mathrm{d}k \tag{5.3}$$

Each spectrum was integrated up to the wavenumber corresponding to 2Δ, and the result was scaled by the integral of the truth spectrum (also up to 2Δ). This results in the ratio of resolved energy to idealised total energy, which was then plotted against 2Δ scaled by the inversion height, $z_i = 1080$ m. Note that in Figure 5.3 the $\Delta = 200$ m, 400 m and 800 m runs have larger domains, resulting in an artificial increase in the resolved energy for the LES of the dry CBL at these grid spacings.

Figure 5.3 shows the amount of energy being resolved by each MONC run. As each energy is then scaled by an equivalent truth value for the energy, simulations that are not in the grey-zone are seen to have an energy ratio of around 0.9. This demonstrates that in the LES regime, the MONC model is able to resolve approximately 90% of the total energy. Note that the $\Delta = 5$ m run shows less energy being resolved than the $\Delta = 20$ m, 25 m, and 50 m runs. This is due to the use of an FFT solver for the $\Delta = 5$ m run, whereas all other simulations use an iterative solver. It was necessary to use the FFT solver for the $\Delta = 5$ m run as, at the time of implementation, there was a bug in the MONC model which prevented very high resolutions from running with the iterative solver, and once this bug was resolved, the simulation was not re-run with an iterative solver due to the computational expense of a simulation with $\Delta = 5$ m run. The $\Delta = 20$ m, 25 m, and 50 m runs are clustered together showing they are resolving about the expected amount of energy for their given resolution, indicating these simulations are not heavily impacted by any grey-zone issues. However, the $\Delta = 100$ m shows a slight deviation from the higher resolution data, indicating the onset of the grey-zone. Meanwhile, the $\Delta = 200$ m run is resolving more energy than the other MONC simulations, though this is probably due to the compensating effect of a larger domain. The MONC runs with $\Delta \geq 400$ m clearly show a decrease in resolved energy, demonstrating the transition to the grey-zone.

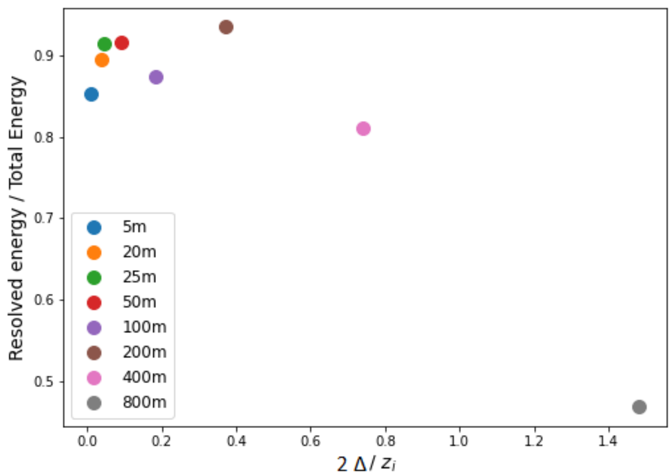

Figure 5.3: *Integral of the power spectra of w scaled by the "true spectrum", plotted against* $2\Delta/z_i$.

5.3 The Gaussian Filter in Spectral Space

The defining equation of the Gaussian filter allows it to reduce energy mostly at the small scales (Equation 4.1 demonstrates that this filter primarily affects large wavenumbers), similar to that of the Smagorinsky scheme. The extent of this similarity will be investigated throughout this section. To compare the Gaussian filter, with its filter scale defined by σ in Fourier space, to the Smagorinsky filter, which has its filter scale defined by Δ in physical space, the relationship between σ and the effective grid scale Δ_{eff} must first be derived. Note that the effective grid scales focused on in this study correspond to data which has been filtered twice using the Gaussian filter, and thus $\Delta_{\text{eff}} = \widehat{\widehat{\Delta}}$ in this case. Before this relation is deduced, the effect of applying a Gaussian filter to LES data will be analysed.

5.3.1 Gaussian Filtered Spectra

The spectra of w and cospectra of the vertical heat flux from the LES output data with a grid spacing of $\Delta = 5\,\mathrm{m}$ are plotted in Figures 5.4 a & b respectively. This $\Delta = 5\,\mathrm{m}$ dry CBL data was filtered to various resolutions using the Gaussian filter with multiple different σ values, and the spectra and cospectra for these filtered fields are also plotted in these figures. It is evident that, as the filter scale increases and the resolution becomes coarser, the dissipation of energy from the data increases on a scale proportional to the filter scale. This results in the filtered spectra appearing to move to the left as the filter scale σ increases. Lower values of σ show the corresponding spectra in Figure 5.4 a maintaining their $k^{-5/3}$ Kolmogorov slope in the ISR. Spectra in the dry CBL have been observed to follow this law, with numerous studies showing this behaviour, including those by Wyngaard and Coté (1971), Kaimal et al. (1972), and Honnert et al. (2020). Meanwhile, Figure 5.4 b shows the cospectra of vertical heat flux maintaining a -7/3 power law when σ is small, similar to cospectra of vertical heat flux from Moraes et al. (2008) and Kaimal and Finnigan (1994). However, when large filter scales are applied to the data, the Gaussian filter begins to heavily affect the energy-containing scales at the spectral peak, and scales in the ISR are no longer resolved fully meaning the Kolmogorov law no longer holds. At very large filter scale values, the Gaussian filter begins to dissipate energy excessively from the large scale motions and the spectral peak sees a reduction in both height and width.

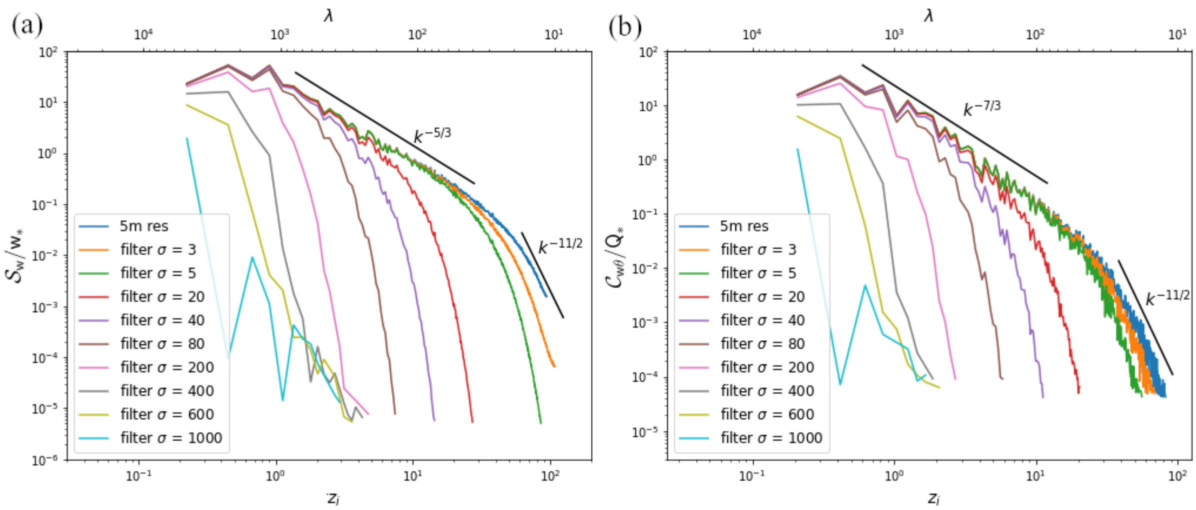

Figure 5.4: *(a) Spectra of the w values from the middle of the dry CBL, filtered by the Gaussian filter to various resolutions and scaled by w_*. (b) Cospectra of the vertical heat flux values from the middle of the CBL, scaled by $\mathcal{Q}_*$. Both plots have wavenumber k ($k = 1/\lambda$) on the x-axis, and an additional axis showing wave length λ. Various power laws are shown by the black lines.*

The spectra have been scaled by w^*, the free-convection scaling velocity, while the cospectra have been scaled by $\mathcal{Q}^*$, the constant surface buoyancy flux, given by:

$$\mathrm{w}^* = \left[\frac{g z_t}{\theta_\mathrm{s}} \, \mathcal{Q}^*\right]^{1/3} \tag{5.4}$$

$$\mathcal{Q}^* = \frac{\mathrm{H}}{\rho_0 c_p} = 0.24\,\mathrm{K\,m\,s^{-1}} \tag{5.5}$$

Where $g = 9.81\,\mathrm{m\,s^{-1}}$ is the acceleration due to gravity, θ_s is the temperature at the surface, $\mathrm{H} =\ 241\ \mathrm{W\,m^{-2}}$ is the heat flux at the surface, $\rho_\mathrm{air} \approx 1\,\mathrm{kg\,m^3}$ is the density of air, and $c_p = 1004\,\mathrm{J\,kg^{-1}\,K^{-1}}$ is the specific heat of air at constant pressure.

The Gaussian filter primarily dissipates energy from the smallest scales. After a point known as the dissipation length scale, this enhanced dissipation from the filter results in a change from the $k^{-5/3}$ Kolmogorov law (for spectra) or a $k^{-7/3}$ law (for cospectra) to a steeper power law that can be approximated by $k^{-11/2}$ for low σ values. However this is not an exact power law, and the $k^{-11/2}$ approximation does not hold true for all sigma values, because the Gaussian filter is not linear. The exact effect of this filter on the power spectra is calculated in Section 5.3.4.

5.3.2 Effective Grid Scale of the Gaussian Filter

The filter scale of the Gaussian filter is determined by the σ value, with larger σ values producing smoother, coarser resolution fields. In order to compare the filtered data to the LES runs, the σ parameter must be related back to a corresponding grid-based scale, and so an appropriate relation must be derived. This scale is defined as the "effective grid spacing", and is denoted by Δ_eff. The $\widehat{\overline{\Delta}}$ scales, which correspond to specific σ values, can be determined by examining the Gaussian equation in Fourier space. The Gaussian filter reduces energy by a factor of $\exp(-\sigma^2 k^2)$ where wavenumber $k = (2\pi)/\lambda$. It is desirable to have the Gaussian filter acting on the smallest scale where very little energy is contained, chosen here to be less than 1.8% of the energy for ease of computation. Therefore, if we assume that only 1.8% of the total energy in the filtered spectrum is contained at scales affected by the Gaussian filter, this corresponds to the filter acting upon scales with a wavenumber of $k \geq 2/\sigma$ or greater:

$$e^{-\sigma^2 k^2} = e^{-\sigma^2 (2/\sigma)^2} = e^{-4} = 1.8\% \tag{5.6}$$

The value of 1.8% was chosen as it leads to a simple equation for k_max while upholding the condition that k_max marks the wavenumber where it is assumed that the majority of energy, in this case 98.8%, is well-resolved. Therefore, the grid scale which corresponds to this k_max defines the scale that this filtered data would have in a physical grid space.

As very little energy is being resolved when $k = 2/\sigma$, this point in the filtered spectra can be assumed approximately analogous to the Nyquist frequency in the spectrum of grid-discretised data. Therefore, it is assumed that the k corresponding to the wavelength where the Nyquist frequency occurs, that is $\lambda_{\text{min}} = 2\Delta$, can be related to the k_{max} where just 1.8% of the total energy in a Gaussian filtered spectrum remains. Recalling the relationship between wavelength λ and wavenumber k:

$$\lambda = \frac{2\pi}{k_{\text{max}}} \tag{5.7}$$

$$\implies 2\Delta_{\text{eff}} = \lambda_{\text{min}} = \frac{2\pi}{k_{\text{max}}} \tag{5.8}$$

By relating the k_{max} of the LES data in discretised space to the k_{max} of the Gaussian filtered data in Fourier space, using the finding from Equation 5.6:

$$k_{\text{max}} = \frac{2}{\sigma} \tag{5.9}$$

Then from Equation 5.8 and 5.9

$$\implies \Delta_{\text{eff}} = \frac{\pi}{2}\sigma \approx 2\sigma \tag{5.10}$$

Therefore, by assuming that the smallest scales in the model contain only 1.8% of the energy, the effective grid spacing of each filter scale σ value can be estimated. This relationship between the Gaussian filter scale and the effective grid-scale $\Delta_{\text{eff}} = \widehat{\overline{\Delta}}$ is supported by experiments carried out in Section 5.4 to verify this formula. These experiments compared the energy spectra of filtered data, defined by an effective grid spacing of $\widehat{\overline{\Delta}}$, with those of LES data on a grid with spacing Δ, under the condition that $\widehat{\overline{\Delta}} = \Delta$. At sufficiently high resolution, where both filtered and LES datasets resolve the dominant turbulent scales, the resulting spectra show strong agreement as illustrated in Figure 5.6. These findings support the validity of using Equation 5.10 to relate the Gaussian filtered fields to their equivalent grid-based representation.

5.3.3 Dissipation Length Scale

From the spectra in Figures 5.1 and 5.4, an increase in the dissipation of energy is seen occurring at lower wavenumbers as the resolution is coarsened. This corresponds to a "dissipation lengthscale", below which the numerical dissipation results in the spectra falling faster than the $k^{-5/3}$ Kolmogorov law. Beare (2014) defined a dissipation length scale, ℓ_d, which is calculated using the second moment of the turbulent kinetic energy spectrum, $\mathcal{S}_{\mathrm{e}}$:

$$k_d = \sqrt{\frac{\int_{k_0}^{k_1} k^2 \mathcal{S}_{\mathrm{e}} dk}{\int_{k_0}^{k_1} \mathcal{S}_{\mathrm{e}} dk}} \tag{5.11}$$

$$\ell_d = \frac{k_d}{2\pi} \tag{5.12}$$

where k_0 is the smallest wavenumber, corresponding to the largest length scale, here $k_0 = \lambda_{\mathrm{max}}$, and k_1 is the largest wavenumber. The dissipation length scale corresponds to the wavelength in the power spectrum where excessive dissipation begins to take over. It can be used as an indication of the grey-zone: if ℓ_d is of the same order as the dominant coherent structures in the flow, then these eddies are not being well resolved. The dissipation length scale has been calculated for each filter scale of the Gaussian filtered dry CBL data, and the resulting plot is presented in Figure 5.5.

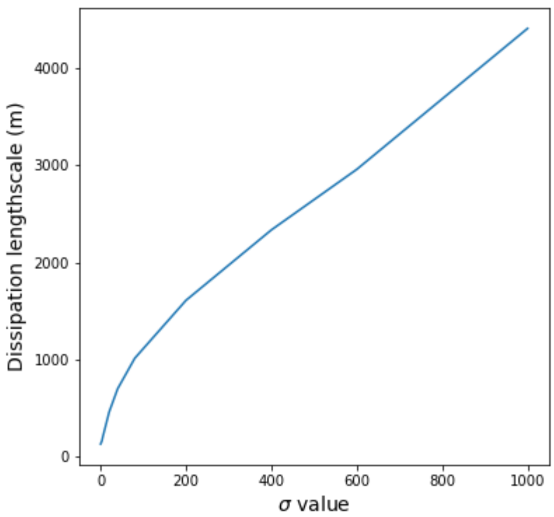

Figure 5.5: *Filter scale, given by σ, versus the dissipation length scale ℓ_d for the power spectrum of the w values in the middle of the dry CBL.*

Figure 5.5 outlines how, at lower value filter scales σ, the dissipation lengthscale is observed to increase rapidly as σ increases. However, as the filter scale goes to large values, approximately $\sigma \geq 200$, the increase in dissipation length scale, ℓ_d, slows. At these larger filter scales ℓ_d increases much more gradually, at a more linear rate. From Section 5.3.2, it is clear that $\sigma = 200$ corresponds to an effective grid scale of $\widehat{\Delta} = 400\,\mathrm{m}$. As previously discussed, this scale is fully within the grey-zone regime, and the change in the slope of ℓ_d with respect to σ at this scale is likely due to the transition from the LES regime to the grey-zone. This hypothesis is further supported by Figures 5.4 a & b, which show that both the spectra and cospectra begin to deviate from the common peak only when $\sigma \geq 200\,\mathrm{m}$.

5.3.4 Slope of the Gaussian-Filtered Spectra

The slope that the spectrum follows past the point that corresponds to ℓ_d can also be calculated. The equation for the power spectrum of Gaussian filtered data, from Equation 4.4 (see Section 4.4.1 for further details), is defined as.

$$\mathrm{E}_{\hat{\mathrm{f}}(k)} = \exp(-\sigma^2 k^2)\alpha\epsilon^{2/3}k^{-5/3} \tag{5.13}$$

When the filtered energy density $\mathrm{E}_{\hat{f}}$ is plotted against wavenumber k on a logarithmic graph, for wavenumbers greater than k_d the slope can be determined by:

$$\begin{aligned} \log(e^{-\sigma^2 k^2} k^{-5/3}) &= -\sigma^2 k^2 - \tfrac{5}{3}\log(k) \\ \implies & [-\sigma^2 k^2 (\log(k))^{-1} - \tfrac{5}{3}]\log(k) \end{aligned} \tag{5.14}$$

Therefore, the slope of the filtered spectra to the right of the dissipation length scale no longer follows an exact power law, and this slope gets steeper as the filter scale (determined by σ) increases. This is further confirmed by Figure 5.2, which shows that the Gaussian filter not only dissipates energy at the small scales, but also reduces energy at large wavelengths, especially for large σ values.

5.4 Comparison of Smagorinsky to the Gaussian Filter in Spectral Space

LES data from the dry CBL case with a grid spacing of $\Delta = 25\,\mathrm{m}$ has been filtered using the Gaussian filter with various different values for σ, as listed in the legend of Figure 5.6. The spectra from both LES and filtered data can be compared using the effective grid spacing relation derived in Section 5.3.2. Plotting the spectra of both datasets enables testing of this relationship, as the spectra of LES data should resemble the spectra of the filtered data when $\Delta = \widehat{\Delta}$. This test is carried out using data from the mid-CBL of the dry CBL case and the resulting plot is presented in Figure 5.6.

The dissipation scheme used in the MONC LES is the Smagorinsky scheme, and from this spectral analysis in Figure 5.6, it is clear that the Smagorinsky scheme behaves similarly to a

Gaussian filter. When both the LES and filtered spectra are at reasonably high resolution, such as $\Delta = 50\,\text{m}$ and $\sigma = 25\,\text{m}$ ($\widehat{\widehat{\Delta}} = 2\sigma = 50\,\text{m}$), their spectra are very similar. This similarity holds for both high and low wavelengths, though the Gaussian filter maintains a smooth slope throughout, while the Smagorinsky spectra show a discontinuity at $\lambda = 2\Delta$, corresponding to the Nyquist frequency. Beyond this point, the Smagorinsky filter exhibits a steeper slope than the Gaussian filter. However, this has minimal impact on the total energy shown by the spectra, as this difference occurs at small wavelengths where very little energy is present.

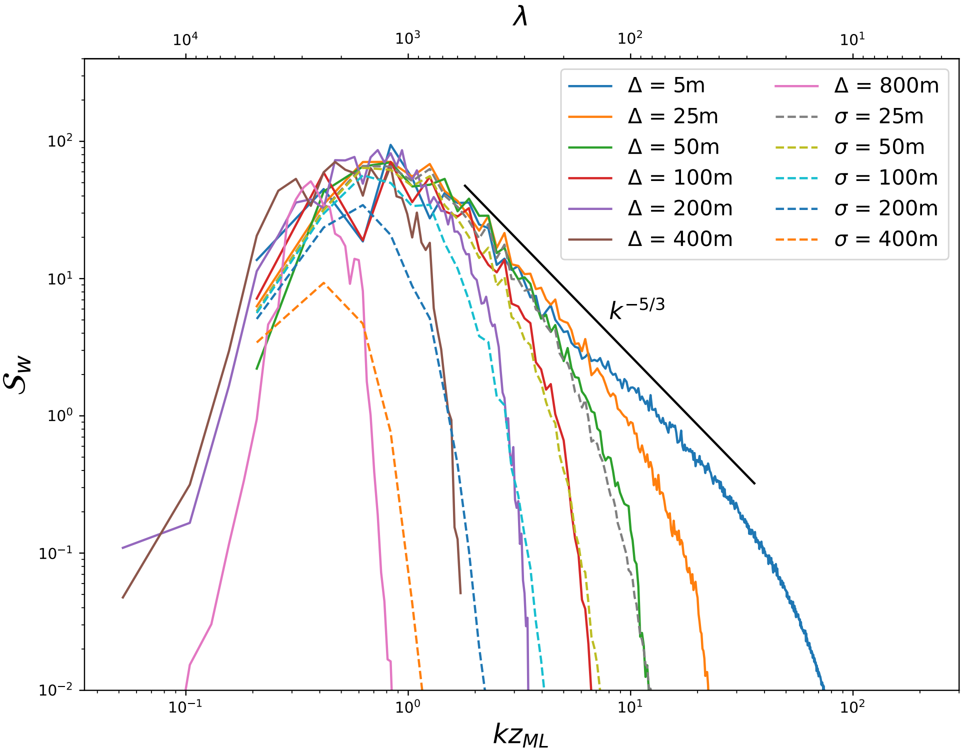

Figure 5.6: *Spectra of the mid-ML level of the w data from the dry CBL case. The LES data is plotted in the solid line, while the dashed lines denote the spectra from the filtered data. The x-axes are the same as Figure 5.4, while an idealised Kolmogorov law is indicated by the black line. The filter scales can be related back to the grid-based space using* $\widehat{\widehat{\Delta}} = 2\sigma$.

A large difference in energy between the corresponding LES and filtered spectra is evident at very coarse resolutions. It appears that when σ is large, the Gaussian filter affects the large-scale features. This is clear by the reduction in the height of the spectral peak for the $\sigma = 200\,\text{m}$, and especially clear in the $\sigma = 400\,\text{m}$ data where an order of a magnitude difference is seen between the peak of the LES and filtered spectra. While both the Smagorinsky scheme and Gaussian-filtered data are affected by grey-zone issues at coarse scales, the Gaussian filter allows the grey-zone to influence larger scales more readily. Recalling the argument in Section 5.3.2, the filtering of large-scale structures results from the k_{max} wavenumber tending toward the wavenumber value corresponding to these dominant structures, k_{i}. Therefore, the assumption that only a negligible amount of energy resides at wavenumbers greater than or equal to k_{max}

becomes invalid, and the Gaussian filter primarily acts on motions with wavelengths $k_{\max} \approx k_{\mathrm{i}}$ or greater, therefore impacting the large-scale components of the flow as well as the smaller scale eddies. It is important to note that, in this example, the Smagorinsky scheme is applied to data computed on a domain four times larger than that used for the Gaussian-filtered data (see Table 4.2). Therefore, this larger domain may enhance the Smagorinsky scheme's ability to retain energy at larger scales, a benefit which is not afforded to the Gaussian-filtered data.

The spectra of the filtered data can also be used to verify that the properties of the Gaussian filter are consistent within the framework employed to compute the spectra. The commutative property of the filter scales can be assessed by applying two of the same filters, separately in different orders, to the same data and then plotting the resulting spectra. Data from the ARM case is used to produce Figure 5.7 to ensure similar behaviour of spectra exists across the case studies.

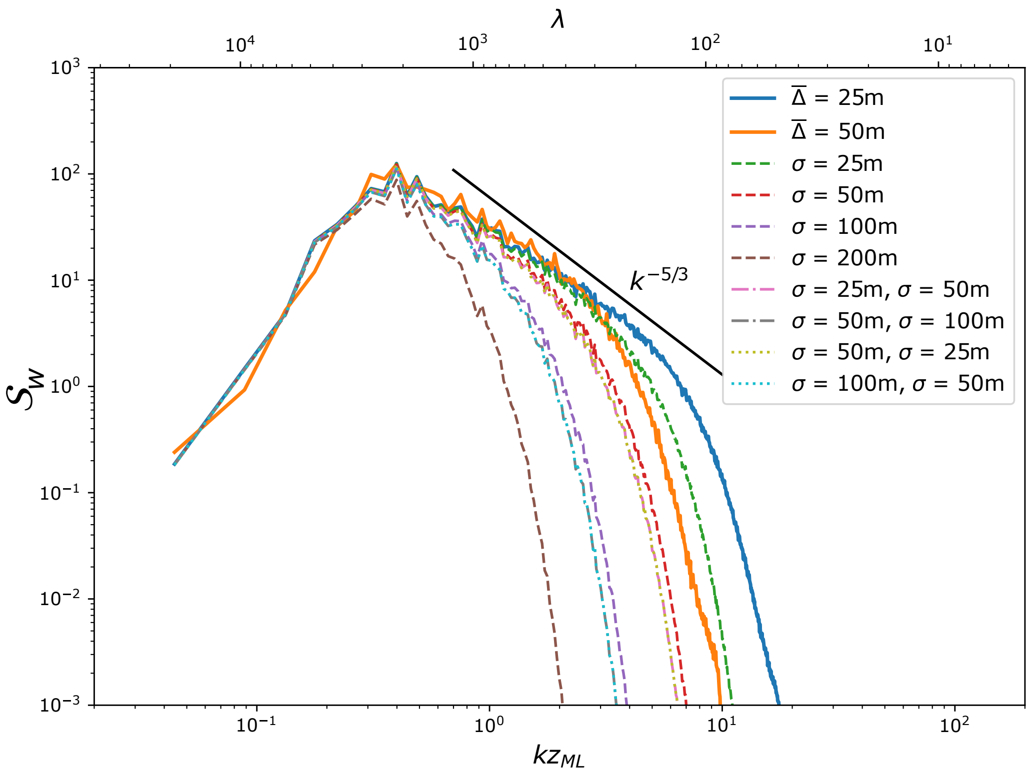

Figure 5.7: *Spectra of w from the mid-CBL of the ARM case at 14:30L, plotted against wavenumber. The solid lines show the unfiltered LES output with grid spacings. The dashed lines show the spectra resulting from filtering the* $\Delta = 25\,m$ *to coarser scales by imposing a Gaussian filter on the data with filter scale* σ*. The dash-dot and dotted lines show spectra which has been filtered a second time.*

Firstly, it is evident from Figure 5.7 that the spectra from the ARM case follows similar trends as those from the dry CBL case (Figure 5.6); at large wave lengths, the spectra from the $\Delta = 50\,\mathrm{m}$ LES output for the ARM case shows similar values as the spectra from the filtered data with

$\sigma = 25\,\mathrm{m}$ (with a corresponding effective grid scale of $\widehat{\Delta} = 2\sigma = 50\,\mathrm{m}$). This indicates that, at large wavelengths, the Smagorinsky filter exhibits a similar behaviour to that of a Gaussian filter, but differing by a factor of approximately two. A more detailed discussion is provided in Section 5.4.1 It should be noted that the BOMEX case was also found to produce similar spectra plots, though these results are not presented here. Secondly, the commutative properties of the Gaussian filter are evident in Figure 5.7, with the spectra from the $\sigma = 25\,\mathrm{m}$, $\sigma = 50\,\mathrm{m}$ data being completely overlaid by the $\sigma = 50\,\mathrm{m}$, $\sigma = 25\,\mathrm{m}$ data. This results in the beige coloured spectra just to the left of the $\sigma = 50\,\mathrm{m}$ data. A similar result is observed for the $\sigma = 50\,\mathrm{m}$, $\sigma = 100\,\mathrm{m}$ data and $\sigma = 100\,\mathrm{m}$, $\sigma = 50\,\mathrm{m}$ data shown by the blue-grey spectra to the left of the $\sigma = 100\,\mathrm{m}$ spectra. Therefore, the property of commutativity holds when applying two different Gaussian filters to a dataset.

5.4.1 Estimating the Smagorinsky Filter Scale

The Smagorinsky scheme also acts as a filter, removing energy at small scales. Therefore Smagorinsky has an associated filter scale σ_{smag}, and this should be accounted for when calculating the effective grid spacing. By likening the Smagorinsky scheme's behaviour to that of a Gaussian filter, a similar method as that used in Section 5.3.2 can be employed to calculate the filter scale imposed on the LES run by the Smagorinsky subgrid scheme. Using a similar method as that employed throughout Section 5.3.2, assuming that Smagorinsky only effects the smallest scales where less that 1.8% of the energy is contained, and recalling Equation 5.6, it can be assumed that:

$$e^{-\sigma_{\text{smag}}^2 k^2} = e^{-4}$$

$$\implies -\sigma_{\text{smag}}^2 k^2 = -\sigma_{\text{smag}}^2 \left(\frac{2\pi}{\lambda}\right)^2 = -\sigma_{\text{smag}}^2 \left(\frac{2\pi}{2\Delta}\right)^2 = -4$$

$$\implies \sigma_{\text{smag}} = \frac{2}{\pi}\Delta \approx \frac{1}{2}\Delta \tag{5.15}$$

This estimation of the Smagorinsky filter scale relies on the assumption that it is similar to the Gaussian filter, and therefore the Gaussian equations can be used to approximate Smagorinsky behaviour. From the spectra in Figure 5.6, it was confirmed that at high resolution, the Smagorinsky filters behaved very similarly to the Gaussian filters in Spectral space, therefore this approximation of the Smagorinsky filter scale is presumed to be valid.

5.5 Summary and Discussion

The findings from this chapter indicate that the Smagorinsky subgrid scheme functions as a grid-scale-based filter. Throughout this chapter, the filter scale, along with the filter shape in spectral space, for the Smagorinsky subgrid scheme has been identified. Furthermore, the Smagorinsky filter was found to exhibit similar behaviours as the Gaussian filter. The Gaussian filter, meanwhile, has been linked to the grid-based domain through a relationship between the filter scale σ and the corresponding effective grid-scale Δ_{eff} which in the case of the data analysed in this work, which has been filtered by two separate Gaussian filters, $\Delta_{\text{eff}} \equiv \widehat{\overline{\Delta}}$. The work carried out in this chapter provided answers to the previously posed research questions.

5.5.1 Responses to the Research Questions

(1) Does the Smagorinsky scheme demonstrate a specific filter shape in spectral space that maintains the same shape consistently for every grid spacing and case study? If so, does this filter shape scale uniformly with grid spacing?

The Smagorinsky scheme shows a distinct shape in spectral space, between the length scale corresponding to effective resolution $R_{\text{eff}} \approx 8\Delta$, and the scale corresponding to the Nyquist frequency 2Δ. The filter maintains the same shape across all data analysed, regardless of the grid scale. Furthermore, the Smagorinsky filter appears to scale linearly with grid spacing.

(2) If the Smagorinsky scheme functions as a filter, how closely does its filter shape correspond to that of a Gaussian filter in spectral space?

The shape of the Smagorinsky filter was found to closely resemble that of the Gaussian filter in spectral space. The greatest similarities were observed at high resolutions, while the resemblances diminished slightly as the resolution decreased to grey-zone scales. Nonetheless, the similarities between the two filters enabled the estimation of a filter scale for the Smagorinsky scheme. Furthermore, the resemblance between the two filters at corresponding scales confirmed that the relationship between σ and Δ, which had been derived from approximations, was indeed valid.

(3) How can the Gaussian filtered data be related back to grid-based domain?

The Gaussian filter scale can be related back to the grid-based domain by the following relation:

$$\widehat{\overline{\Delta}} = 2\sigma$$

This relationship was derived based on assumptions regarding the energy distribution at the smallest scales of the flow, and was found to effectively associate the Smagorinsky filter to the Gaussian filter. When comparing the spectra of both filters, a high degree of agreement was observed when this function was applied to relate their respective scales.

6 How Turbulent Mixing Differs Across Distinct Flow Regimes

Turbulent mixing in LES models governs the rate of energy dissipation from the resolved scale to the SGS (unresolved) scale. The mixing length is determined by the grid spacing and the Smagorinsky parameter. In standard models, the Smagorinsky parameters are set to fixed constants. This approach does not allow the Smagorinsky parameter to adapt in order to account for the well-known fact that turbulent mixing varies within different parts of the flow. For instance, mixing in-cloud is expected to be much more intense than mixing in the non-cloudy environment of the cloud layer. Additionally, in these models, the Smagorinsky parameter for all scalars depends on the momentum parameter C_s, using the Prandtl number as a constant of proportionality. Consequently, turbulent mixing is forced to depend on fixed constants rather than the scalar-specific behaviour within the flow.

In contrast, the dynamic model allows the Smagorinsky parameters to adapt to changes in the flow regime. The dynamic model also has the benefit of enabling variable-dependent parameters to be calculated. The Germano identity and equations from the dynamic model can be applied to high-resolution LES of CTBLs to investigate if key dependencies of the Smagorinsky parameters, and thus the turbulent mixing, can be identified. Using high-resolution data allows for the identification of these relationships when the flow is well-resolved. These relationships can then be adapted and scaled to coarser grid spacings which are traditionally considered grey-zone resolutions, with the aim of improving the accuracy of coarse, grey-zone resolution LES of CTBLs.

This chapter will analyse the behaviours of the Smagorinsky parameters for the variables of interest which are transported and diffused in the flow. Fields of independent Smagorinsky parameter values are calculated for each variable, and their responses to different flow regimes are assessed. The research questions focused on in this chapter are as follows:

- Does the Smagorinsky parameter exhibit different systematic behaviours, depending on the flow regime? Can these behaviours be quantified into general relationships between the regime and the parameter value?
- Are there significant differences between the Smagorinsky parameters for momentum, heat, and moisture in the CTBL?
- Considering that the conservation properties of a scalar affect its corresponding Smagorinsky parameter value, which thermodynamic variable would be most appropriate for the new Smagorinsky based parametrization?

To begin addressing these questions, the dynamic Smagorinsky equations were applied to high-resolution LES of the BOMEX and ARM case studies. Cross-sections of the resulting fields of

C^2 values are presented in the following section. Although the dry CBL case was also analysed, its results are not presented here as the behaviour of C_s and C_θ is very similar to what is observed in the ML of the BOMEX and ARM cases. Additionally, numerous previous studies have already examined LES of dry CBLs at grey-zone resolutions (Efstathiou and Beare, 2015; Shin and Hong, 2015; Zhou et al., 2014; Ito et al., 2015; Kealy et al., 2019; Honnert et al., 2011) and the results from the dry CBL are in keeping with the findings in these studies. Since the primary goal of this work is to develop a parametrization scheme for CTBLs, the focus will now shift to the cases with shallow cumulus clouds present.

6.1 High-Resolution Fields of Dynamic Smagorinsky Parameters

The first step in developing a new parametrization scheme for the mixing length is to determine the flow-dependent fields of Smagorinsky parameters. This is achieved by applying equations from the dynamic model, as outlined in Section 3.4.1, to filtered LES of the BOMEX and ARM cases. The C^2 fields, i.e. fields of the square of the Smagorinsky parameters for momentum, heat, and moisture, were calculated using Equations 3.40 and 3.44. Vertical cross-sections through subsets of the C^2 domains are plotted in Figures 6.1 and 6.2. Corresponding vertical cross-sections through the C fields are presented in Figures 6.3 and 6.4. While the plots in this Section only depict a single snapshot of the fields at a specific time, they have been confirmed to be representative of the entire dataset through visual inspection of multiple snapshots with coordinates varying in both space and time.

6.1.1 Fields of Dynamic C^2 Values

Both the BOMEX (Figure 6.1 a) and ARM (Figure 6.2 a) cases reveal that the C_s^2 fields experience numerous small areas of interchanging positive and negative values throughout both the ML and CL. Meanwhile, the C_θ (Figures 6.1 b & 6.2 b) and C_{q_t} (Figures 6.1 c & 6.2 c) fields show larger and more coherent areas of alternating positive and negative areas compared to the C_s fields. The positive areas of the C_θ and C_{q_t} fields are more abundant than negative areas in all regions apart from the in-cloud (IC) regions of the C_θ field. This suggests that the structures transporting, mixing, and dissipating heat and moisture are larger than the structures for momentum. The majority of non-zero values in the Smagorinsky parameter fields for scalars are confined to occur in the ML thermals, IC, and near the cloud edges in the cloud-free environment (CFE). The majority of the areas outside the clouds in the CL depict near-zero values for scalars, however this is not the case for C_s. The C_s field exhibits non-zero values throughout the ML and CL.

A substantial number of areas have negative C^2 values in both the BOMEX (Figure 6.1) and ARM (Figure 6.2) cases, though to different extents depending on the variable in question. Negative values occur in areas of counter-gradient transport, and the dynamic model produces negative parameters in an attempt to represent the corresponding upscale transport of energy in these regions. However, using negative C^2 values in the Smagorinsky model would yield negative mixing lengths, which is both unrealistic and impractical. These values are therefore "clipped" to zero, though their locations relative to other CTBL features are noteworthy.

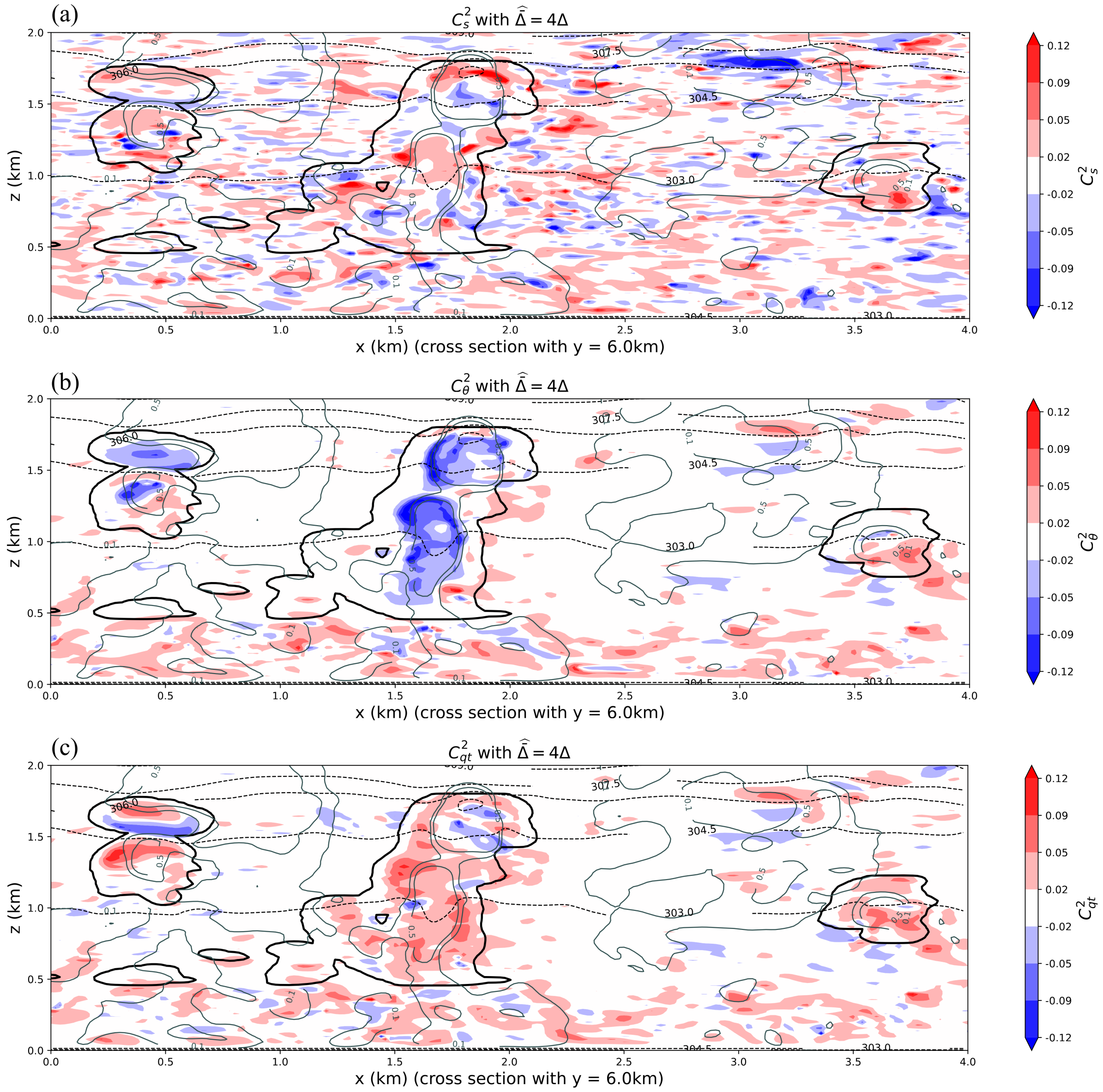

Figure 6.1: *Vertical cross sections through fields from the BOMEX case of the dynamically computed Smagorinsky parameters squared, (a) C_s^2, (b) C_θ^2, and (c) C_{qt}^2, before negative values are removed. These fields have been calculated using data that has been filtered to $\widehat{\bar{\Delta}} = 4\Delta$. Solid black contours denote the presence of clouds (areas where the liquid water content $q_{cl} \geq 10^{-7}\,kg\,kg^{-1}$), the dashed black lines show the θ_v levels, and the grey contours outline areas where vertical velocity $w' \geq 0.1\,m\,s^{-1}$ and $w' \geq 0.5\,m\,s^{-1}$.*

Focusing first on the C_s^2 fields, both the BOMEX and ARM cases portray this as quite a noisy field. However, this field is not without structure, as the extreme C_s^2 values appear to systematically align with the vertical velocity (w′) contours in cross-sections from both cases. This is consistent with the theory, as regions with significant wind speed gradients are expected to experience elevated levels of shear, stress, and strain. Since the dynamic Smagorinsky parameters are functions of stress and strain, higher values of these variables correspond to larger Smagorinsky parameter values. Figures 6.1 a and 6.2 a depict noisy C_s^2 fields, however the MLs (the layer below $z/z_{\mathrm{ML}} = 1$) are primarily positive. In cloud, there are also large consistent areas of positive parameter values, with both cases showing high values focused at the edges of the cloud updraft (as indicated by the w′ contours) and at the cloud top. Furthermore, Figure 6.1 a highlights that the region of elevated C_s^2 values at the top of the actively growing cloud, located at the centre of the BOMEX cross-section, corresponds to a parcel of higher θ_v that has been entrained into the cloud from the overlying warm, dry layer. This is expected, as the entrainment of high thermal energy air would enhance turbulent mixing in that area.

In the CFE of the CL, C_s^2 values are lower, with small sporadic positive and negative patches mainly concentrated near the cloud edges. This is particularly evident in the BOMEX case, which has a mean wind of $u = 8.75\,\mathrm{m\,s^{-1}}$, with Figure 6.1 a depicting higher C_s^2 values on the windward (right-hand) side of the active cumulus cloud. Meanwhile, the ARM case was initialised with a mean wind of $u = 10\,\mathrm{m\,s^{-1}}$ However, there are large coherent areas of positive C_s^2 values in the near-surface areas of the thermals, as evident in the lower left-hand side of Figure 6.2 a. This is likely due to the large sensible heat flux input at the surface, which gives rise to large-scale mixing in these thermals.

Focusing now on the C_θ^2 and C_{qt}^2 fields presented in Figures 6.1 b & c (BOMEX) and 6.2 b&c (ARM). Negative values of the Smagorinsky parameter for both heat and moisture appear to occur at the cloud base of older, dissipating clouds in the ARM case, as seen in the central, passive cloud in Figure 6.2 b & c. This area of negative C_θ^2 and C_{qt}^2 aligns with the incursion of the higher $\theta_v = 312\,\mathrm{K}$ layer, which brings warmer, drier air down from the overlying environment. These negative values can be seen extending downward from the passive, central cloud in the cross-section of the ARM case. This is in contrast to the cloud top of the actively growing cumulus in the BOMEX case, which shows elevated C_s^2 and C_{qt}^2 values (Figures 6.3 a & c respectively). Here, a bubble of higher value θ_v has been entrained into this cloud top, resulting in large, positive values of C_s^2 and C_{qt}^2 parameters in this region, as the entrainment causes an increase in the levels of scalar mixing.

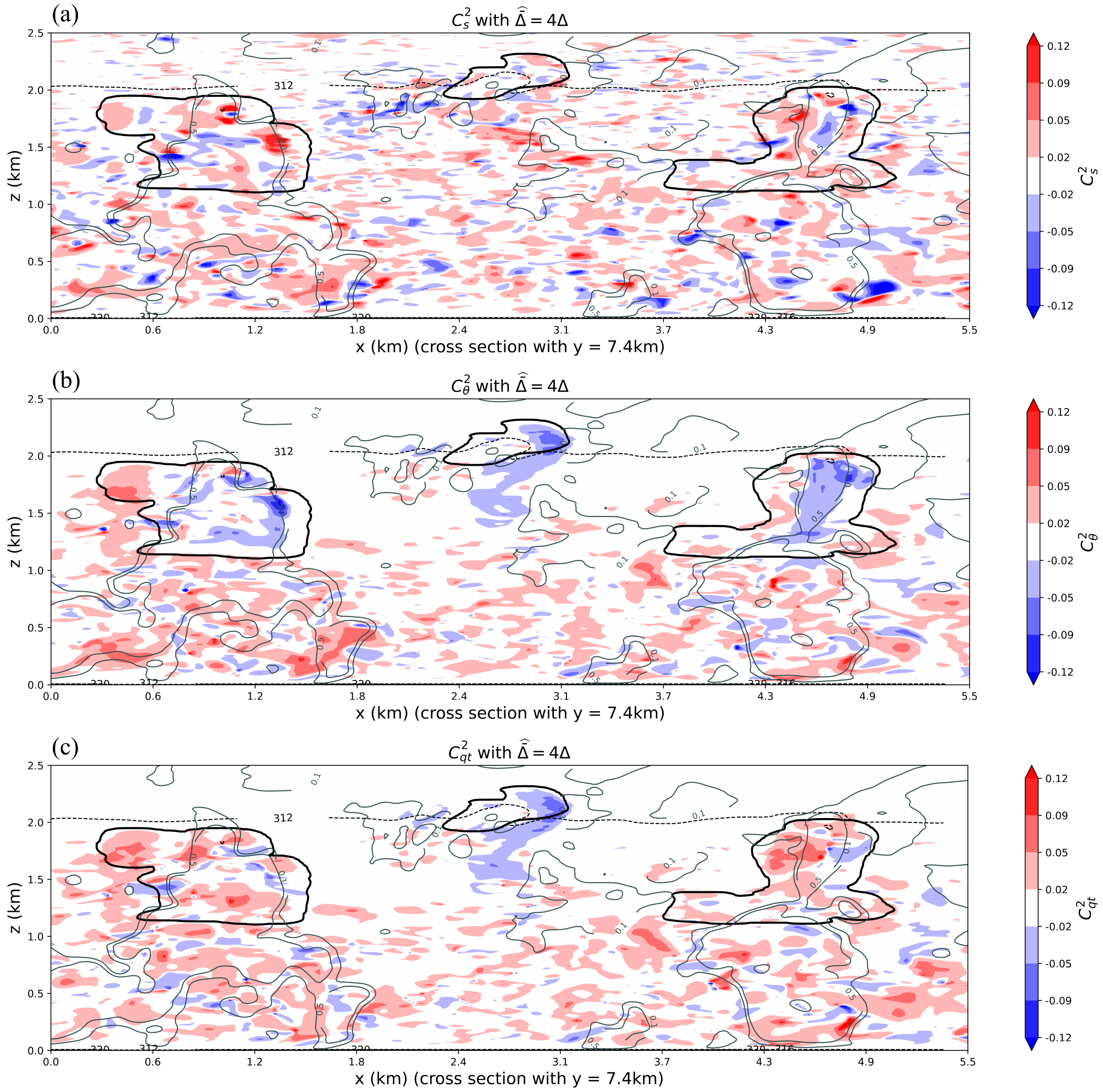

Figure 6.2: *Same as figure 6.1, but for ARM data at 14:30 L, filtered to* $\widehat{\widehat{\Delta}} = 4\Delta$.

While the field of momentum parameters C_s^2 (Figures 6.1 a & 6.2 a) shows numerous small areas of negative values, the moisture parameter C_{qt}^2 displays only a few such regions, with these negative values also being less pronounced (see Figures 6.1 c & 6.2 c). The potential temperature fields from both the BOMEX (Figure 6.1 b) and ARM (Figure 6.2 b) cases reveal that the majority of negative C_θ^2 values are concentrated in the centre of the clouds, or the "cloud core" region. In contrast to this, the moisture field shows mostly positive $C_{q_t}^2$ values in-cloud, particularly in actively growing cumulus clouds. The surrounding CFE of the CL has near-zero values for both the C_θ and C_{q_t} parameters, however this same behaviour is not evident in the CFE of the C_s^2 parameter field. The ML of both the BOMEX and ARM cases show similar behaviours for all Smagorinsky parameters, with regions of negative values aligning, and similar alignment of the positive value areas. This is especially evident in the ML of the C_θ and C_{q_t} parameter fields, with Figures 6.3 b & c and 6.2 b & c showing similar features throughout the MLs.

The large, actively growing clouds show thermals feeding these clouds, evident from the w$'$ $\geq$ $0.5\,\mathrm{m\,s^{-1}}$ contours. Examples of this include the central cloud in the BOMEX domain (Figure 6.3) and the clouds on the left and right-hand sides of the ARM domain (Figure 6.2). These thermal updrafts transport heat and moisture from the ML into the CL, enabling the formation and growth of the cumulus clouds. This upward motion can be seen extending from the surface and being pushed to the left in the BOMEX plots, following the direction of the prevailing easterlies, until it reaches the cloud base. On the right-hand side (RHS) of the active BOMEX cloud, Figure 6.3 b & c show positive C_θ^2 and C_{qt}^2 values just outside the cloud on the windward side. This is likely due to the mixing of heat and moisture into the non-cloudy environment by smaller eddies that form as they detach from the main overturning circulation due to wind shear. Meanwhile, the leeward side of this cloud shows no such trend, with Smagorinsky parameter values being near-zero here.

6.1.2 Dynamic Smagorinsky Parameter Fields

While negative values are useful for highlighting differences between the parameter fields, such results are unrealistic and the Smagorinsky model would be unable to deal with negative mixing length values. Therefore, negative C^2 values are set to zero, completely inhibiting turbulent mixing and dissipation from occurring in the region. The more conventional Smagorinsky parameter C can then be computed for each variable of interest using Equations 3.41 and 3.45 and the resulting fields can be analysed.

The coloured contours in Figures 6.3 and 6.4 highlight regions in the CTBL with high levels of mixing, which leads to greater dissipation of energy from the resolved scale to the subgrid scale (SGS). Figure 6.3 a depicts the largest values of C_s occurring at the top of the actively growing cumulus clouds in the BOMEX case (between $x = 1.5\,\mathrm{km}$ and $2\,\mathrm{km}$ at $z \approx 1.75\,\mathrm{km}$). High C_s values are also observed within the same large BOMEX cloud at mid-cloud levels, particularly along the periphery of the updraft, as indicated by the positive w$'$ contours. In contrast, Figure 6.2 a shows that the ARM case does not exhibit similarly extensive regions of elevated C_s values near the cloud top. However, areas near the edges of the cloud core still provide clear evidence of significant momentum mixing.

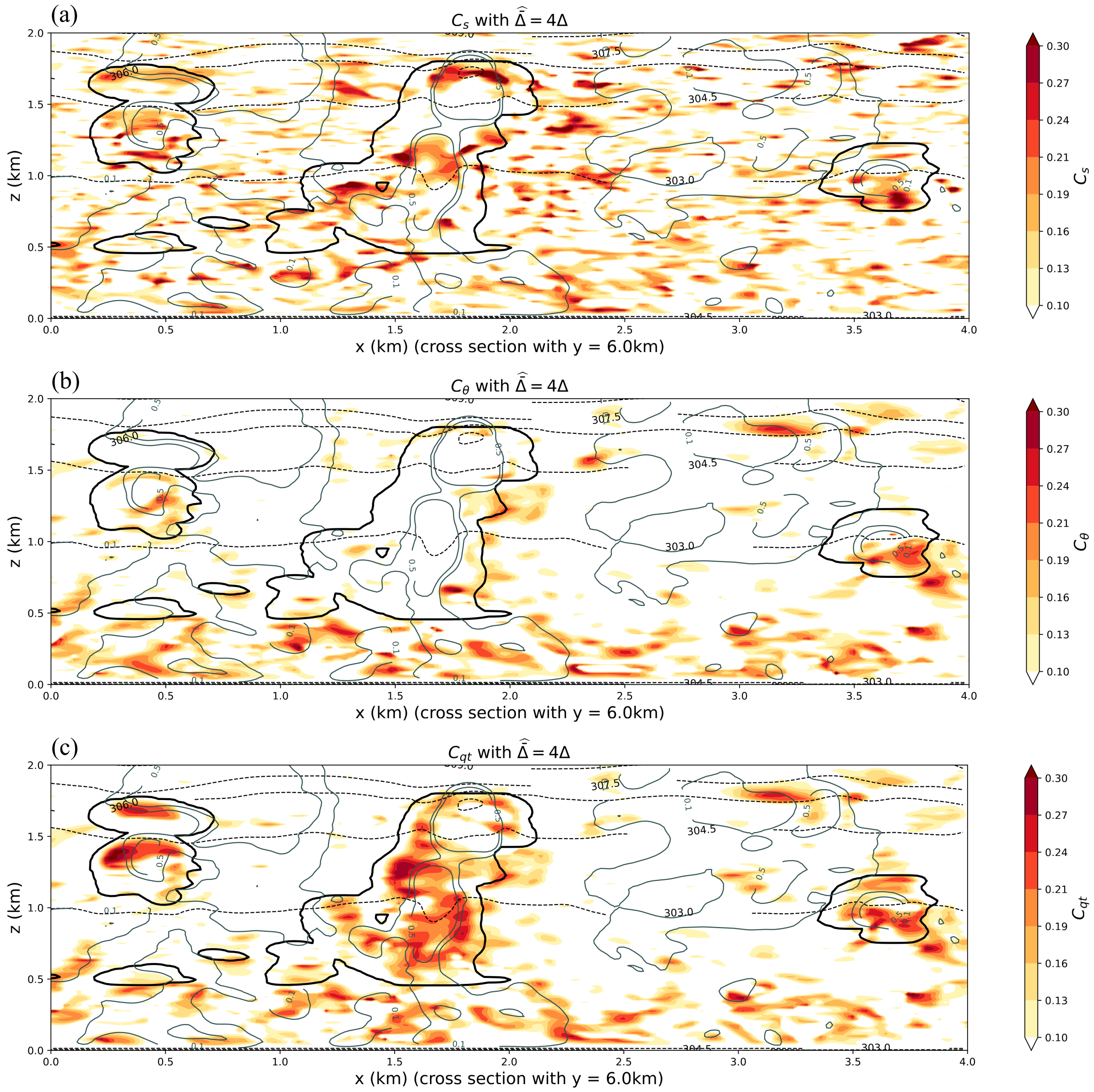

Figure 6.3: *Vertical cross sections through fields of the dynamically computed Smagorinsky parameters for the BOMEX case, (a) C_s, (b) C_θ, and (c) C_{qt}, calculated using data that has been filtered to $\widehat{\overline{\Delta}} = 4\Delta$. The same contours as in Figure 6.1 are used here.*

The scalar fields do not show this alignment of enhanced mixing along thermal edges. Further to this, each parameter shows different behaviours in the cloudy regions, depending on its corresponding scalar. While this highlights the difference between the mixing of heat and moisture, the discrepancy is primarily due to differences in the conservation properties of θ (unconserved) and q_t (conserved) when water condenses in the cumulus clouds. The C_θ parameter exhibits minimal to no mixing within the clouds (see Figures 6.3 b and 6.4 b), though this is assumed to be an artefact of θ being non-conserved within the clouds. Meanwhile, the C_{qt} parameter shows the highest levels of mixing occurring within the cloudy areas (as shown in Figures 6.3c and 6.4 c). In the ML, the fluxes driving the thermals appear to influence the scalar-specific Smagorinsky parameters. The BOMEX case in Figure 6.3 c, with its high latent heat fluxes, shows similar fields in the ML for both C_{q_t} and C_θ. The BOMEX case also shows higher C_{q_t} values in-cloud compared to the in-cloud values of the ARM case in Figure 6.4 c. This discrepancy between the two cases is likely due to ARM's significantly lower latent heat fluxes compared to BOMEX, as ARM is primarily driven by sensible heat fluxes. This is also apparent in the ML of the ARM case, with Figure 6.4 c demonstrating lower in-thermal values of C_{q_t} compared to the C_θ field, as depicted in Figure 6.4 b. This is further evidenced by the higher C_θ values for the ARM case (Figure 6.4 b) compared to the BOMEX case (Figure 6.3 b) when looking in the thermals of the lower level ML.

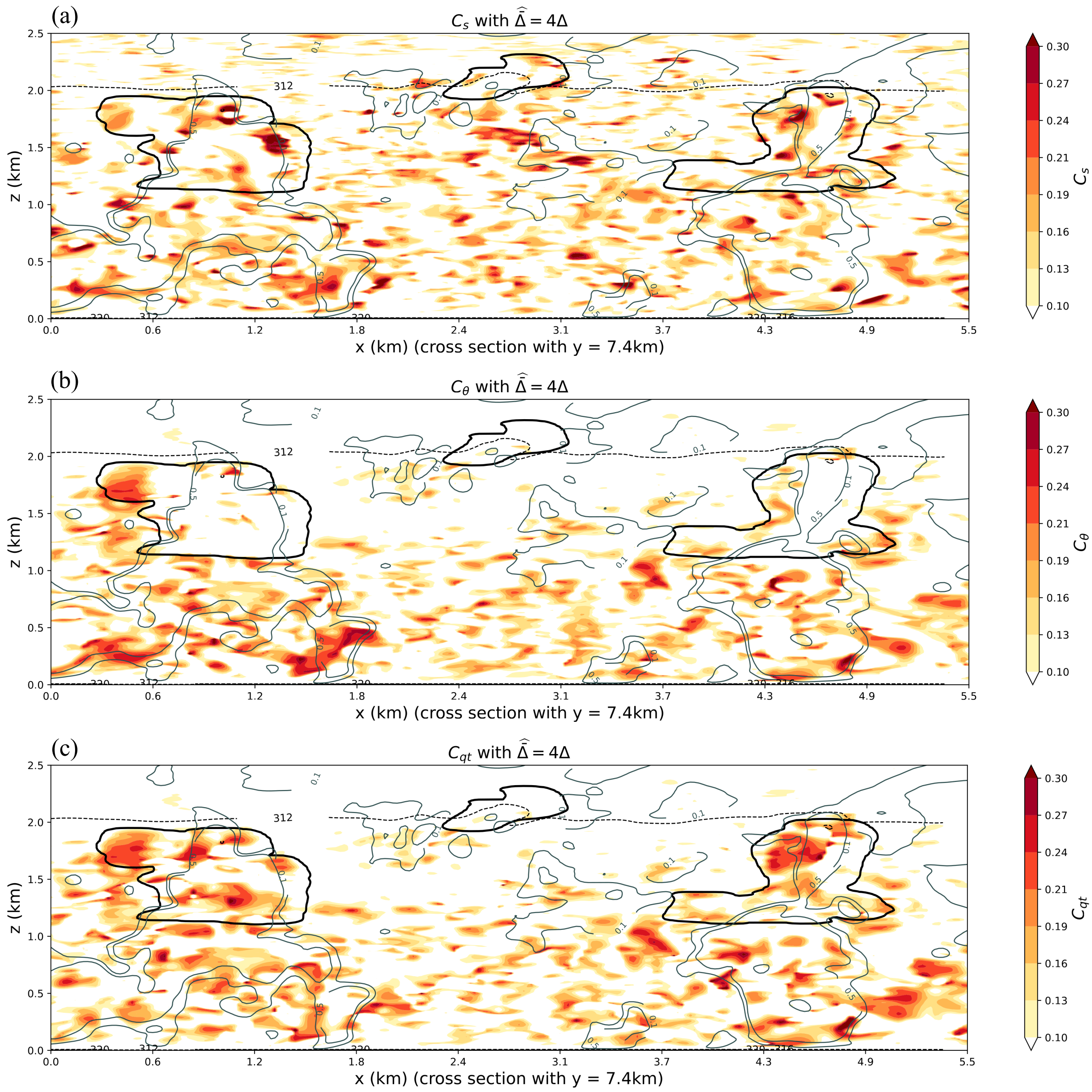

Figure 6.4: *Vertical cross sections through fields of the dynamically computed Smagorinsky parameters for the ARM case, (a) C_s, (b) C_θ, and (c) C_{qt}, calculated using data that has been filtered to $\widehat{\overline{\Delta}} = 4\Delta$. Solid black contours denote the presence of clouds, the dashed black lines show the θ_v levels, and the grey contours outline the areas where $w' \geq 0.1\,m\,s^{-1}$ and $w' \geq 0.5\,m\,s^{-1}$.*

6.2 Partitioning the CTBL Flow into Distinct Regions

The analysis detailed in Section 6.1 clearly demonstrates that the Smagorinsky parameters for each variable exhibit distinct behaviours depending on the flow regime. The flow within the CTBL will now be partitioned into distinct regions, each defined by specific characteristics.

6.2.1 Distinct Layers Within the CTBL

The first and most intuitive division of the CTBL is into the ML and the CL. Clear differences are observed between these two layers, with the mean profiles of various diagnostics exhibiting distinct behaviours depending on the layer. In particular, the profiles for the heat flux ($\overline{\mathrm{w}'\theta}$) and the cloud cover display characteristic features that make them especially useful for identifying the boundaries of each layer. Figures 6.5 and 6.6 show the main profiles of diagnostics focused on throughout this work. Each profile has been scaled by the corresponding mixed layer capping inversion height, z_{ML}. The exact height of this inversion was calculated by finding the height at which the planar-averaged heat flux was at a minimum. In all cases (BOMEX and the four ARM timestamps) except one, the cloud base aligns with the top of the ML. The exception is the early morning ARM profile (t = 10:30L), where the cloud base is located a few metres above the ML top. This minor discrepancy is disregarded, and the boundary between the ML and the CL is defined at the height of the ML capping inversion, z_{ML}.

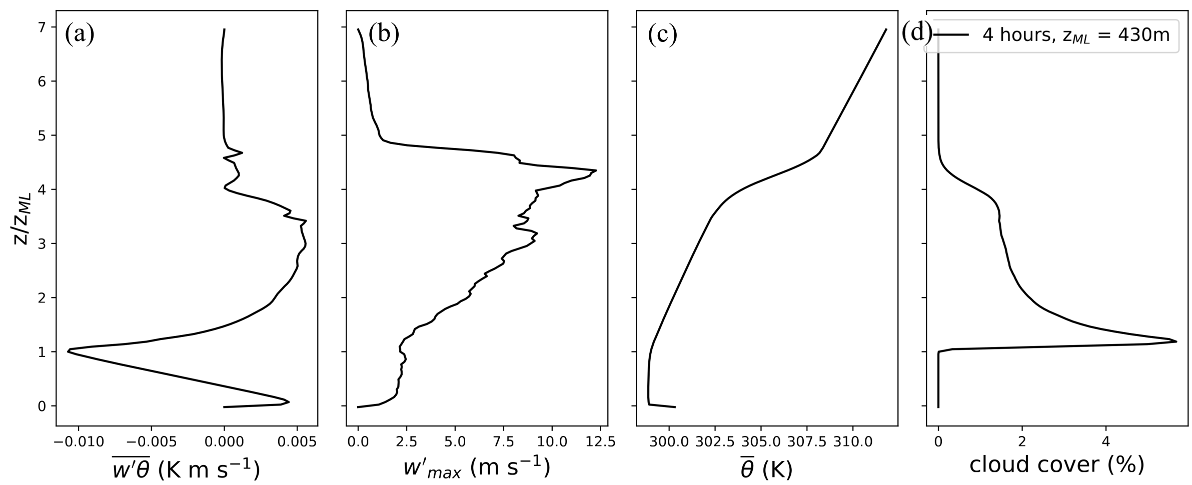

Figure 6.5: *Profiles showing (a) the average value of the vertical eddy heat flux $w'\theta'$ with height, (b) the maximum vertical velocity at each height level, (c) the mean potential temperature profile and (d) the percentage of cloud cover with height for the BOMEX case at t = 4 hours.*

The effect of the capping inversion is apparent not only in the mean heat flux (Figures 6.5 a and 6.6 a) and cloud cover profiles (Figures 6.5 d and 6.6 d) but also in the average potential temperature profile (Figures 6.5 c and 6.6 c). The θ profiles show each case demonstrating a well-mixed layer extending from the surface level up to the capping inversion, while the heat input at the surface is evident by the spike in θ values at the lowest levels. This is also apparent in the heat flux values, which experience a distinct maximum value at the surface of the ARM case (Figure 6.6 b). Meanwhile, the BOMEX case shows a much lower value peak in $\overline{w'\theta}$ values at the surface, reaching values similar to those in the CL (Figure 6.5 a). This is most likely due to the lack of sensible heat fluxes at the surface, as BOMEX is primarily driven by latent heat fluxes. This also causes the maximum vertical wind speeds in the ML to be much lower in the BOMEX case, with $\mathrm{w}'_{\mathrm{max}} = 2.5\,\mathrm{m\,s^{-1}}$, compared to $\mathrm{w}'_{\mathrm{max}} = 5\,\mathrm{m\,s^{-1}}$ in the ARM case (see Figures 6.5 b and 6.6 b respectively).

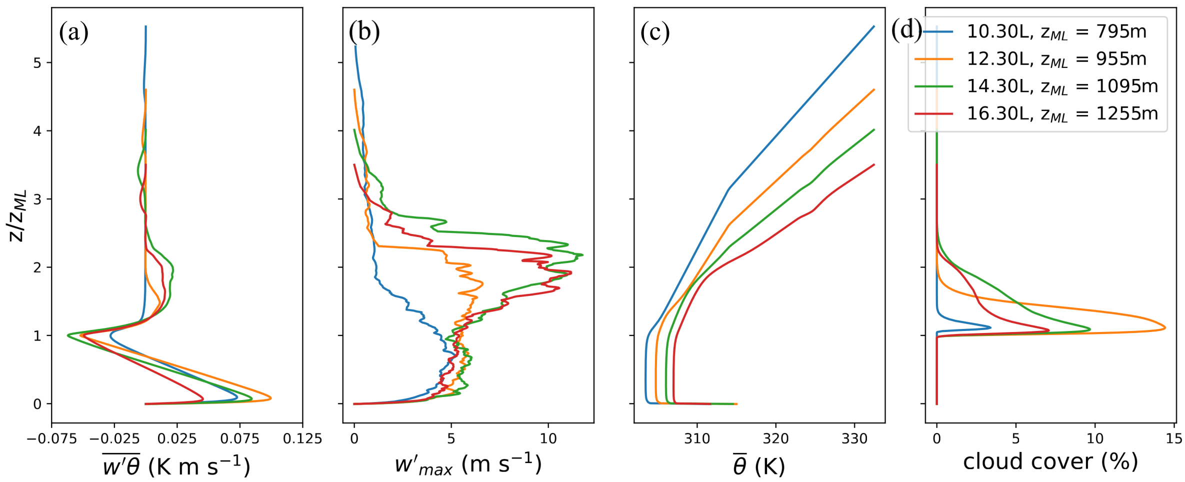

Figure 6.6: *Same as Figure 6.5, but for four different times during the diurnal cycle of the ARM case.*

Above the ML inversion, the θ profiles show all cases exhibiting a conditionally unstable layer, which is overlaid by a statically stable layer. It is in this conditionally unstable layer that the cumulus clouds form. The highest values of maximum upward velocities occur in these clouds, with ARM experiencing maximum speeds of $\mathrm{w}'_{\mathrm{max}} \approx 12\,\mathrm{m\,s^{-1}}$ at the cloud tops during the rapid cloud development stage (see the profile corresponding to 14:30L in Figure 6.6 b). BOMEX reaches comparable maximum w' speeds at the cloud top (Figure 6.5 b), underscoring the role of latent heat release in compensating for the lower velocities of thermals in the BOMEX ML. Another notable difference between the CLs in the ARM and BOMEX cases is the extent of cloud cover. Figure 6.5 d illustrates that clouds cover just over 5.5% of the BOMEX domain, while Figure 6.6,d shows the cloud cover in ARM increases from 3% to 15% over a 6-hour period. It is important to note that, because BOMEX is in a quasi-steady state, its cloud cover remains relatively constant over time.

6.2.2 Different Flow Regimes Within the Cloud Layer

The flow within a CTBL has now been separated into two main layers: the ML and the CL. However, the analysis of the Smagorinsky parameter (C) fields conducted in Section 6.1 suggests that it would be beneficial to split the CL up into areas with similar characteristics. The CL can be divided into the CFE and the IC regions, the latter defined as areas where the cloud liquid water content $q_{\text{cl}} \geq 1 \times 10^{-7}\,\text{kg}\,\text{kg}^{-1}$. While distinguishing between cloudy and non-cloudy points in the CL is important, Figures 6.3 and 6.4 also suggest that position within the cloud itself may have an important influence on the Smagorinsky parameter value. Therefore, to investigate this further the IC region can be further subdivided into the cloud updraft (CU), defined as points where $q_{cl} \geq 10^{-7}\,\text{kg}\,\text{kg}^{-1}$ & $\text{w}' \geq 0.5\,\text{m}\,\text{s}^{-1}$, and the cloud core (CC), defined as points where $q_{cl} \geq 10^{-7}\,\text{kg}\,\text{kg}^{-1}$, $\text{w}' \geq 0.5\,\text{m}\,\text{s}^{-1}$, & $\theta_v' \geq 0\,\text{K}$. These conditions are summarised in Table 6.2.2 below.

	Mixed Layer	Cloud Layer			
	Non -	Cloudy Areas	In-Cloud		
	ML	CFE	IC	CU	CC
Criteria	$0 \leq z \leq z_{\text{ML}}$	$z > z_{\text{ML}}$	$q_{\text{cl}} \geq$	1×10^{-7}	$\text{kg}\,\text{kg}^{-1}$
		$q_{cl} < 10^{-7}\,\text{kg}\,\text{kg}^{-1}$		$\text{w}' \geq 0.5$	$\text{m}\,\text{s}^{-1}$
					$\theta_v' \geq 0\,\text{K}$

Table 6.1: *Summary of the flow regimes into which the CTBL has been categorized, their corresponding positions within the CTBL, and the criteria required for classification as each specific regime.*

To illustrate the areas where these conditions are fulfilled in the CTBL, Figure 6.7 shows fields of (a) TKE, (b) w$'$, and (c) w$'\theta'$ taken from the rapid cloud development stage (14:30L) of the ARM diurnal cycle. Areas with the highest values of C_s, as set by the 98$^{\text{th}}$ percentile, are shown by the grey contours in these plots. These contours often align with large values of TKE, w$'$, and/or w$'\theta'$, demonstrating the effect that these fields have on the mixing length. The w$'$ field shows the upward (red) and downward (blue) motion of the fluid in the CTBL, highlighting the position of thermals in the ML and the CU within active cumulus clouds. The w$'\theta_v'$ flux field can be taken as a proxy for buoyancy. Further to this, in the centre of clouds, where it is certain that w$'$ is positive, the w$'\theta_v'$ field can also be used to highlight the location of the CC.

When comparing the fields in Figure 6.7 with the square of the Smagorinsky parameter fields in 6.2, it is evident that the regions with the most extreme C^2 values are concentrated within the thermals and cloudy areas. The concentration of the significant C^2 values IC, in upward motions, and in positively buoyant regions of the fluid is particularly pronounced in the fields of Smagorinsky parameters for scalars. Figure 6.2 b shows that the C_θ^2 values are at their most extreme in the ML thermals (see Figure 6.7 b), with C_θ^2 values predominantly positive in this region. Meanwhile, in the CL, Figure 6.7 c portrays the most extreme C_θ^2 values occurring IC, where they are mostly negative. While the $C_{q_t}^2$ parameter (6.2 c) exhibits similar behaviour to C_θ^2 in the ML and CFE, it demonstrates a completely different behaviour IC. Extreme $C_{q_t}^2$ values do not show a strong preference for occurring only in the CU or CC, but instead show large, positive values throughout the majority of the IC region.

The general behaviours of the individual Smagorinsky parameters for momentum, heat, and moisture appear self-consistent within each flow regime. It is therefore reasonable to partition the CTBL flow into distinct regions and compute horizontal averages within these areas to facilitate further analysis of these extensive data sets.

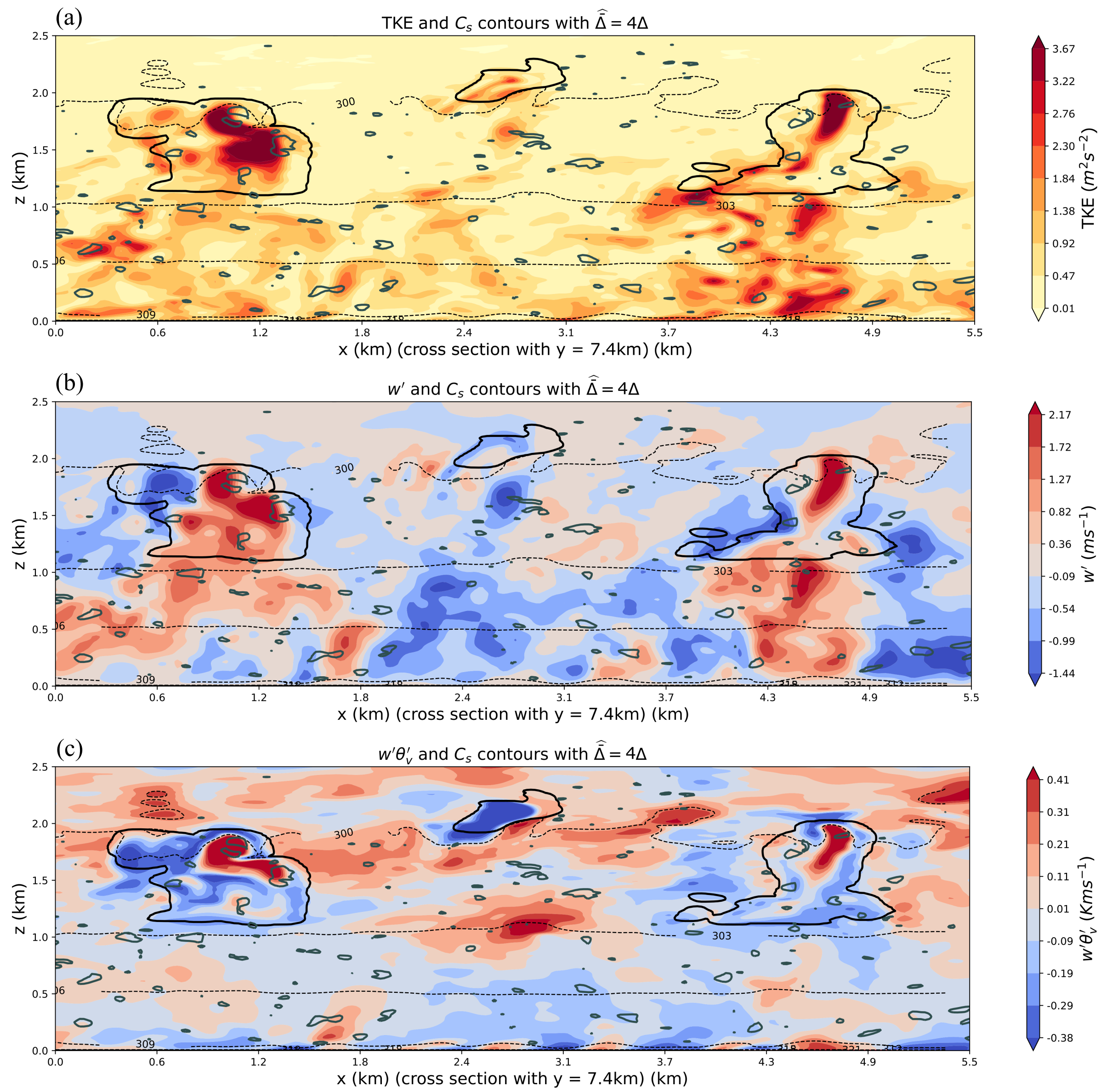

Figure 6.7: *C_s contours outline areas where C_s values are greater than or equal to the 98th percentile of values.*

6.3 Height Dependencies of the Smagorinsky Parameters in the CTBL

Horizontal averaging was imposed across the entire domain (DA), as well as the following distinct flow regimes: the areas with no cloud (NC: encompassing the ML and the CFE within the CL), the IC areas, the CU, and the CC. The definitions of each area are the same as those described in Table 6.2.2. The Smagorinsky parameter fields have been taken and conditionally averaged at each height level within these partitions, following the method detailed in Equations 4.13 and 4.14. The resulting profiles will be analysed and discussed throughout the following section.

6.3.1 Impact of Conservation Properties of Thermodynamic Variables on C_θ

The focus in Section 6.1 was on analysing the fields of the Smagorinsky parameter for heat based only on potential temperature (θ) data. However, Smagorinsky parameters for other thermodynamic variables have also been examined to evaluate the impact that conservation properties have on the parameter. The objective of this comparison study is to determine the most appropriate thermodynamic variable for the new Smagorinsky parametrization. Smagorinsky parameters were calculated for the following thermodynamic variables:

- Potential temperature θ
- Virtual potential temperature θ_v
- Equivalent potential temperature θ_e
- Liquid-water potential temperature θ_L

Fields of C_θ^2, $C_{\theta_L}^2$, $C_{\theta_v}^2$, and $C_{\theta_e}^2$ were computed using high-resolution $\widehat{\overline{\Delta}} = 4\Delta$ data from the ARM and BOMEX cases. These fields were horizontally averaged across each flow regime, with this averaging operation being denoted by angle brackets. The resulting plots are shown for the BOMEX case in Figure 6.8 and the ARM case in Figure 6.9. The C^2 profiles conditioned to lie within the IC, CU, and CC regions exhibited a high degree of similarity in shape. Consequently, only the IC profiles are presented. However, it is important to note that the mean C^2 values were marginally higher in the CU region and increased further in the CC region. The C_{θ_e} parameter was not calculated for the ARM case as the required data was unavailable. In the diurnal CTBL, C_{θ_e} profiles are expected to behave similarly to C_{θ_L} due to their shared conservation properties, along with the observed similarity between them in the BOMEX case.

The domain average (DA) of the Smagorinsky parameter was computed for each thermodynamic variable. The resulting plot for the DA in the BOMEX data is shown in Figure 6.8 a, while the DA for the ARM case at each time step is shown in the top panel (plots a-d) of Figure 6.9. The ARM case shows very similar values for the horizontal averages $\langle C^2_\theta \rangle$, $\langle C^2_{\theta_L} \rangle$, $\langle C^2_{\theta_v} \rangle$ in the ML. In contrast to this, the BOMEX case shows $\langle C^2_\theta \rangle$ and $\langle C^2_{\theta_L} \rangle$ behaving similarly in the ML, with peak values evident at the surface. This peak is thought to be a result of the mismatch between the inclusion of stability functions in the raw data, in contrast to the lack of stability functions in the equations of the offline dynamic model (see Section 4.5 for further details). Artefacts from stability functions are thought to have been present in the raw MONC output fields, which were then unaccounted for when used as input data to the offline dynamic model. These stability functions would have a pronounced effect on the θ and θ_L scalars. Meanwhile, the $\langle C^2_{\theta_v} \rangle$ profile shows a higher value peak at the surface than $\langle C^2_\theta \rangle$ and $\langle C^2_{\theta_L} \rangle$, followed by a steep decrease in value with increasing distance from the surface. The final thermodynamic parameter which was analysed in the BOMEX case is $\langle C^2_{\theta_e} \rangle$, whose profile exhibits a well-defined parabolic shape within the mixed layer. The varying range of behaviours of the thermodynamic parameters in the BOMEX case is presumably due to moisture's significant influence in the BOMEX case compared to ARM. Differences in the treatment of moisture when accounting for its contributions to the thermodynamics in the ML are evidently significant enough to cause the observed differences in profiles when moisture is present in substantial quantities.

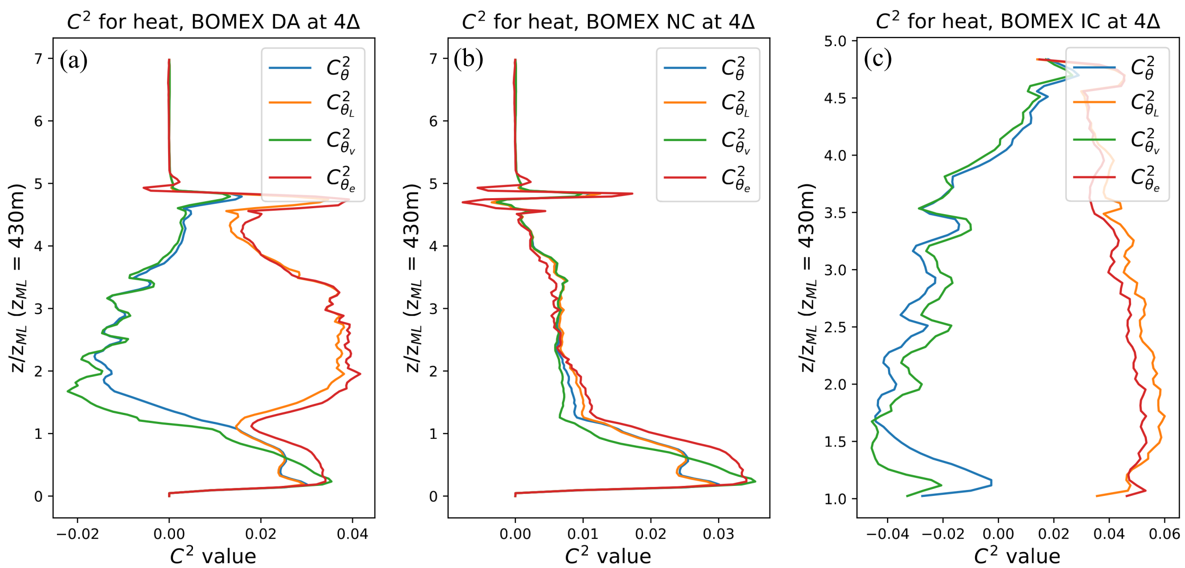

Figure 6.8: *Horizontally averaged C^2 values for each of the heat variables from the BOMEX case, calculated by averaging across (a) the entire domain, (b) areas with no cloud, (c) areas in-cloud.*

In the cloud layer, the differences between the conservation properties of the various thermodynamic variables are even more apparent. In the BOMEX case, the cloud layer ranges between $1 \leq z/z_{\mathrm{ML}} \leq 5$. In the ARM case the CL begins at a height of $z/z_{\mathrm{ML}} = 1$ with the cloud top reaching (a) $z/z_{\mathrm{ML}} = 1.5$, (b) $z/z_{\mathrm{ML}} = 2.3$, (c) $z/z_{\mathrm{ML}} = 2.6$ and (d) $z/z_{\mathrm{ML}} = 2.4$ at each stage of the diurnal cycle respectively. Note that the height of the ML inversion grows with time, and its exact height is recorded on the y-axis of the plots in Figure 6.9. The DA profiles for $\langle C_\theta^2 \rangle$ and $\langle C_{\theta_v}^2 \rangle$ can be seen veering to negative values in the CL for both the BOMEX and ARM cases (Figures 6.8 a and 6.9 b - d). This is consistent with the extensive areas of negative C_θ^2 values seen in the CC regions of the cross-sections shown in Figures 6.1 & 6.2. Due to the non-conservative nature of both the θ and θ_v scalars, it is reasonable for their Smagorinsky parameter profiles to behave similarly in the CL, where moisture begins to actively condense. Conversely, the DA Smagorinsky profiles of the conserved thermodynamic variables do not tend toward negative values within the CL. This is evident in the BOMEX case, with Figure 6.8 a showing both $\langle C_{\theta_e}^2 \rangle$ and $\langle C_{\theta_L}^2 \rangle$ averaging to slightly larger values in the CL ($\langle C_\theta^2 \rangle \approx \langle C_{\theta_L}^2 \rangle \approx 0.04$) than in the ML ($\langle C_\theta^2 \rangle \approx 0.035$ and $\langle C_{\theta_L}^2 \rangle \approx 0.025$). In the ARM case however, the maximum $\langle C_{\theta_L}^2 \rangle = 0.03$ value in the CL remains less than or equal to the peak ML value of $\langle C_{\theta_L}^2 \rangle = 0.035$.

Since condensation, and therefore issues with conservation, only occur within the clouds, it is useful to partition the CL into IC and CFE regions. The profiles resulting from this decomposition are shown in Figures 6.8 and 6.9. The horizontal averages in the CFE for both the BOMEX case (Figure 6.8 b) and the ARM case (Figures 6.9 e - h) all show the $\langle C^2 \rangle$ profiles from each thermodynamic variable collapsing along the same profile. This supports the argument that differences in the DA $\langle C^2 \rangle$ profiles in the CL can be attributed to inconsistencies in the conservation properties between parameters, which becomes apparent when water condenses in the clouds. This is further confirmed in the distinction between the $\langle C_{\theta_e}^2 \rangle$ & $\langle C_{\theta_L}^2 \rangle$ IC profiles, and the IC profiles for $\langle C_\theta^2 \rangle$ & $\langle C_{\theta_v}^2 \rangle$ (with θ_e present in the BOMEX plots only). The BOMEX case (Figure 6.8 c) shows the non-conservative thermodynamic variables, $C_{\theta_e}^2$ and $C_{\theta_L}^2$ going to values as negative as -0.04 in the MBL. Meanwhile the ARM case in Figure 6.9 c depicts these parameter values reaching values as negative as -0.02 throughout the rapid cloud development stage (14:30L) and the quasi-stead state (16:30L). These figures also show that the IC averages of $\langle C_{\theta_L}^2 \rangle$ and $\langle C_{\theta_e}^2 \rangle$ reach values of 0.06 in the BOMEX case, while $\langle C_{\theta_L}^2 \rangle$ attains values of 0.04 in the ARM case.

This suggests that it would be sensible to use a conserved thermodynamic variable in the new parametrization, and as such, the parameter based on liquid water potential temperature θ_L will be analysed further. However, since potential temperature is commonly used in dynamic LES analysis, parameters based on the θ variable will also be examined.

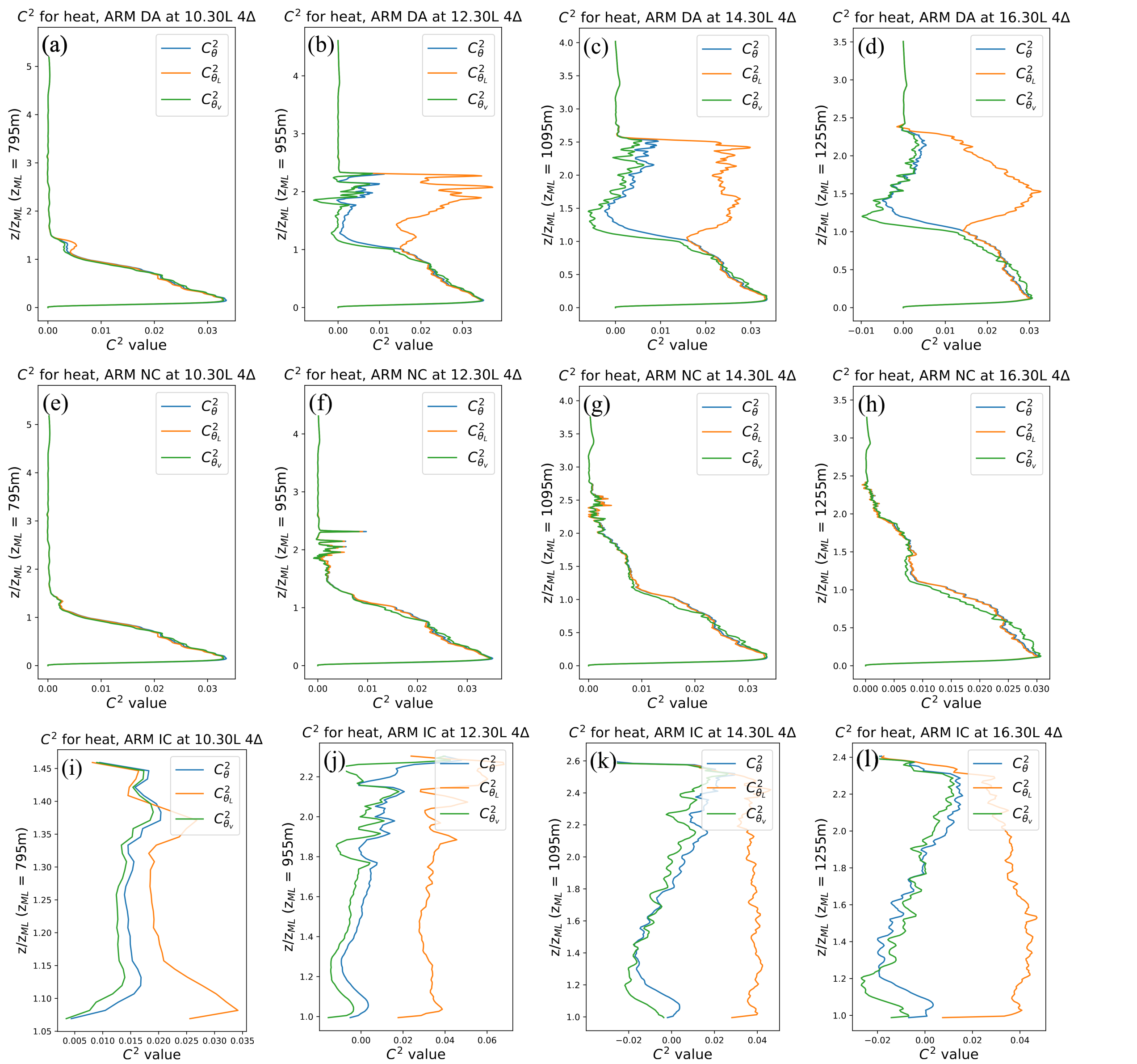

Figure 6.9: *Horizontally averaged values for the various heat parameters squared from the ARM case at four different times throughout the diurnal cycle, calculated by averaging across (a) the entire domain, (b) the areas with no cloud, (c) the areas in-cloud.*

6.3.2 Variable-Dependent Smagorinsky Parameters

The dynamic analysis focuses on the following variables from this point onward: momentum, potential temperature θ, liquid water potential temperature θ_L, and total moisture q_t. Any significant differences between the behaviour of the Smagorinsky parameters for each variable will be identified relative to the flow regime they are in. Again the conditional horizontal average, specific to each distinct region in the CTBL, is used to identify general trends in the behaviour of each C parameter. The parameter profiles for each partition are presented in the plots below. Figure 6.10 illustrates the profiles based on the BOMEX case, while Figure 6.11 shows the profiles derived from the ARM dataset.

The horizontally averaged profiles show that in the CL, differentiating between IC, CU, and CC does not have a significant impact on the majority of the Smagorinsky parameter values analysed across all datasets. This is logical because only the C_θ parameter showed a clear distinction between the IC, CU, and CC regions (recall Figures 6.1 b & 6.2 b). However, as this distinction was based on negative C_θ values in the CC, which have since been clipped and set to zero, the resulting IC $\langle C_\theta \rangle$ profile's ability to differentiate between the IC regions has been diminished. This suggests that position in-cloud (ie: relative to updrafts and the cloud core) does not seem to have a large effect on the planar averaged Smagorinsky parameters for a quasi-steady marine CTBL.

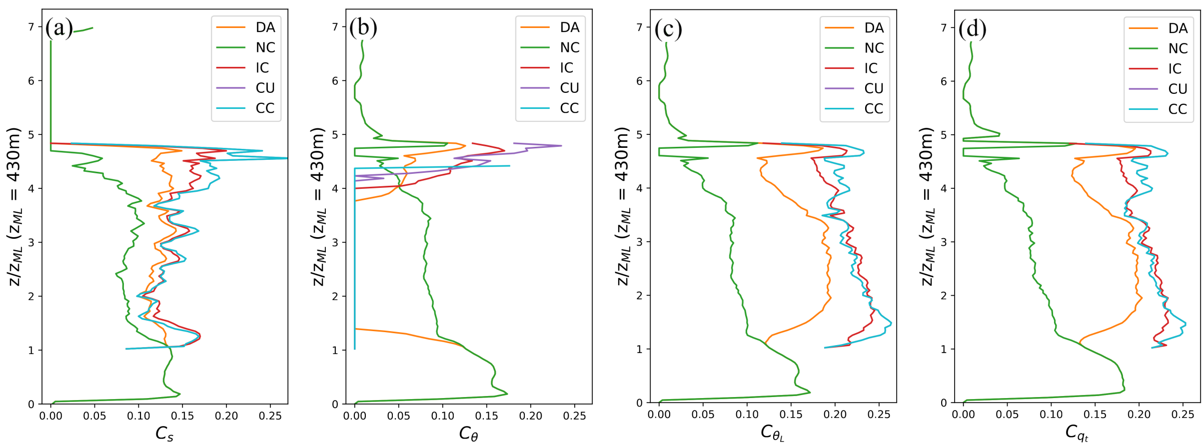

Figure 6.10: *From left to right, profiles of (a) $\langle C_s \rangle$, (b) $\langle C_\theta \rangle$, (c) $\langle C_{\theta_L} \rangle$, (d) $\langle C_{qt} \rangle$ for the BOMEX case, 4 hours after spin-up, with a filter width of $\widehat{\overline{\Delta}} = 4\Delta$. These profiles were calculated by planar averaging points that satisfy the conditions specific to each flow regime, as previously outlined (see Table 6.2.2).*

The BOMEX case as illustrated in Figure 6.10 a, shows the average DA values largely in agreement with the IC, CU, and CC values for $\langle C_s \rangle$. Profiles from the ARM case presented in the top panel of Figure 6.11 show that, while the DA $\langle C_s \rangle$ values do tend toward the IC, CU, and CC values after cloud has developed, there is a clear distinction between the DA $\langle C_s \rangle$ profile and the profiles from IC regions in the early morning as the cloud field initiates. This is due to the lack of cloud in the early morning, resulting in the DA value being strongly influenced by Smagorinsky parameter values in the CFE. Additionally, the DA $\langle C_s \rangle$ profiles in the CL remain approximately 0.02 lower than the IC/CU/CC profiles for all times after 12:30L in the ARM case, while BOMEX only shows a 0.01 difference between the DA and IC/CU/CC.

The behaviour of the Smagorinsky parameters for scalars in the CL is heavily influenced by the conservation properties of each scalar. Recall from Section 6.1 that scalars which are not conserved generate negative C^2 values in the clouds, which must be clipped and set to zero before any averaging procedures can be performed. As expected, data from the BOMEX case presented in Figure 6.10 shows the $\langle C_\theta \rangle$ profiles averaging to zero for all the flow regimes in-cloud. This induces the secondary effect of forcing the DA for $\langle C_\theta \rangle$ in the CL to zero also. Interestingly, the ARM case does not show the same behaviour for the $\langle C_\theta \rangle$ profiles. The most marked difference is apparent in the morning, when cumulus clouds are just beginning to develop. Figure 6.11 e demonstrates that the DA of $\langle C_\theta \rangle \approx 0.5$ in the ARM case at 13:30L is approximately equivalent to $\langle C_\theta \rangle$ in the CFE at the same time. Furthermore, the IC and CU show similar non-zero values of $\langle C_\theta \rangle \approx 0.14$ at 10:30L in the ARM case, while there is a striking increase in the CC average, with values reaching over $\langle C_\theta \rangle\, 0.25$ in this region at this time. This is significant as no other parameter shows a notable difference between the CU and CC average values at any other stage of the diurnal cycle, nor is it seen in the BOMEX case.

As the CTBL develops throughout the day, Figures 6.11 f - h show average $\langle C_\theta \rangle$ values fixed at zero in the lower half of the CL. However, in the upper half of the CL the $\langle C_\theta \rangle$ values can be seen oscillating between near-zero values, up to ML $\langle C_\theta \rangle$ values for the corresponding time. The high $\langle C_\theta \rangle$ values are most likely due to the high levels of mixing occurring at cloud tops where warmer air is being entrained from above. The large $\langle C_\theta \rangle$ parameter values may also be a result of the gravity waves forming due to disturbances in the overlying layer resulting from the overshooting cloud tops. Figure 6.10 b shows the top layer in BOMEX experiencing increases in $\langle C_\theta \rangle$ from 0 to maximum values of 0.25 in some flow regimes near the CL top. However, the maximum values of $\langle C_\theta \rangle$ occur in the CU region and not in the CC areas of the cloud top. with $\langle C_\theta \rangle \approx 0.25$ being the peak value reached. This again aligns with the theory of high amounts of entrainment occurring and/or the effects of gravity waves in the region.

The conserved variables, θ_L and q_t, display similar profiles for the Smagorinsky parameter, both in the ML and throughout the CL. This similarity persists across all four stages of the diurnal cycle in the ARM case, as shown by Figures 6.11 i - p, as well as in the quasi-steady marine BL, as shown by Figures 6.10 c & d. The IC profiles of $\langle C_{\theta_L} \rangle$ and $\langle C_{q_t} \rangle$ are very similar to the CU and CC profiles in the BOMEX case (Figures 6.10 c & d), and are near identical after 14:30L in the ARM case (Figures 6.10 k, l, o, & p). Note that the CC profile almost perfectly overlays the CU profile, making it difficult to see in most plots. In the BOMEX case $\langle C_{\theta_L} \rangle$ and $\langle C_{q_t} \rangle$ values in the cloudy regions are approximately 0.24, while in the ARM case the parameters grow from morning time values of 0.16 to 0.2. The only exception to the alignment of $\langle C_{\theta_L} \rangle$ and $\langle C_{q_t} \rangle$ profiles for the in-cloud regimes occurs during the early stages of the diurnal cycle, at 10:30L and 12:30L in the ARM case. This is illustrated in Figures 6.11 i & m where the IC has lower values of $\langle C_{\theta_L} \rangle \approx \langle C_{q_t} \rangle \approx 0.14$ at 10:30L compared to the CU and CC values of 0.16. The early afternoon also depicts the profiles for the θ_L and q_t Smagorinsky parameters aligning, with the 12:30L profiles in 6.11 j & n showing IC values of $\langle C_{\theta_L} \rangle \approx \langle C_{q_t} \rangle \approx 0.16$ when the CU and CC values reach 0.2. This suggests that accounting for different scalar mixing lengths between IC and CU/CC may be important during the cloud initiation and onset stages of the shallow cumulus development. However, this factor does not seem as significant after the initial development phase.

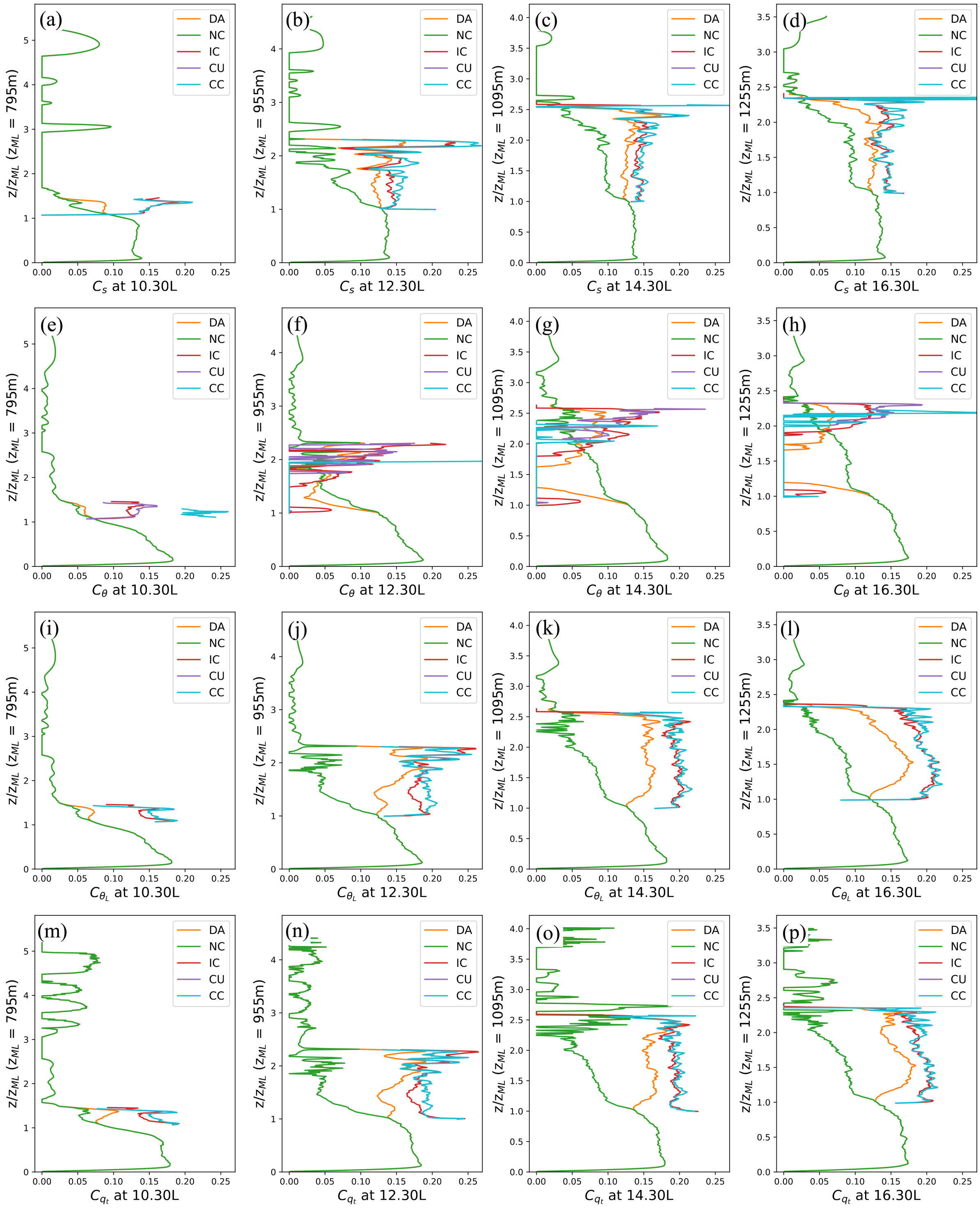

Figure 6.11: *Profiles of planar averaged Smagorinsky parameters in each distinct flow regime in the ARM case, using data with a filter width of* $\widehat{\overline{\Delta}} = 4\Delta$*. From left to right the 4 different timestamps can be seen, while the top panel shows* $\langle C_s \rangle$*, 2nd panel from the top is* $\langle C_\theta \rangle$*, 3rd panel from the top is* $\langle C_{\theta_L} \rangle$*, and the bottom panel details the* $\langle C_{q_t} \rangle$ *profiles.*

The CFE of the CL show similar behaviours and values across all parameters, suggesting that this flow regime may not be variable-specific. In the BOMEX case, Figure 6.10 depicts the profiles for the CFE regimes averaging to values of $\langle C \rangle \approx 0.1$. This is in contrast to the average values in the ML area of the non-cloudy (NC) profiles, which show different values depending on the variable in question. For instance the $\langle C_s \rangle \approx 0.14$ in the ML, while the $\langle C_{\theta_L} \rangle \approx \langle C_{\theta_L} \rangle \approx 0.16$, and the $\langle C_{q_t} \rangle \approx 0.185$ in the ML. Therefore, although scalar dependence appears unimportant in the CFE of the CL, it appears to play a significant role in the ML of the marine CTBL. Examining the CFE averages in the ARM case, the trend of having $\langle C \rangle \approx 0.1$ for all parameters also holds true once the cloud field has developed, specifically for times after 14:30L. However, for earlier times, while $\langle C \rangle$ remains unaffected by variable choice, it takes smaller values in the CFE of the CL. This suggests that $\langle C \rangle$ in the CFE is affected by the eddy size. Figures 6.11 a, e, i, m illustrate this, with the morning time (10:30L) profiles having values of 0.05 in the CFE of the CL. Later on in the day, at 12:30L the top half of the CFE profile ($1.5 \leq z/z_{\mathrm{ML}} \leq 2.4$ in Figures 6.11 b, f, j, n) the parameter profiles maintain the $\langle C \rangle \approx 0.05$, however in the lower levels of the CL ($1 \leq z/z_{\mathrm{ML}} \leq 1.5$ in the same figures) the profile increases in value as it approaches the ML capping inversion. This indicates that, as shallow cumulus clouds develop, they promote mixing in the surrounding non-cloudy environment. This behaviour is consistent with the high levels of mixing observed in the CFE near the cloud edge of actively growing cumulus clouds, as seen in Figure 6.3.

6.3.3 Mean Profiles of the Prandtl and Schmidt Numbers

The profiles for the Prandtl and Schmidt numbers were calculated for each variable using Equations 4.15 and 4.16, recalling that the horizontal averages in each flow regime are calculated using the corresponding $\langle C_s \rangle$ and $\langle C_\psi \rangle$ values for each scalar ψ. The profiles from the BOMEX data are shown in Figure 6.12, while the data from the ARM case are given in Figure 6.13. The resulting profiles from these two case studies are compared and contrasted in the following section.

Firstly, the ML values of the Pr and Sc profiles from the ARM case are analysed. Figure 6.13 shows the average values for all Pr and Sc numbers collapsing along the same profile in the ML. The average ML value in the ARM case can be seen increasing gradually from 0.5 at 10:30L in Figure 6.13 a to 0.6 by 16:30 in Figure 6.13 d. However, not one of the parameters sees their profile reach a value of 0.7 in the mixed layer of the ARM case. This is significant as the standard Smagorinsky scheme in the MONC model assumes that Pr $= 0.7$ (shown by the dashed vertical line in each of the plots), and this number sets the ratio of scalar to momentum mixing for each scalar being modelled. On the other hand, the BOMEX case does show the Prandtl number for θ and θ_L following the 0.7 vertical line in the middle of the ML (Figure 6.12 a).

The profiles of Pr_{θ_v}, Pr_{θ_e}, and Sc_{q_t} are more similar to those seen in the ARM case, with these parameters taking average values of 0.5 in the ML also.

Above the ML, the DA of the non-conservative scalars, θ and θ_v go to unrealistically large values due to the issues posed by condensation of moisture in the CL, and therefore these parameters are overlooked. At the capping inversion ($z/z_{\mathrm{ML}} = 1$), the divide between the ML and the CL, Figure 6.12 a shows the DA of both Pr_{θ_e} and Sc_{q_t} peaking at unity in the BOMEX case, indicating that the average mixing of θ_e, q_t, and momentum are identical at this boundary between flow regimes. Interestingly, this case shows the Prandtl number for θ_L peaking at $\mathrm{Pr}_{\theta_L} \approx 1.2$, indicating that θ_L has smaller structures and therefore undergoes less mixing than momentum at the ML capping inversion. This is contrary to the common assumption that the structures mixing thermodynamic scalars in the CTBL are larger, therefore resulting in their corresponding Prandtl numbers being less than one (conventionally $0.5 \le \mathrm{Pr} \le 1$). The ARM data in Figure 6.13a - d also demonstrates that the DA of Pr_{θ_L} consistently takes larger values than the DA of Sc_{q_t} at the capping inversion, with the exact values taken by each parameter varying in time. The morning BL shows these parameters peaking at much larger values of $\mathrm{Pr}_{\theta_L} \approx 1.9$ and $\mathrm{Sc}_{q_t} \approx 1.4$ at the ML capping inversion. However, as the cloud layer develops the DA values at the capping inversion demonstrate a reduction to $\mathrm{Pr}_{\theta_L} \approx 1$ and $\mathrm{Sc}_{q_t} \approx 0.85$ from 12:30L onward. This difference between Pr_{θ_L} and Sc_{q_t} at the ML capping inversion implies that the mixing length may benefit from being scalar-specific at this particular interface in the CTBL flow.

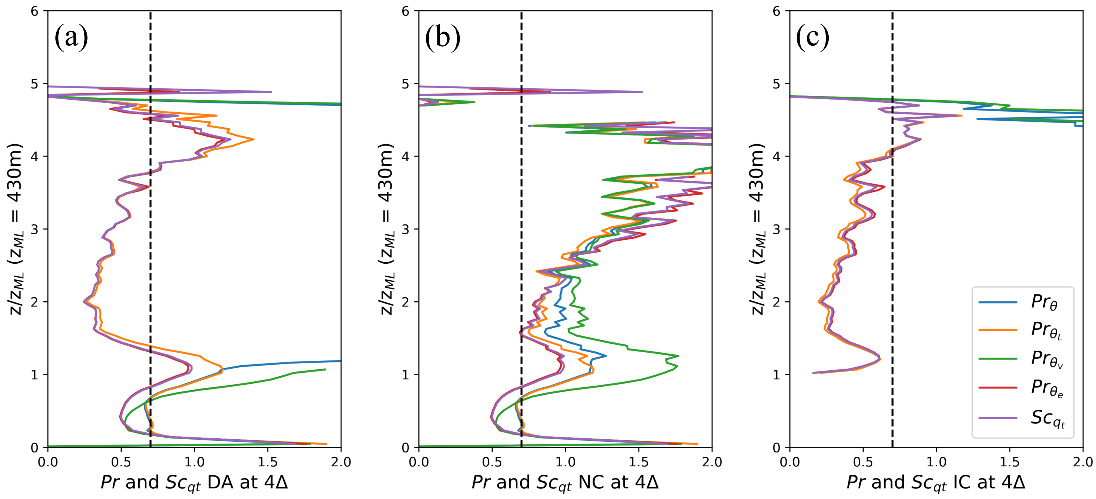

Figure 6.12: *Vertical profiles of the Pr_θ, Pr_{θ_L}, Pr_{θ_v}, Pr_{θ_e}, and Sc_{qt} from the BOMEX case, 4 hours after initiation. The plots show the profiles for the horizontal average across the (a) total domain, (b) areas with no cloud, and (c) regions in-cloud, and scaled by the height of the mixed layer: $z_{ML} = 480\,m$. The black dashed vertical line shows where $Pr = Pr_{air} = 0.7$.*

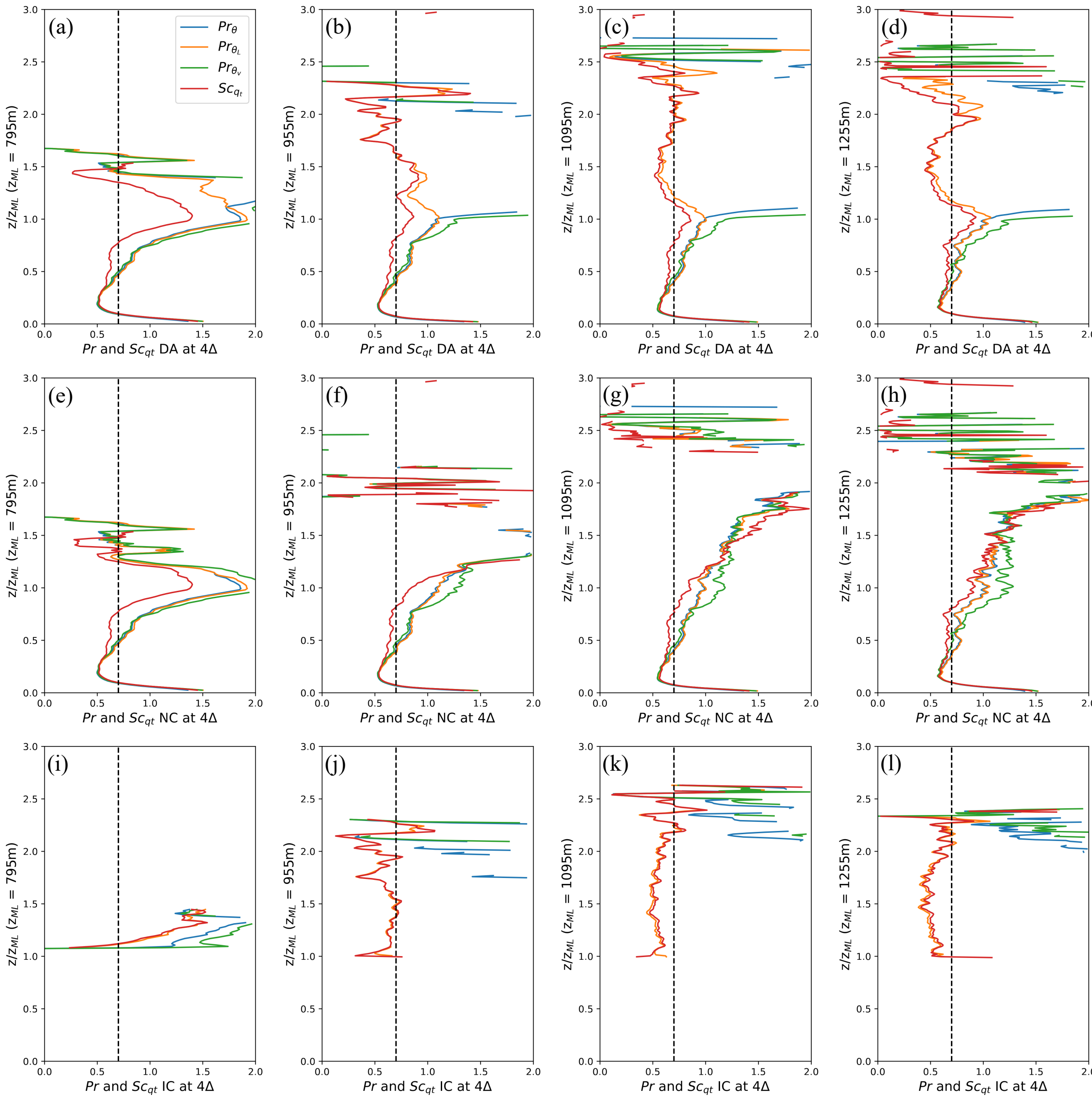

Figure 6.13: *Vertical profiles of the Pr_θ, Pr_{θ_L}, Pr_{θ_v}, and Sc_{q_t} from the ARM case, at four different time steps throughout the diurnal cycle. The plots show the profiles for the horizontal average across the (top) total domain, (middle) areas with no cloud, and (bottom) regions in-cloud, and scaled by the height of the mixed layer*

When looking at the CL above this capping inversion, it is more productive to decompose the layer into its IC and CFE counterparts before conducting any analysis. In the CFE, both the BOMEX data (Figure 6.12 b) and the ARM data (Figures 6.13 e - h) show the averages of all the Prandtl and Schmidt numbers following similar trajectories. However, in the CFE just above the capping inversion, Figures 6.13 e & f outline how, compared to the other scalars, lower average values of $\mathrm{Sc}_{q_t} \approx 1.3$ at 10:30L and $\mathrm{Sc}_{q_t} \approx 0.9$ at 12:30L are experienced in the earlier stages of the diurnal cycle. Meanwhile, the average Prandtl number for all other scalars in the CFE is $\mathrm{Pr} \approx 1.5$ at 10:30L and $\mathrm{Pr} \approx 1.2$ at 12:30L. The BOMEX case also shows a difference between profiles just above the capping inversion, however it is the Prandtl numbers corresponding to the non-conserved variables θ and θ_v which show larger Pr values, while those relating to the conserved variables exhibit similar values of $\text{Pr}_{\theta_e} \approx \text{Pr}_{\theta_l} \approx \text{Sc}_{q_t} \approx 0.8$. As we ascend in the CL, average Pr and Sc values in the CFE increase, as shown by the BOMEX case in Figure 6.12 b, and across all times in the ARM case, as presented in Figures 6.13 e - h. This indicates a strong discrepancy between the size of momentum structures and heat/moisture structures in the CFE which only increases with height.

The IC regions exhibit the opposite behaviour, with Figure 6.12 c illustrating the Prandtl and Schmidt numbers of the conserved variables taking nearly identical profiles. These IC profiles have lowest values of 0.3 in the lower levels of the CL, which gradually increase to 0.5 in the mid-to-upper level, and reach 0.8 at the cloud top. This indicates that within the marine cumulus clouds, the q_t moisture structures and θ_L & θ_e heat structures are, on average, the same size, while the momentum structures in these clouds are much smaller. Figures 6.13 i - l demonstrate how the ratio of momentum structures to scalar structures evolves in-cloud throughout the diurnal cycle. In the morning-time BL, Figure 6.13 i reveals how the average in-cloud Prandtl and Schmidt numbers are only slightly smaller than the CFE averages for the same time. However by 12:30 Figure 6.13 j shows that the conserved scalars have near-identical profiles IC, with average values having reduced to $\text{Pr}_{\theta_l} \approx \text{Sc}_{q_t} \approx 0.7$. Once the active cumulus clouds have fully developed (see Figures 6.13 k & l) the average in-cloud values stabilise at $\text{Pr}_{\theta_l} \approx \text{Sc}_{q_t} \approx 0.5$. This indicates that as a boundary layer transitions from a CTBL with rapidly developing cumuli to a quasi-steady state, the heat and moisture structures within the cumulus clouds grow to twice the size of the IC momentum structures.

6.4 Summary of Variable-Dependent Parameter Responses to the Flow Regime

Throughout this chapter, the effect that various flow regimes have on the Smagorinsky parameter value is assessed. The conditionally averaged profiles of Smagorinsky parameters suggest that differentiating between the cloud, cloud updraft, and cloud core regions does not appear to have a significant influence on the dynamic parameter value. It was found that decomposing the CTBL into a ML, CFE, and CL gave the best results, with all Smagorinsky parameters showing distinct and systematic behaviours within these three regimes.

In the ML regime, the Smagorinsky parameters for each variable behave similarly, with fields of alternating positive and negative parameter squared values. However, the C_s^2 fields exhibit smaller, more sporadic patches of alternating positive and negative value regions than the C_θ^2 and C_{qt}^2 fields. The average profiles demonstrate how C_s, on average, also has lower values than the C_θ and C_{qt} parameters in the ML. This is further emphasized by the low values of the Prandtl and Schmidt numbers in the ML, with the majority of scalars having Pr $\approx$ Sc ≈ 0.5. This value is notably lower than the current value of Pr $= 0.7$ which is prescribed to the flow in the MONC model, suggesting that the current model underestimates the scale of the heat and moisture structures. This is also evidenced in the C_θ^2 and C_{qt}^2 fields, which show similar patterns as C_s^2, but the patches for the scalar parameters are larger and more consistent than those for momentum. This indicates that the same turbulence is present across the fields, but it is impacting coherent structures of different scales, with heat and moisture having larger structures than momentum.

The C_s^2 fields demonstrate that the noise in the momentum field persists into the CL. This is evident in the IC areas of the C_s^2 which still show alternating patches of positive and negative values, though the positive values here are slightly larger IC than in the ML, and this is confirmed by the conditional profiles. In contrast to this, the C_θ^2 and C_{q_t} fields show much more consistency within the IC regions. The C_θ^2 fields show extensive areas of negative values within the CU/CC, indicating counter-gradient transport occurring in these areas, with near-zero values outside this specific region of the cloud, signalling a lack of mixing in these areas. In sharp contrast to this, the C_{qt}^2 fields show predominantly large, positive values of C_{qt}^2 throughout the IC regions. This is confirmed by the large values of the C_{q_t} profile conditioned on the IC region, with the Sc_{q_t} indicating that the moisture structures within the cumulus clouds are over twice the size of the momentum structures in the same region.

The noisy patchwork of positive and negative C_s^2 values continues to persist in the CFE, however, the values appear smaller and slightly more areas show near-zero values. The suspected lower values of C_s in the CFE are confirmed by the conditional profiles, which show the CFE as having the lowest average values compared to all other regimes in the CTBL. In the parameter fields for the scalars, a much clearer distinction is evident between the CFE and the rest of the CTBL, with the C_θ^2 and $C_{q_t}^2$ values near zero throughout almost the entirety of the CFE. This is confirmed by the very large values of the Prandtl and Schmidt numbers presented in the conditional profile plots. These profiles show values of $1 \leq \mathrm{Pr}, \mathrm{Sc} \leq 2$ in the CFE, with larger values higher up in the CL. This implies that the heat and moisture structures are the same size as the momentum structures in the lower and middle of the CL, but diminish in size to scales half that of the momentum structures near the top of the cloud layer.

Focusing on the scalar parameters which were analysed, it is clear that the conservation properties of the scalars have a large influence on the Smagorinsky parameter value. Unconserved variables experience counter-gradient transport in regions where the water vapour condenses, causing the dynamic Smagorinsky equations to generate negative value parameters. This typically occurs in the IC regions as the Smagorinsky model attempts to account for the upscale transport of energy. However, negative mixing lengths are unrealistic and unusable in the model and therefore are clipped to zero, shutting off mixing completely in these areas. The main differences between the conserved variables are now examined. If the scalars are conserved, their associated Smagorinsky parameters and Prandtl/Schmidt numbers exhibit similar behaviour regardless of the flow regime, suggesting that scalar-specific parameters may not be necessary as long as the properties of the scalar align. The Prandtl and Schmidt number profiles suggest that the only substantial difference between conserved variables, that is consistent across both the BOMEX and ARM cases, occurs in the layers above and below the ML capping inversion. In this region, the profile indicates that $\mathrm{Pr}_{\theta_L} \geq \mathrm{Sc}_{q_t}$, implying that the moisture structures are bigger than the heat structures. This is consistent with findings from previous studies (see Section 2.2 for details). The analysis also underscores that setting $\mathrm{Pr} = 0.7$ in the standard model fails to capture the variation in scales of the coherent structures for momentum, heat, and moisture across the different flow regimes.

6.4.1 Responses to the Research Questions

(1) Does the Smagorinsky parameter exhibit different systematic behaviours, depending on the flow regime? Can these behaviours be quantified into general relationships between the regime and the parameter value?

Yes, the Smagorinsky parameters for all variables are observed to take on key characteristics when in a given flow regime in the CTBL. In the ML, each parameter yields profiles that take on a slightly parabolic shape for all parameters. The largest C values occur in the middle of the mixed layer, with a marked decrease in value as the surface and ML capping inversions are approached. Meanwhile, the IC Smagorinsky parameter values for all conserved variables are notably larger than the parameters in the ML and CFE, with the IC profiles being relatively constant with height. In contrast to this, the unconserved variables experience the IC Smagorinsky values being set to zero. This is because the Smagorinsky model is not equipped to handle the counter-gradient transport of non-conserved variables that occurs IC as the moisture condenses. In the CFE there is a marked decrease in mixing for all variables, with the Prandtl and Schmidt number profiles suggesting that the heat and moisture structures are the same scale as the momentum structures, and even become smaller than the momentum structures near the CL top.

(2) Are there significant differences between the transport and diffusion of momentum, heat, and moisture in the CTBL?

The momentum parameter exhibits the most anomalous behaviour, characterised by noisy fields of alternating positive and negative values that lack the degree of organisation and structure observed in the heat and moisture parameter fields. The C_s fields also lack a clear overall distinction between the ML, IC, and CFE as the noise camouflages any differences. Meanwhile, the heat and moisture parameters show clear variations between these three regimes. Furthermore the heat and moisture fields also both show a higher degree of structure than the momentum field. While the structure within the ML and CFE of the C_θ and C_{q_t} is very similar, the IC regions are where these parameters show a high degree of difference. The extreme values of C_θ^2 demonstrate a clear preference for the CU/CC region, whereas the $C_{q_t}^2$ field sees its extreme values occurring throughout the entirety of the cloud. Further to this the IC C_θ^2 values are entirely negative, whereas the $C_{q_t}^2$ values are mostly positive throughout the cloud. The differences between conserved variables, such as θ_L and q_t are less apparent. The main distinctions between these parameters occur in the area above and below the ML capping inversion. Here the moisture structures are observed to be larger than the heat structures. Therefore, to a certain extent, each variable exhibits its own unique characteristics.

(3) Considering that the conservation properties of a scalar affect its corresponding Smagorinsky parameter value, which thermodynamic variable is most appropriate for the new parametrization?

Ideally, a conserved variable would be used as the basis for heat mixing and dissipation in any new parametrization for the Smagorinsky scheme. However, the majority of LES models, MONC included, are built using θ as the basis for the mixing length of heat. Switching to a conserved thermodynamic variable would require significant restructuring of the LES framework, which is impractical. As a result, the θ parameter remains in use. Despite this, it is preferable to base the Smagorinsky subgrid parametrization on a conserved variable, as non-conserved variables like θ encounter counter-gradient transport IC. This presents a challenge for the Smagorinsky scheme, which cannot handle transporting energy upscale. Therefore, the behaviour of the C_{θ_L} parameter is still analysed, with the intention of testing if the parametrization can be improved by imposing the characteristics of a conserved thermodynamic variable on the unconserved θ fields being modelled. That being said, further testing is needed to determine whether θ or θ_L provides a more suitable basis for the new Smagorinsky coefficient parametrization.

7 The Effect of Filter Scale on the Smagorinsky Parameters

This chapter focuses on determining the local filter scale dependencies of the flow. In the previous chapter, the effect of certain flow regimes on the values of the Smagorinsky parameter was investigated. However, this analysis focused on filtered LES with an effective grid spacing of $\widehat{\overline{\Delta}} = 4\Delta$, the highest resolution dynamic data available. Therefore, the next step in this investigation is to assess the impacts of the grey-zone on each flow regime as the filter scale increases. Note that in this chapter, the terms filter scale and effective grid spacing are used interchangeably. This is because, in this analysis, the effective grid spacing $\widehat{\overline{\Delta}}$ of a filtered field is directly determined by the Gaussian filter scale σ which was applied to the field. While the standard Smagorinsky model has a scale-aware mixing length, as demonstrated in Equation 3.30, it is clear that this alone is insufficient for accurately handling dissipation when in the grey-zone regime. Recall that the mixing length benefits from the use of the Blackadar formula (see Equation 8.1) as a limiting function in the near-surface region, decreasing the mixing length in response to the reduction in eddy size as the surface is approached. This can be considered as a grey-zone adaption, prompting the question of whether additional modifications can be incorporated into the Smagorinsky model to address the challenges encountered in the grey-zone regime. This chapter aims to evaluate and quantify the impact that filter scale has on LES of CTBLs as resolution coarsens to grey-zone scales. The objective of this is to identify relationships between the filter scale and the Smagorinsky parameter value, which can then be used in grey-zone parametrization schemes.

This chapter presents the findings from the analysis of LES of the ARM and BOMEX cases, representing an evolving case and an equilibrium case respectively. The primary objective of this chapter is to investigate the scale dependencies of the Smagorinsky parameter, focusing on the distributions and central tendencies of its values. However, before conducting this analysis, it is necessary to address the negative parameter values that arise during the dynamic analysis. The key research questions guiding the work in this section are as follows:

- To what extent does the filter scale systematically influence the prevalence of negative C^2 values across the different flow regimes in the CTBL?
- How does the variability of the Smagorinsky parameters change with scale in each distinct flow regime?
- What are the key scale dependencies of the Smagorinsky parameters in each of the flow regimes?

7.1 The Prevalence of Negative C^2 Values at Different Scales

The dynamic method is known to produce negative Smagorinsky parameter values, primarily due to the inability of the Smagorinsky scheme to capture up-gradient transport. Understanding how coarsening the filter scale influences the prevalence of these negative values is important, as it has the potential to cause issues when the model operates within the grey-zone regime. As we move to coarser resolutions the model reaches a point where the grid spacing is of the same order as the energy-producing scales of the flow. This means that the model is no longer resolving and dissipating energy at scales in the ISR, but rather at the large scales. Therefore the problem presented by upscale transport is expected to be much more common at coarser resolutions. To test this, the frequencies of negative-valued Smagorinsky parameters at various effective grid spacings have been analysed. The flow was decomposed into the IC areas and the NC regions (which is made up of the ML and the CFE), and the proportion of negative C^2 values in each distinct region were counted at each height level. Figures 7.1 and 7.2 present the findings from the BOMEX and ARM cases, respectively. Therefore these profiles show the relative frequency of negative points as a percentage of the entire horizontal domain, categorised by whether they were identified as being IC or in the CFE.

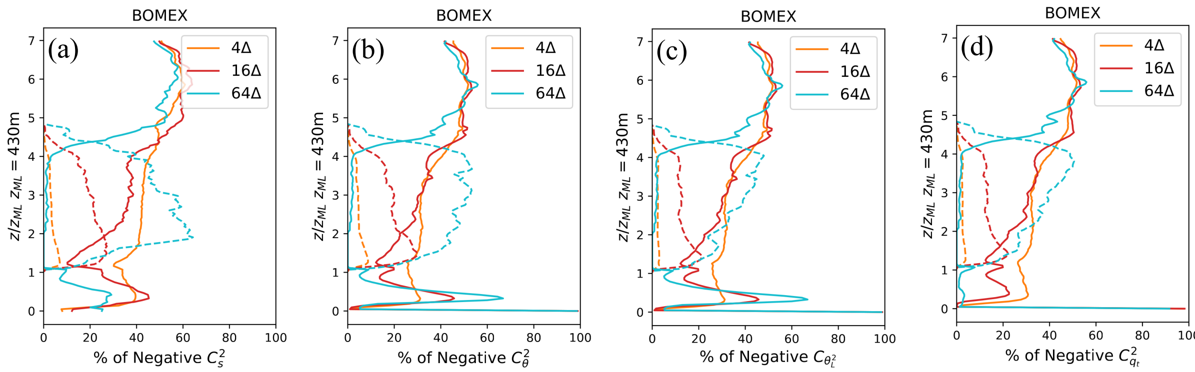

Figure 7.1: *Profiles showing the percentage of the domain which takes negative C^2 values in the BOMEX case. The solid lines indicate data from the NC environment, while the dashed lines indicate the IC data. The profiles from both these flow regimes have been scaled by the total number of points at each level.*

The sole similarity observed across all cases, parameters, and resolutions is that, in the free troposphere (FT) above the cloud top (CT), the NC profiles indicate that approximately 55% of the C^2 values are negative. However, this area is not the focus of this study; rather, the other regions of the CTBL are of primary interest for the new Smagorinsky parametrization. These areas, the ML, CFE, and IC areas of the CTBL show various dependencies on regime, variable, and scale. The percentage of negative C_s^2 values is consistent with height throughout the ML, with approximately 40% of these parameters identified as being negative when $\widehat{\widetilde{\Delta}} = 4\Delta$ in both the BOMEX case (Figure 7.1 a) and during all times during the ARM case (Figures 7.2 a - d). The BOMEX case then shows that the percentage of negative C_s^2 points in the ML decreases as the effective grid spacing ($\widehat{\widetilde{\Delta}}$) increases. However, the opposite occurs in ARM, where the percentage of negative C_s^2 values increases as $\widehat{\widetilde{\Delta}}$ increases. The difference in the number of negative values between scales reduces as time progresses, with the morning time ML (Figure 7.2 a) showing $\widehat{\widetilde{\Delta}} = 64\Delta$ having approximately 10% more negative C_s^2 values than $\widehat{\widetilde{\Delta}} = 4\Delta$, but by late afternoon (Figure 7.2 d) the ML shows only 5% difference between the scales.

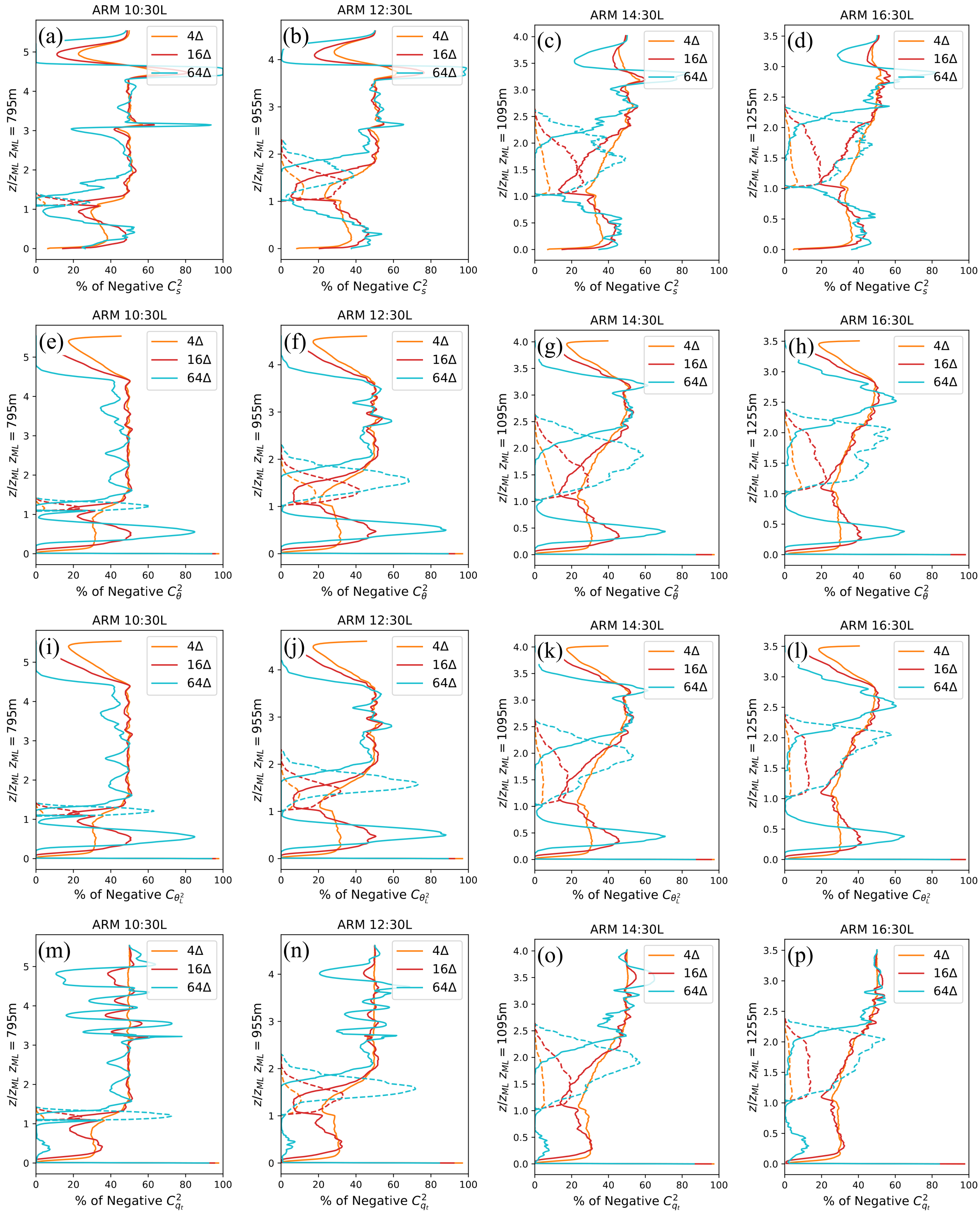

Figure 7.2: *Same as in Figure 7.1, but instead showing the ARM data at 4 different times. The rows show the profiles for a certain parameter, while the columns show each of the time stamps.*

The percentage of negative scalar parameters in the ML shows a clear scalar dependence as the grey-zone is approached. Both the BOMEX and ARM cases demonstrate that at $\widehat{\overline{\Delta}} = 4\Delta$, for every scalar ψ, approximately 30% of the scalar flux coefficients C_{ψ}^2 take negative values throughout the ML. However, the scale-dependent behaviour of each of the C_{ψ}^2 parameters varies significantly depending on the specific scalar quantity, ψ, being considered. As condensation does not readily occur in the ML, the C_{θ}^2 and $C_{\theta_L}^2$ parameters behave the same in this flow regime as conservation does not become an issue. In the BOMEX case, the percentage of negative C_{θ}^2 and $C_{\theta_L}^2$ values in the ML increases from 30%, to 45%, to 70% as $\widehat{\overline{\Delta}}$ coarsens from 4Δ, to 16Δ, to 64Δ respectively (Figures 7.1 b & c). Similar percentages of negative parameters are observed in the ML during the later stages of the diurnal cycle for C_{θ}^2 (Figures 7.2g & h) and $C_{\theta_L}^2$ (Figures 7.2k&l). However, earlier in the day, Figures 7.2 e, f, i & j show the percentage of negative C_{θ}^2 and $C_{\theta_L}^2$ increase to just below 90% in the ML.

The scaling trends set by the C_{θ}^2 and $C_{\theta_L}^2$ parameters in the ML are in complete contrast to the trends seen in the C_{qt}^2 parameter values in the same region. In the ML of the BOMEX case, Figure 7.1 d illustrates the percentage of negative C_{qt}^2 in the ML decreasing from 30% to 20%, and eventually to less than 5% as $\widehat{\overline{\Delta}}$ increases from 4Δ, to 16Δ, to 64Δ. The ARM case however shows a similar percentage of negative C_{qt}^2 values in the ML when $\widehat{\overline{\Delta}} = 4\Delta$ as $\widehat{\overline{\Delta}} = 16\Delta$ (Figure 7.2 m - p). The similarity between these two scales becomes more pronounced as the CTBL evolves throughout the day, with the 16:30L data showing near identical percentages of negative values occurring in the ML, as evident in Figure 7.2 p. However, this similarity does not hold as the effective grid spacing is pushed further into the grey-zone, with the $\widehat{\overline{\Delta}} = 64\Delta$ ARM data showing negative C_{qt}^2 values occurring in only 5% to 10% of the ML.

There is a distinct decrease in the number of negative C^2 values for all variables evident in the levels above and below the ML capping inversion z_{ML} for every $\widehat{\overline{\Delta}}$. This indicates that at the inversion the dynamic method does not detect many instances of counter-gradient transport in the levels about the inversion. Above this capping inversion, the flow in the CL has been partitioned into two distinct regions: the IC areas and the CFE. Note that the IC profiles depict the number of negative values in-cloud as a percentage of the entire horizontal domain at that level, and the same for the CFE profiles. Therefore the impact of changes to the percentage of total cloud cover with scale is not accounted for in these profiles. However, the growth in cloud cover with increasing $\widehat{\overline{\Delta}}$ is driven artificially by the filtering procedure (see Section 7.1.1 for more details), by factors which are not expected to occur in LES models. Therefore the following plots display the frequency of negative values as a percentage of the entire horizontal domain, rather than as a percentage of the area of each specific flow regime. Additionally, representing the negative values as a percentage of the total domain makes the distinction between scales more evident.

In Figures 7.1 a - d the BOMEX data, with an effective grid spacing of $\widehat{\overline{\Delta}} = 4\Delta$, indicates that the CFE exhibits approximately 5% more negative values than the ML for each of the Smagorinsky parameters. As the effective grid spacing increases in size, both the BOMEX and ARM cases show the percentage of negative C^2 values identified in the CFE decreasing. This is evident in the slight decrease between 4Δ and 16Δ, however, by the time that 64Δ is reached, the percentage of negative C^2 parameters occurring in the CFE has dropped significantly to near zero. This decrease in the quantity of negative C^2 values as $\widehat{\overline{\Delta}}$ increases is seen in the ARM case also. However, this trend likely arises from the increase in the number of points classified as being in-cloud as resolution decreases. This results in fewer points being classed as being in the CFE, artificially causing a reduction in points with a negative C^2 value from occurring in the CFE. The impact of $\widehat{\overline{\Delta}}$ on cloud cover, and the knock-on effects of this on the percentage of negative C values is further discussed in Section 7.1.1.

The percentage of negative C^2 values detected in the CFE gradually increases with height. This trend is less pronounced in the BOMEX case; however, it is clearly evident during the rapid development stages of the cumulus clouds in the ARM case, as illustrated in Figures 7.2 c, g, k & o. The difference between the percentage of negative C^2 values near the CB and at the CT of the CFE becomes more evident as the effect grid spacing increases in size.

As anticipated, the percentage of negative C^2 values in the CL increases in the cloudy regions as $\widehat{\overline{\Delta}}$ increases. This phenomenon is again believed to be a direct result of the growing number of points classified as being in-cloud with the increasing filter scale. At high resolution, the percentage of negative C^2 values identified in-cloud is very low, with Figures 7.1 a, c & d indicating that less than 5% of the negative C_s^2, $C_{\theta_L}^2$, and C_{qt}^2 values occur IC. Meanwhile, approximately 10% of the negative ${C^2}_\theta$ values occur in-cloud in the BOMEX case, as presented in Figure 7.1 b. A similar percentage of negative C^2 values are detected in-cloud in the BOMEX case as in the ARM case at each resolution once the diurnal CTBL reaches a quasi-steady state (QSS), as evidenced in Figures 7.2 d, h, l & p. However, during the developmental stages of the cumulus clouds (12:30 L and 14:30), a higher percentage of negative C^2 points are classified as being IC, rather than in the CFE, as indicated in Figures 7.2 a, b, e, f, i, j, m & n. During the onset of clouds at 12:30L, the $\widehat{\overline{\Delta}} = 4\Delta$ data reveals that approximately 10% of the negative C_s^2, $C_{\theta_L}^2$ and C_{qt}^2 points are IC, as indicated in Figures 7.2 c, k & o). Meanwhile, at $\widehat{\overline{\Delta}} = 4\Delta$, the C_θ^2 parameter indicates that 20% of the negative CL values are classified being IC during the onset of cumulus convection in the ARM case (see Figures 7.2f).

At high resolutions, when the system is in QSS (as in Figures 7.1 and 7.2 d, h, l & p), the proportion of negative C^2 values determined to be in-cloud are relatively constant with height, although a slight decrease in percentage with height is evident, particularly at coarser scales such as $\widehat{\overline{\Delta}} = 16\Delta$. However, at the coarsest scale analysed ($\widehat{\overline{\Delta}} = 64\Delta$), data from cases in QSS show that the proportion of negative IC C^2 values occurring follows a trend with height which differs from that observed in the high-resolution data. The IC profiles at this coarse resolution instead exhibit a similar trend to that of the CFE profiles, with the percentage of negative C^2 values increasing with height, as shown in Figures 7.2 d, h, l & p. The percentage of negative C^2 values identified as being in-cloud is also much higher in the $\widehat{\overline{\Delta}} = 64\Delta$ data, compared to the higher resolution data for all cases. The coarsest resolutions see the percentage of negative C^2 values reaching counts of between 40% to 60% in-cloud, depending on the case and parameter.

In contrast to the QSS data, the ARM case demonstrates a different behaviour occurring in-cloud during the developmental stages of the cumulus clouds. During these times, the percentage of negative IC C^2 values IC peaks in the low-to-mid levels of the CL. The coarse $\widehat{\overline{\Delta}} = 64\Delta$ data shows these peaks appearing to increase in height with time until, by the time the system has reached a QSS, the peak in percentage of negative values classified as being IC occurs at the CT level.

7.1.1 The Impact of Effective Grid Scale on Cloud Cover

As mentioned in previous sections, the number of points being classified as in-cloud increases with scale. This affects the proportion of negative values identified as either IC or part of the CFE, as the ratio of IC points to CFE points changes with scale. Consequently, this section explores the influence of scale on cloud cover. The way this dependency is addressed impacts how the number of negative values in a regime is interpreted. Two methods of presenting this information are examined: (1) the ratio of negative to positive C^2 values within a given flow regime, and (2) negative values expressed as a percentage of the entire horizontal domain (this is the method employed in the previous section). This section highlights the differences in cloud cover as the filter scale varies and examines how these changes affect the two methods of data presentation.

Figure 7.3 a shows the percentage of the CL that is classified as being IC for each case study at various effective grid spacings. At an effective grid spacing of $\widehat{\overline{\Delta}} = 4\Delta$ the BOMEX case, along with ARM at 10:30L and 16:30L show approximately 12% of the CL being cloudy. At the same resolution the ARM case at 14:30L shows 19% cloud cover, while 35% of the cloud layer is considered IC at 12:30L. As the $\widehat{\overline{\Delta}}$ values increase, the percentage of cloud cover also increases. This is because, as the filter is applied to the cloud field, the q_{cl} content is dispersed from the clouds to the surrounding area, effectively enlarging the clouds but diluting their water content. This continues as the filter scale increases until the entire cloud layer has some positive value of $q_{\text{cl}} \geq 1 \times 10^{-7}$ at each point. As a result, in the majority of cases, 100% of the CL is classified as being IC by the time $\widehat{\overline{\Delta}} = 128\Delta$. The cases which have a higher percentage of cloud cover at $\widehat{\overline{\Delta}} = 4\Delta$ show the CL reaching 100% cloud cover at smaller filter scales than the other cases. For example, the 12:30L ARM case shows the entire CL domain being classed as IC at $\widehat{\overline{\Delta}} = 32\Delta$. However, the analysis of the coarse resolution data conducted so far focuses on the $\widehat{\overline{\Delta}} = 64\Delta$ scale. At this effective grid scale, the majority of cases show near 100% cloud cover, but still show some signatures of the typical behaviour of a Smagorinsky parameter within a CTBL.

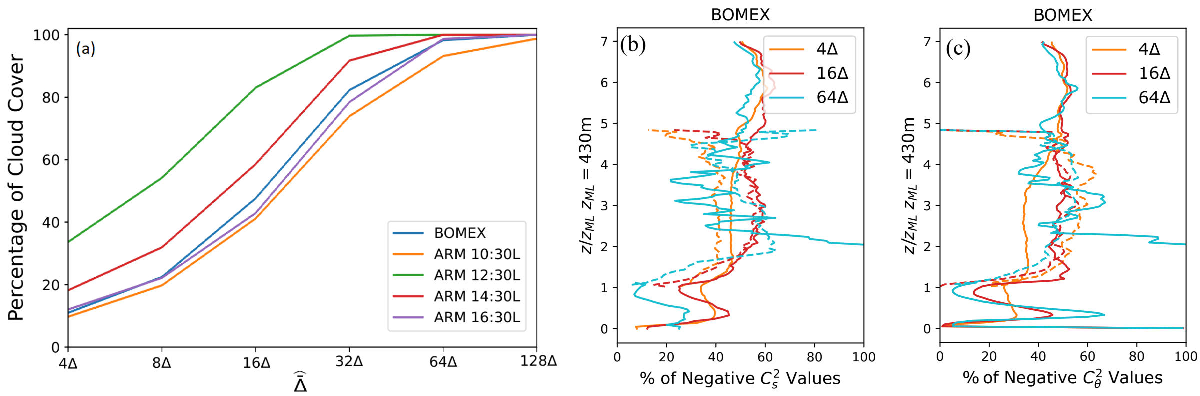

Figure 7.3: *(a) The percentage of the total CL domain that is classified as IC at each $\widehat{\overline{\Delta}}$ for every case. (b) Profiles of the percentage of each distinct region of the CTBL that show negative C_s^2 values. The number of negative C_s^2 values occurring IC, as a percentage of the total number of IC points at each height level, are shown by the dashed lines. The solid lines shows the equivalent for the NC areas.*

The variation in cloud cover between scales can be accounted for when plotting the number of negative C^2 values. This adjustment can be achieved by displaying the C^2 data as a ratio of negative to positive values in each specific regime, rather than as a percentage of the entire CL. The plots resulting from this method show very similar percentages of negative values occurring IC as in the CFE for each scale analysed. The percentage of negative C_s^2 values in each flow regime of the BOMEX case is given in Figure 7.3 b as an example. There is a 2% to 7% difference between the percentage of IC values that are negative and the percentage of CFE values that are negative, when noise in the CFE at coarser resolutions is ignored. The high resolution $\widehat{\overline{\Delta}} = 4\Delta$ data shows approximately 45% of points IC and in the CFE experiencing negative C_s^2 values. This increases to approximately 60% for $\widehat{\overline{\Delta}} = 16\Delta$ and $\widehat{\overline{\Delta}} = 64\Delta$, however the $\widehat{\overline{\Delta}} = 64\Delta$ data shows a high level of noise.

It was found that the majority of profiles, both for different parameters and different cases, displayed the same tendencies in the CL when the number of negative C^2 values were plotted as a percentage of the specific regions. However, the exception to this trend is, as expected, in the C_θ^2 IC data shown in Figure 7.3 c. At high resolution, the number of negative C_θ^2 values in-cloud are approximately 25% higher than those in the CFE. As a result, the $\widehat{\overline{\Delta}} = 4\Delta$ data showed that 35% of the CFE regions experienced negative C_θ^2 values, while over 60% of the in-cloud areas displayed negative values. However, by the time the effective grid scale had coarsened to $\widehat{\overline{\Delta}} = 16\Delta$ there was very little difference between the percentage of negative C_θ^2 values and the percentage in the CFE. This similarity between the two regions at coarse resolution is most likely a result of the dilution of cloudy areas and dispersion of the q_{cl} into the CFE. Figure 7.3 a confirms that, at $\widehat{\overline{\Delta}} = 16\Delta$, 45% of the CL is classified as being IC. Therefore it would be expected for these two regions to begin to show shared characteristics at this filter scale. This remains true as the scale is coarsened further to $\widehat{\overline{\Delta}} = 64\Delta$, though noise was once again present in the CFE at this resolution. This noise is a result of the lack of points remaining which are still classified as being in the CFE.

7.2 Distributions of C Values at Various Filter Scales

The values taken by the Smagorinsky parameters for each variable across the model domain will now be analyzed in this section. The negative C^2 values in the field were clipped and set to zero, allowing for the calculation of the C fields. The distributions of values taken by each of the Smagorinsky parameters can then be plotted and analysed. These histograms of the C values have been produced at three different effective grid spacings for each case: $\widehat{\overline{\Delta}} = 4\Delta,\ 16\Delta,\ 64\Delta$. The histograms shown here are specific to the following flow regimes: ML, IC, and CFE. The frequencies have been normalized by the total number of points within each regime, meaning that the distributions represent the percentage of points in a given region that correspond to a specific C value. Note that all of the histograms presented in the following plots show a bimodal distribution, with one of the peaks occurring at zero. This peak is ignored throughout this analysis as it is an artefact of the clipping process applied to the negative C^2 values.

In the ML at $\widehat{\overline{\Delta}} = 4\Delta$, all four of the Smagorinsky parameters show an approximately symmetric distribution with a high standard deviation about a mean value of $C \approx 0.13$ (as depicted in Figures 7.4 a, d, g & j). In the ML, the BOMEX case is observed to have a flatter distribution than the ARM distributions for all parameters. This suggests that BOMEX has a higher variance of C values, despite BOMEX and ARM having similar mean values in the ML at $\widehat{\overline{\Delta}} = 4\Delta$. As the effective grid spacing increases toward grey-zone regimes, the distributions exhibit a rightward skew. This clearly demonstrates the reduction in the mean and mode of the Smagorinsky parameter values as the grid scale increases in size. When the effective grid scale is coarsened to $\widehat{\overline{\Delta}} = 64\Delta$, the ML Smagorinsky parameters for scalars θ, θ_L, and q_t exhibit not only a pronounced rightward skew, but also significantly higher leptokurtosis, particularly during the earlier stages of the diurnal cycle in ARM, compared to the higher-resolution distributions (see Figures 7.4 f, i & l). These distributions indicate that, at very coarse resolutions, the scalar parameters C_ψ experience a very high percentage of points in the ML tending towards zero, as shown in Figures 7.4 f, i & l. The momentum parameter however, shows a slightly weaker tendency toward zero at this scale. This is evidenced in Figure 7.4 c, which reveals that, while the distribution of C_s in the ML of the $\widehat{\overline{\Delta}} = 64\Delta$ data is also heavily skewed to the right, the mode is non-zero. This is especially evident when comparing the C_s distributions to the distributions for C_θ, C_{θ_L} and C_{q_t} in the ML for the same resolution.

The analysis now shifts to examining the distributions of the Smagorinsky parameter values within the CL. This layer has again been decomposed into two distinct flow regimes: the IC areas (Figures 7.6) and the CFE (Figures 7.5). It should be noted that, in the CL, the distributions of C values at coarse resolutions exhibit significantly lower densities, particularly for scalar parameters. This can be attributed to the fact that most C values in the CL are zero when $\widehat{\overline{\Delta}} \geq 64\Delta$. Evidence of this trend is seen in the profiles detailing the percentage of negative values with height in Figures 7.1 and 7.2. These figures outline that a high proportion of negative values occur in the CL, with approximately 60% of the CL being negative at this scale, and these values are subsequently clipped to zero. Additional evidence supporting the prevalence of zero values is provided in the next section, where Smagorinsky parameter profiles indicate that the mean values of $\langle C \rangle$ approach zero at $\widehat{\overline{\Delta}} \geq 64\Delta$ in both the IC regions and the CFE (Figures 7.7, 7.8 and 7.9).

The majority of the CL is made up of the homogeneous, near-laminar CFE. When the Gaussian filter is applied to the cloud field it spreads the properties of the CFE into the neighbouring small, but very turbulent cloudy areas, diminishing the in-cloud turbulence. Meanwhile, the q_{CL} content, which is used to define cloud boundaries, is dispersed over large areas which may not exhibit the typical properties of in-cloud regions. As a result of the combined effect of these processes, the areas identified as IC exhibit a near-laminar flow, causing the corresponding C parameters to approach zero as resolution decreases. In contrast, the ML is characterised by a highly turbulent and homogeneous flow across the entire domain. Therefore, while filtering will laminarize the ML flow somewhat, it is not to the same extent as observed in the CL. Consequently, the ML retains a near-normal distribution of non-zero C values (Figure 7.4), while the CFE (Figure 7.5) and IC (Figure 7.6) distributions of the Smagorinsky parameters exhibit a right skewness, even at higher resolutions.

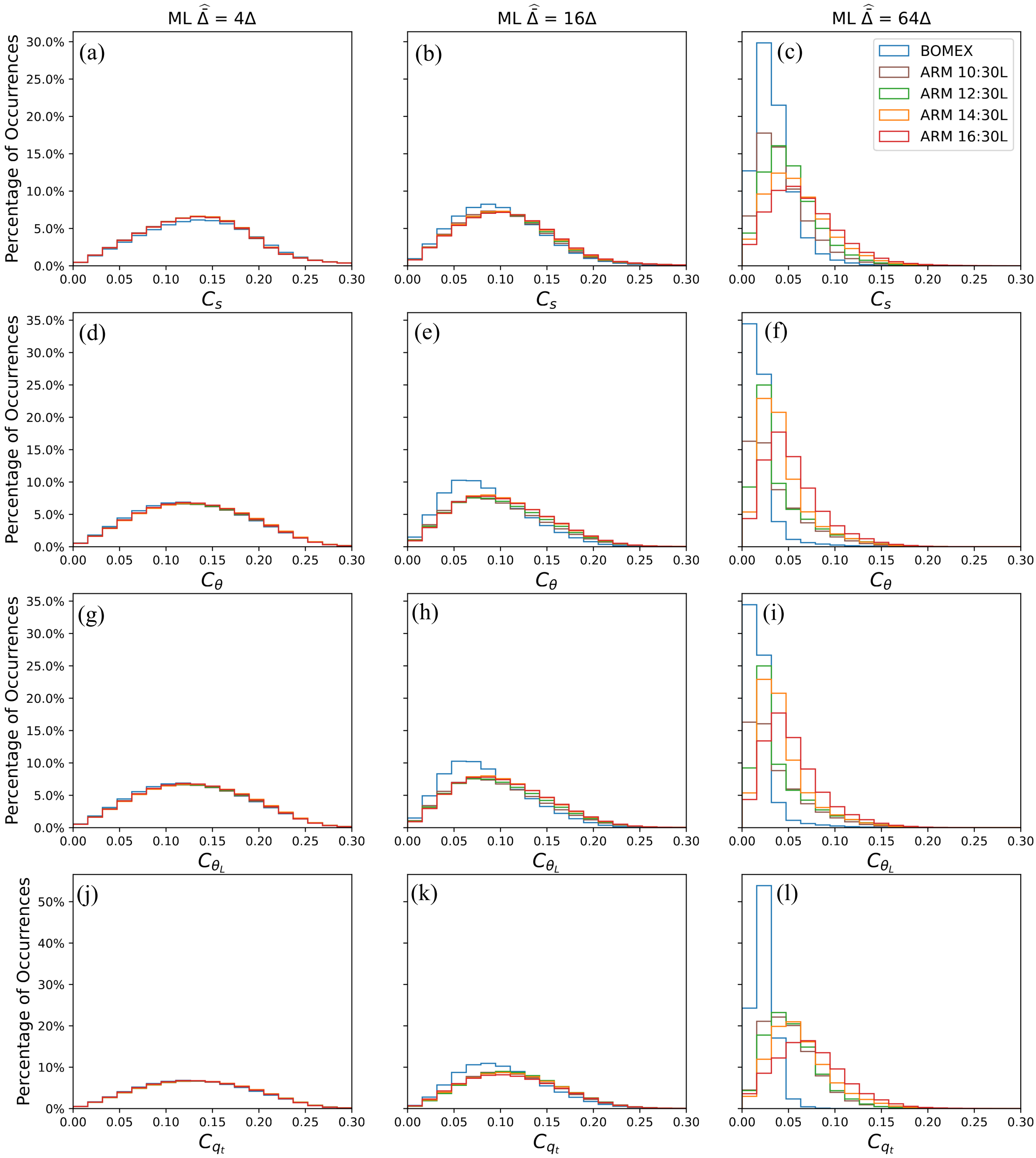

Figure 7.4: *The distributions of values of the Smagorinsky parameters in the ML for data filtered to effective grid spacings of* 4Δ*,* 16Δ*, and* 64Δ*. The rows show the distributions of the Smagorinsky parameters for* C_s*,* C_θ*,* C_{θ_L}*,* C_{q_t} *respectively. The frequency is shown as a percentage of the total ML.*

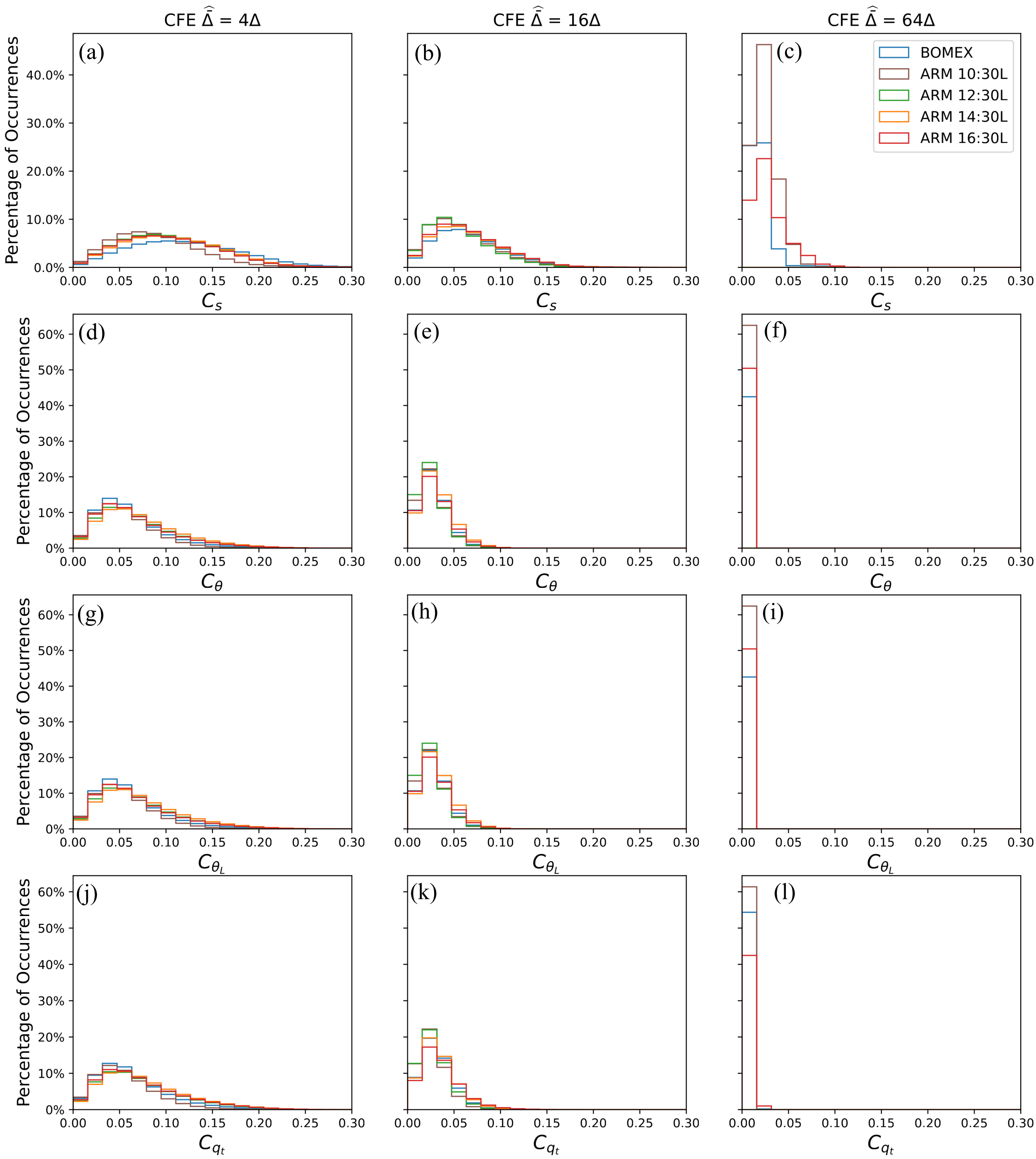

Figure 7.5: *Same as in Figure 7.4, but for C values in the CFE.*

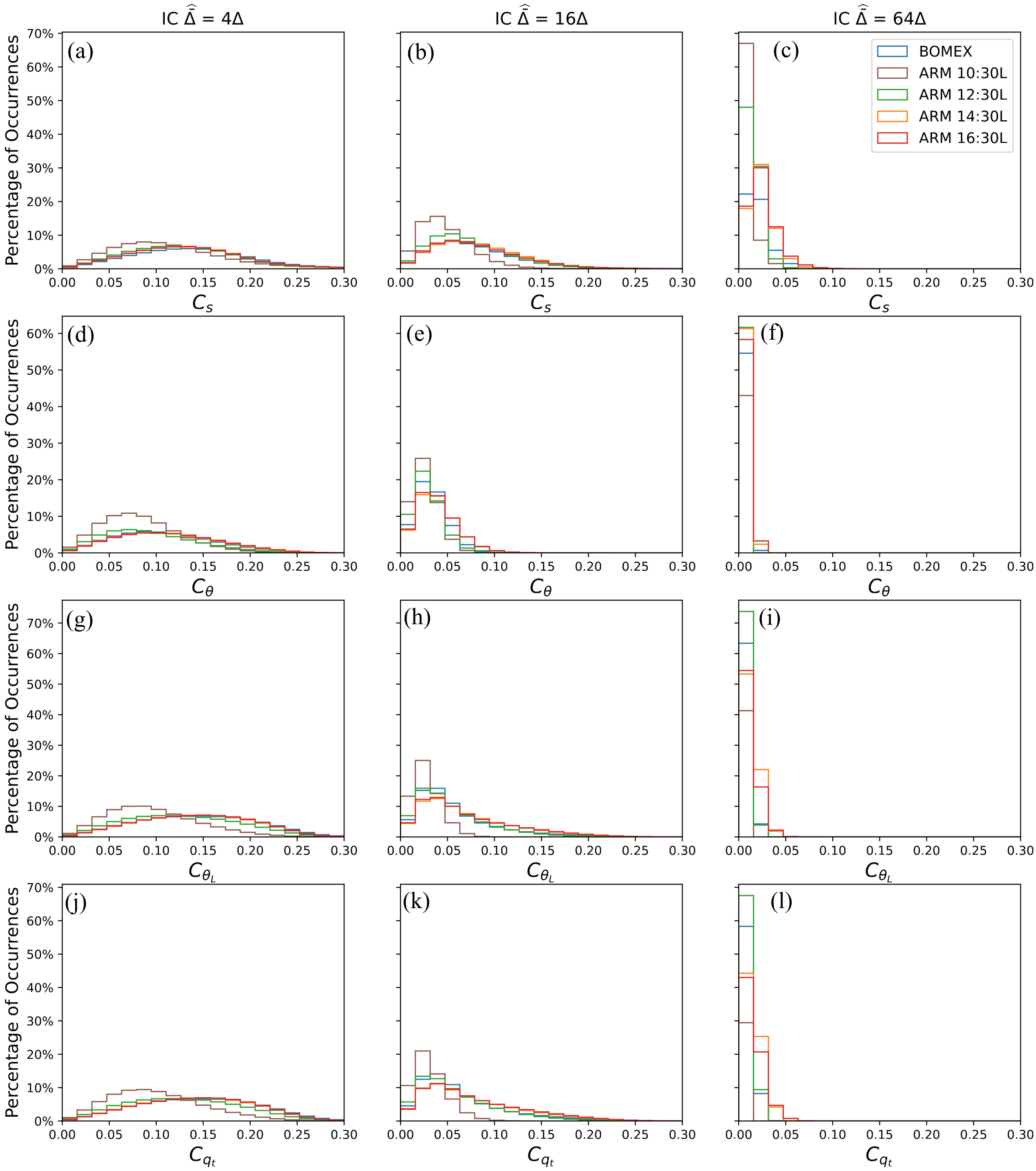

Figure 7.6: *Same as in Figure 7.4, but for C values IC.*

Firstly, Figure 7.5 demonstrates that in the CFE, all of the parameters at each filter scale show a right-skewed distribution, though this skew is least severe in the high resolution C_s data. The momentum parameter in the CFE shows a mode of $C_s \approx 0.08$ at $\widehat{\Delta} = 4\Delta$, reducing by half each time the effective grid spacing is raised to the power of two (see Figures 7.5 a, b, & c). This drives the CFE Smagorinsky parameters toward zero as the effective grid spacing is increased. The distributions for C_s in the majority of cases demonstrate a high variance, with Figure 7.5 a showing some extreme values reaching $C_s \approx 0.3$ in the CFE when $\widehat{\Delta} = 4\Delta$.

Meanwhile, the distributions of the scalar parameters C_ψ, for $\psi = \theta, \theta_L$, & q_t, exhibit a much more severe tendency to skew to the right, as presented in Figures 7.5 d, g, & j. All cases also show the distributions of the scalar parameter values as more leptokurtic, peaking at a modal value of $C_\psi \approx 0.05$. The tendency of the scalar parameters to show right-skewed distributions indicates that, while the majority of C_ψ values are near zero in the CFE, some points experience high Smagorinsky parameter values. Analysis conducted in Section 6.1 suggests that in the CFE, large C_ψ values are expected to occur near the cloud edges, where detrainment of heat and moisture from the cumulus cloud causes excess mixing and dissipation of these scalars in this region. Comparing the high resolution distributions of C_s to C_ψ demonstrates that, while the momentum structures persist in the CFE (evidenced by the higher mode and right-skewed tail in Figure 7.5 a), the heat and moisture structures do not persist in the CFE to the same extent (as demonstrated by the near-zero values of C_ψ, even at high resolution). Therefore the mixing and dissipation of heat and moisture is at a minimum in the CFE.

Secondly, the distributions of Smagorinsky parameter values from the cloudy regions of the CL are examined. It is evident that as $\widehat{\overline{\Delta}}$ increases and approaches the grey-zone, the Smagorinsky parameter values in-cloud again tend toward zero, particularly for the C_θ parameter (Figure 7.6 f). The distributions also become more skewed to the right when $\widehat{\overline{\Delta}}$ is large, with the majority of C parameters taking near-zero values as the resolution decreases. This behaviour is consistent with the previously discussed trends observed in the CFE, and occurs for the same reasons. Figure 7.6 a shows a strong agreement in the frequency of IC C_s values across all cases when compared to other flow regimes, which can display significant discrepancies in the C_s distributions between cases. This indicates that the mixing and dissipation of momentum is consistent within all shallow cumuli at high resolutions, rather than being case-dependent. The high resolution IC C_s values are approximately normally distributed about a mean of $C_s \approx 0.135$ with a large variance, as indicated by the long tails on the distribution, as seen in Figure 7.6 a. The exception to this is the 10:30L ARM case, which shows a distribution that is skewed slightly to the right, rather than being symmetric. This is likely due to the limited presence of clouds in the domain, as they are just beginning to form at this time. Consequently, the mixing and dissipation occurring in areas designated to be IC is relatively low.

The distributions of the C_ψ parameters in-cloud demonstrate the scalar dependence of the Smagorinsky parameter value. As resolution decreases, Figure 7.6 d e & f show that the C_θ values are decreasing toward zero slightly more rapidly than the other scalars. This reduction in non-zero values directly results from the lack of conservation of potential temperature θ in-cloud, giving rise to counter-gradient transport of θ within the clouds and causing the C_θ values to be forced toward zero in these regions. At high resolution the C_θ distribution is slightly skewed to the right while the C_{θ_L} and C_{q_t} parameters show flat, normal distributions of values in-cloud. This is similar to that of the IC C_s values, although the morning-time ARM C_{θ_L} and C_{q_t} values again show a right-skewed distribution. The distributions of C_{θ_L} and C_{q_t} are centred around 0.15, and both distributions are platykurtic (signifying a large variance). Meanwhile the C_θ distribution, while skewed to the right, shows a much shorter tail compared to the other scalar parameters. This is a result of the C_θ parameter tending to zero due to its lack of conservation in-cloud; in contrast the distributions of C_{θ_L} and C_{q_t} show longer tails to the right as the conservation properties of these scalars enable the Smagorinsky parameter to remain non-zero for longer as resolution is decreased.

Overall, it is evident that the Smagorinsky parameters depend strongly on the flow regime. The distributions analysed above also highlight the effect that resolution has on the parameter values, particularly in the IC and CFE areas of the CL. The role of scalar dependencies for the Smagorinsky parameters appears to be limited to the IC regions of the CTBL. The majority of the high-resolution Smagorinsky parameter data yields symmetric distributions, indicating that using the mean as a representative value within the different flow regimes is a reasonable choice for well-resolved data. Therefore the following section will focus on the mean values of C in the different regions of the CTBL flow and at different scales.

7.3 The Effect of Filter Scale on the Smagorinsky Parameter Profiles

The conditional averages of Smagorinsky parameter values in each flow regime have been calculated. Parameter values across multiple resolutions are analysed, using filtered data from the BOMEX and ARM test cases. The relationship between the Smagorinsky parameter value and filter scale as the system moves deeper into the grey-zone regime is investigated in this section. Additionally, this chapter introduces results from the scale-dependent dynamic Smagorinsky parameter calculation method, described in detail in Section 3.4.2. The resulting profile calculated using this approach is denoted with a subscript β, referring to the β parameter which measures the degree of scale dependence. Since the focus is on the relationship between C and $\widehat{\overline{\Delta}}$ within the distinct flow regimes, the DA plots are not presented here. However, the DA plots showed a similar general trend as the other flow regimes, with average C values decreasing fractionally when the effective grid spacing increases.The CU and CC profiles exhibit the same trends as the IC profiles, and because there is minimal distinction between the CU/CC areas and the overall IC average, these profiles are also not included. The IC and NC (encompassing the ML and CFE regions) profiles are presented in Figure 7.7 for the BOMEX case, and Figures 7.8 and 7.9 for the ARM case (NC and IC respectively).

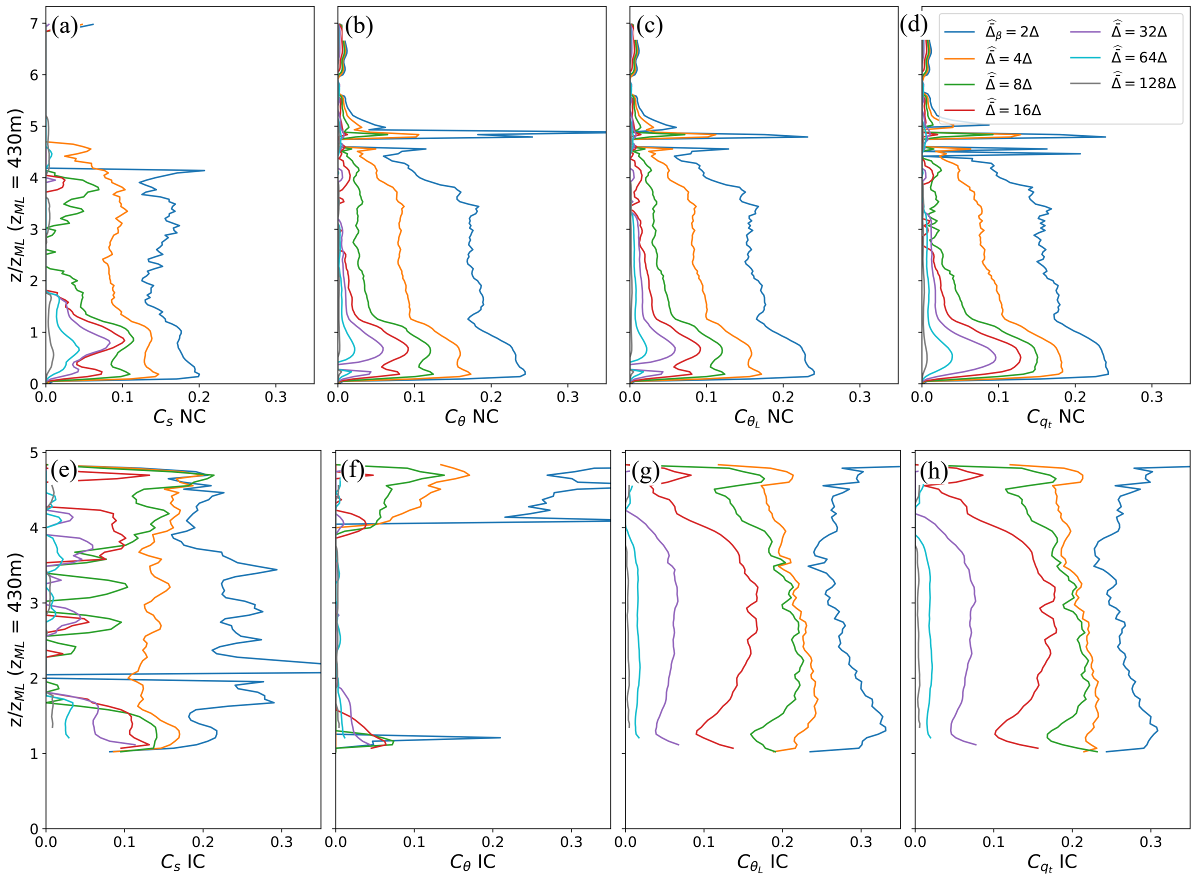

Figure 7.7: *Vertical profiles from the BOMEX case calculated by conditionally averaging over the (top panel) NC regions for (a) C_s, (b) C_θ, (c) C_{θ_L}, (d) C_{q_t} and (bottom panel) IC areas for the same variables.*

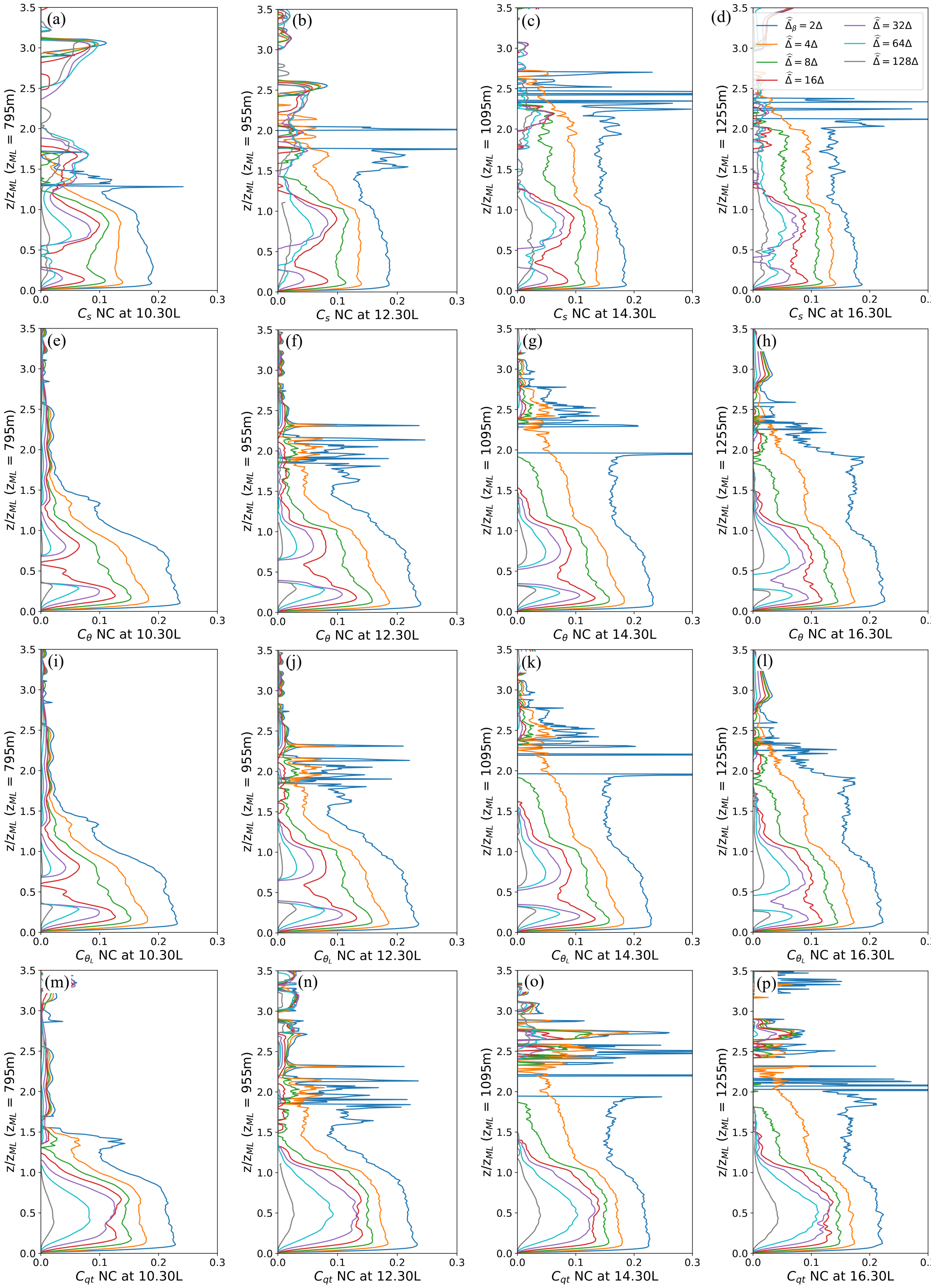

Figure 7.8: *Same as in Figure 7.7 but vertical profiles are from the ARM case at 4 different timestamps, conditionally averaged over the NC regions (the ML and the CFE of the CL).*

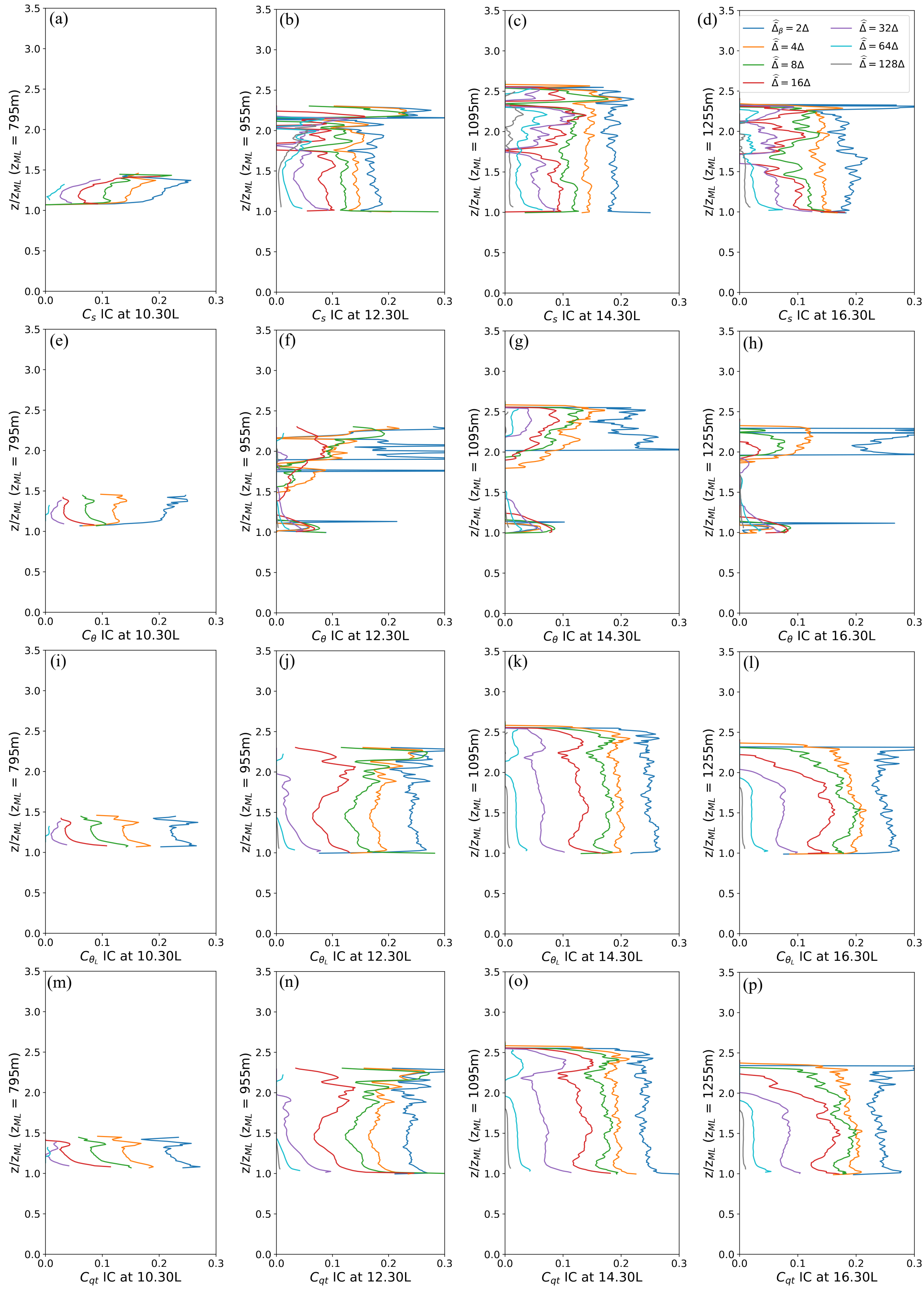

Figure 7.9: *Same as in Figure 7.8 but with profiles from the ARM data which have been averaged over the in-cloud areas.*

7.3.1 The Response of the Momentum Parameter to Changes in Filter Scale

The behaviour of the horizontally-averaged Smagorinsky parameter for momentum, $\langle C_s \rangle$, across multiple filter scales is examined first. In the ML, the $\langle C_s \rangle$ profiles exhibit a peak at the surface, which is clearly evident at all filter scales in the BOMEX (Figure 7.7 a) and ARM cases (Figures 7.8 a - d), becoming more apparent as the resolution coarsens. Above this surface-level peak, the higher resolution $\langle C_s \rangle$ profiles for each case exhibit a relatively constant value with height throughout the ML at high resolutions. At the effective grid spacings of $\widehat{\overline{\Delta}} = 16\Delta$ both cases display a local minimum occurring in the $\langle C_s \rangle$ profile at the mid-ML level, which is exacerbated as the resolution coarsens further. The highest resolution profile, $\widehat{\overline{\Delta}}_\beta = 2\Delta$, computed using the scale-dependent method (and therefore depending on both the 4Δ and 8Δ profiles) shows the average ML value of $C_s \approx 0.185$ for both the BOMEX and ARM profiles. The profiles within the ML exhibit incremental differences between scales, with the profile shifting to the left, indicating a decrease in the average value as $\widehat{\overline{\Delta}}$ increases. The increments between resolutions are reasonably consistent across both scales and cases. However, when the resolution reaches 32Δ, the dynamic Smagorinsky method begins to fail, particularly in the mid-ML, where $\langle C_s \rangle$ collapses to zero. This is expected, as such filter scales are within the scales typically considered to be in the grey-zone regime for a CTBL.

The ML shows a distinct reduction in $\langle C_s \rangle$ with height as the capping inversion is approached, above which, the profiles reduce to lower values in the CFE. Both the BOMEX case (see Figure 7.7 a) and the ARM case during the later times of 14:30L and 16:30L (Figures 7.8 c & d) show the high-resolution values of $\langle C_s \rangle$ remaining uniform with height in the CFE, at consistently lower values than their corresponding $\langle C_s \rangle$ in the ML. However, at coarser resolutions, the BOMEX case and the early stages of the diurnal cycle in the ARM case (10:30L and 12:30L) show the C_s profile in the CFE collapsing to zero in the mid-CL ($z/z_{\mathrm{ML}} \approx 2.5$ in the BOMEX case and $z/z_{\mathrm{ML}} \approx 1.5$ in the ARM cases) at $\widehat{\overline{\Delta}} \geq 8\Delta$ (see Figures 7.7 a and 7.8 a & b). The same behaviour emerges in the C_s profiles at the mid-CL of the ARM case during the later times, but for data with $\widehat{\overline{\Delta}} \geq 16\Delta$. Therefore all cases exhibit the dynamic method breaking down in the CFE at higher resolutions than is observed in the ML, indicating that the CFE is more sensitive to grey-zone issues than the ML. Furthermore, the difference between $\langle C_s \rangle$ values at each scale is not uniform in the CFE, as it was in the ML. Instead, in the CFE the difference in $\langle C_s \rangle$ values between scales compounds as the effective grid spacing is increased.

The IC values of $\langle C_s \rangle$ from the ARM case (see Figures 7.9 a - d) show much more consistent increments between scales than the CFE profiles from the same case. The IC profiles are observed decreasing in value from $\langle C_s \rangle \approx 0.2$ at 16:30L (Figure 7.9 d) to zero as the corresponding $\widehat{\overline{\Delta}}$ increases to 128Δ. Despite the systematic and regular intervals at which the $\langle C_s \rangle$ value decreases with increasing scale, the dynamic method still begins to fail in the mid-CL at coarser resolutions. This is evident in the IC profiles of $\langle C_s \rangle$, which indicate that the latest three times during the ARM case (Figures 7.9 b - d) experience $\langle C_s \rangle \to 0$ at the height level corresponding to $z/z_{\mathrm{ML}} \approx 1.7$ when $\widehat{\overline{\Delta}} \geq 16\Delta$. The BOMEX data experiences the failings of the dynamic method at higher resolutions than ARM, with the $\widehat{\overline{\Delta}} = 8\Delta$ case showing the $\langle C_s \rangle$ profile oscillating between zero and values near 0.09. This also aligns with the resolution at which the dynamic method failed in the BOMEX CFE. Subsequently, the calculation of $\widehat{\overline{\Delta}_\beta} = 2\Delta$ for the BOMEX case is affected (see Equation 3.49), making this profile noisy and inconsistent in value with height. At each effective grid scale, the IC $\langle C_s \rangle$ values are higher than the corresponding CFE values, by about 0.03 for BOMEX and 0.04 for ARM.

7.3.2 The Effect of Filter Scale on the Scalar Parameters

The behaviours of the Smagorinsky parameters for scalars are analysed here, both for cloud-conserved (θ_L and q_t) and non-conserved (θ) variables. Firstly, looking at the behaviour of the planar averaged Smagorinsky profiles in the ML of the BOMEX case: Figures 7.7 b - d show that each of the scalar profiles have a distinctive parabolic shape in the ML, which is particularly evident in the $\langle C_{q_t} \rangle$ profile. The profiles of $\langle C_\psi \rangle$, for each scalar ψ, experience maximum values in the mid-ML, while minimum values occur at the surface and ML capping inversion. Furthermore, the $\langle C_\theta \rangle$ and $\langle C_{\theta_L} \rangle$ profiles demonstrate a peak in values at the surface in both the BOMEX and ARM cases. Above the surface, the ARM case demonstrates a more constant profile with height than BOMEX for the higher resolutions (see Figures 7.8 e - p). As the effective grid spacing increases in size, these profiles decrease in value and shift to the left toward zero.

At the coarser resolutions, when $\widehat{\overline{\Delta}} \geq 32\Delta$, the parameter profiles of the θ and θ_L scalars (Figures 7.7 b & c) show the average values in the mid-levels of the ML collapsing to zero. This is a key signature of the dynamic method failing in the mid-ML for the thermodynamic variables and is evident at this resolution in the ARM case also (Figures 7.8 e - l). However, early signs of breakdown are observed at $\widehat{\overline{\Delta}} = 16\Delta$ in the mid-ML of both cases, with the profile approaching zero at these scales. If the dynamic method had not failed for $\langle C_\theta \rangle$ and $\langle C_{\theta_L} \rangle$ in the mid-ML, values from the other levels in the ML suggest that $\langle C_\psi \rangle$ would not have decreased to zero in the ML until $\widehat{\overline{\Delta}} = 128\Delta$ in the BOMEX case, while in the ARM case it would have remained non-zero throughout the ML for all resolutions analysed. The dynamic method does not fail in the ML when calculating the C_{q_t} parameter, suggesting that the failure might result from the averaging data with a peak in values at the surface. The reduction in $\langle C_\theta \rangle$ and $\langle C_{\theta_L} \rangle$ values above the surface peak appears to amplify as $\widehat{\overline{\Delta}}$ increases, until the mid-ML value reduces to zero. Therefore, the peak in values at the surface might be the cause of the dynamic method failure in the mid-ML.

At the top of the ML, both the BOMEX (Figures 7.7 b - d) and ARM (Figures 7.8 e - p) cases show the scalar parameters decreasing rapidly in value to their CFE values. The difference between the average $\langle C_\psi \rangle$ values between the CFE and ML becomes more pronounced as $\widehat{\overline{\Delta}}$ increases. This suggests that the ML and CFE respond at different rates to changes in the filter scale. Above the ML capping inversion, the CFE profile is roughly constant with height. However, the difference between the $\langle C_\psi \rangle$ profiles at each scale in the CFE grows as resolution increases toward $\widehat{\overline{\Delta}}_\beta = 2\Delta$. All the cases show that, for data with an effective grid spacing of $\widehat{\overline{\Delta}} \geq 16\Delta$, the corresponding profiles are at or very near zero. This collapse to zero occurs at a higher resolution in the CFE than in the ML, further supporting the hypothesis that the ML and CFE have different scale dependencies. The spikes in value at the top of the CL which are evident in the CFE scalar parameter profiles are possibly due to the entrainment of warmer, drier air at certain points from the overlying statically stable layer. The high values reached by these spikes in the $\widehat{\overline{\Delta}}_\beta = 2\Delta$ profile are artificial, and are disregarded in this analysis as a result. The spikes result from the misalignment of areas where the $\widehat{\overline{\Delta}} = 8\Delta$ has collapsed to (zero due to the dynamic method failing) when the $\widehat{\overline{\Delta}} = 4\Delta$ has not, therefore forcing the ratio of these two parameters to tend to infinity.

The IC profiles reveal a pronounced difference between conserved and non-conserved variables for both the BOMEX data (Figure 7.7 f - h) and ARM data (Figures 7.9 e - p). In the mid-CL region, the non-conserved θ variable produces IC profiles of $\langle C_\theta \rangle$ which are forced to zero at all scales. This is evident in each case, apart from the earliest stage of the diurnal cycle where the clouds are just initiating (see Figure 7.9 e). During the morning, the cloud is just forming and the IC $\langle C_\theta \rangle$ maintains its value from the ML thermal which is feeding the developing cumulus cloud. This is because the C_θ has not had a chance to respond to the developing counter-gradient transport of θ in the newly formed cloud. Once the cumulus clouds have developed, from 12:30L onwards, the ARM case agrees with the BOMEX case, and shows an IC profile of $\langle C_\theta \rangle = 0$ in the mid-CL. However, at the levels just above the capping inversion z_{ML} and just below the cloud top (CT), non-zero values of $\langle C_\theta \rangle$ are observed in the IC profile. At the cloud base (CB), which occurs just above the z_{ML}, the IC profile of $\langle C_\theta \rangle \approx 0.1$ for all high-resolution cases. This level experiences a non-zero $\langle C_\theta \rangle$ in-cloud as the overshooting thermals transport air from the ML, which has not yet experienced condensation, into the cloud. Therefore, in the region where the thermal retains its essential characteristics (near the cloud base), the heat fluxes of θ are not yet counter-gradient. This results in the corresponding Smagorinsky parameter remaining positive and non-zero in this area.

The IC profiles in the area below the CT also show non-negative $\langle C_\theta \rangle$ values, however, these values near the CT are much higher than those just above z_{ML}. The $\langle C_\theta \rangle$ values in-cloud reach approximately the same value as the IC profiles of $\langle C_{\psi_c} \rangle$ for the conserved scalars ψ_c, when both profiles have the same $\widehat{\overline{\Delta}}$. This is because the positive IC $\langle C_\theta \rangle$ values result from the entrainment of warm air from the statically stable overlying layer. In the BOMEX case, the overlying stable layers were initialised at $z = 2000\,\mathrm{m}$, while the positive $\langle C_\theta \rangle$ values occur between $1720\,\mathrm{m} \leq z \leq 2064\,\mathrm{m}$. Similarly in the ARM case the statically stable layer was initialised at $z = 2500\,\mathrm{m}$, while the positive $\langle C_\theta \rangle$ values occur between $z/z_{\mathrm{ML}} \approx 1800$ and $z/z_{\mathrm{ML}} \approx 2850$, depending on the time during the diurnal cycle. This signifies that the overshooting tops of the cumulus clouds encroach into the statically stable air, and in doing so entrain warm, dry air downwards into the upper levels of the cumulus clouds. Similar to what was observed just above the z_{ML}, this warm, entrained air maintains its characteristics near the cloud top. As condensation has not yet occurred in these entrained air parcels, the $\langle C_\theta \rangle$ profile is positive in these regions. The $\langle C_\theta \rangle$ profile at CT reaches higher values compared to those at CB, due to the significant heat content of the overlying air which is entrained at CT. The resulting IC value $\langle C_\theta \rangle$ have approximately the same value as the $\langle C_{\psi_c} \rangle$ at the corresponding $\widehat{\overline{\Delta}}$. When noise is disregarded, these IC profiles at the CT show the expected trends of a decrease in $\langle C_\theta \rangle$ values as $\widehat{\overline{\Delta}}$ increases.

The conserved variables (ψ_c) show IC profiles of $\langle C_{\theta_L} \rangle$ and $\langle C_{q_t} \rangle$ taking distinctive shapes, depending on the stage of cloud development, as evident from the diurnal cycle depicted in Figures 7.9 i - p. The morning (10:30L) IC profiles of $\langle C_{\theta_L} \rangle$ and $\langle C_{q_t} \rangle$ (Figures 7.9 i & m) show the highest values at CB, presumably as this is where the heat and moisture are being injected into the developing cumulus clouds. The IC C_{ψ_c} profiles show a peak in value near the CT during the morning, while in the mid-CL, values are approximately equal to their corresponding ML values for a given $\widehat{\overline{\Delta}}$. As the cloud layer develops during the diurnal cycle, passive clouds form and the ARM data at 12:30L reveal that the IC profiles of $\langle C_{\theta_L} \rangle$ and $\langle C_{q_t} \rangle$ take a slightly parabolic shape, as presented in Figure 7.9 j & n. At this time, a minimum is evident in mid-CL levels of the C_{θ_L} and C_{q_t} IC profiles, while the maximum values occur at the CT and CB. By the rapid cloud development stage, the ARM data from 14:30L in Figure 7.9 k & o indicates that the IC $\langle C_{\theta_L} \rangle$ and $\langle C_{q_t} \rangle$ profiles are near-constant with height. As the system reaches a quasi-steady state at 16:30L, the ARM data reveals the $\langle C_{\theta_L} \rangle$ and $\langle C_{q_t} \rangle$ IC profiles once again take on a parabolic shape. However, the parabola is the inverse of the one observed at 12:30L, with the mid-CL now experiences the maximum parameter values, while the minimum values occur at CB and CT. This aligns with the BOMEX data, which is also in a quasi-steady state, and shows a similar shaped profile for these parameters when IC. These distinct shapes taken by the IC profiles during each specific phase of the diurnal cycle hold true for every $\widehat{\overline{\Delta}}$ until the profile collapses to zero, which occurs at $\widehat{\overline{\Delta}} = 128\Delta$ for the majority of cases.

7.3.3 Trends in the Smagorinsky Parameter's Response to Filter Scale

The focus of this section is to investigate the impact of the effective grid scale on the average parameter values for every variable in each of the flow regimes. The average value of the profile within each flow regime is calculated and plotted against the corresponding $\widehat{\bar{\Delta}}$. The layer used for computing the profile average excludes any extreme values near the top and bottom boundaries of each regime. This ensures that the average for each regime accurately reflects the typical values throughout the majority of the layer, without being contaminated by boundary effects or overshooting/entraining motions from other layers. The resulting averages for each parameter calculated using the BOMEX data are presented in Figure 7.10, while Figure 7.11 depicts the averages as calculated using the ARM data.

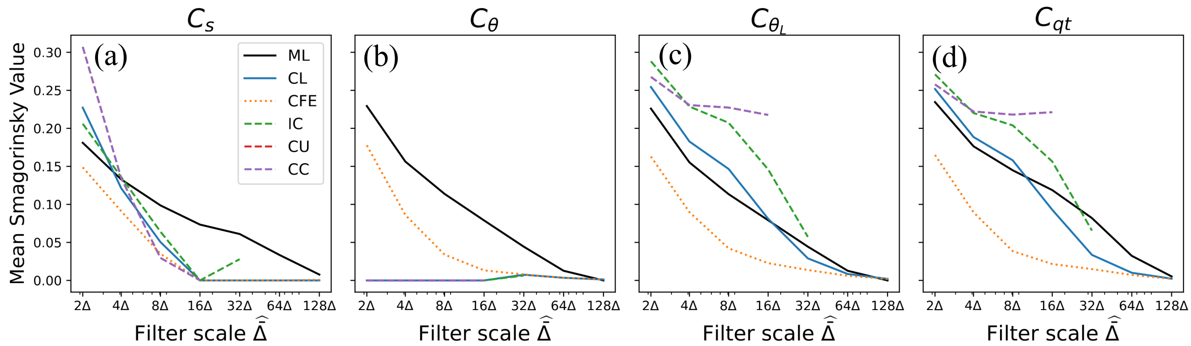

Figure 7.10: *Mean values of the planar averaged Smagorinsky parameters for each variable of interest in the BOMEX case. The averages are conditioned on being in the flow regimes, the criteria of which have been described previously.*

In the ML, the relationship between the average values of each of the Smagorinsky parameters and the effective grid spacings are presented by the solid black line in Figures 7.10 (BOMEX) and 7.11 (ARM). The average C_s parameter value in the BOMEX case shows a gradual decrease in value as $\widehat{\bar{\Delta}}$ increases, as evident in Figure 7.10 a. The ARM case however, experiences a more rapid decrease in C_s value as $\widehat{\bar{\Delta}}$ increases during the morning time (10:30L) as shown in Figure 7.11 a. As time progresses and the CTBL develops, the slope of C_s versus $\widehat{\bar{\Delta}}$ becomes less steep and trends from the ARM case at 12:30L, 14:30L, and 16:30L show much better agreement with the BOMEX data for this parameter. The Smagorinsky parameters for the other scalars in the ARM case also show this trend of having a steeper slope during the morning phase when plotted against $\widehat{\bar{\Delta}}$, compared to the slopes of the same parameter versus scale during the later phases of the diurnal cycle. This suggests that the Smagorinsky parameter experiences a greater sensitivity to scale during the morning, which is consistent with the expectation that eddies in the CTBL are smaller in scale at this time. As a result, these small-scale eddies are more likely to be influenced by the grey-zone compared to larger-scale motions.

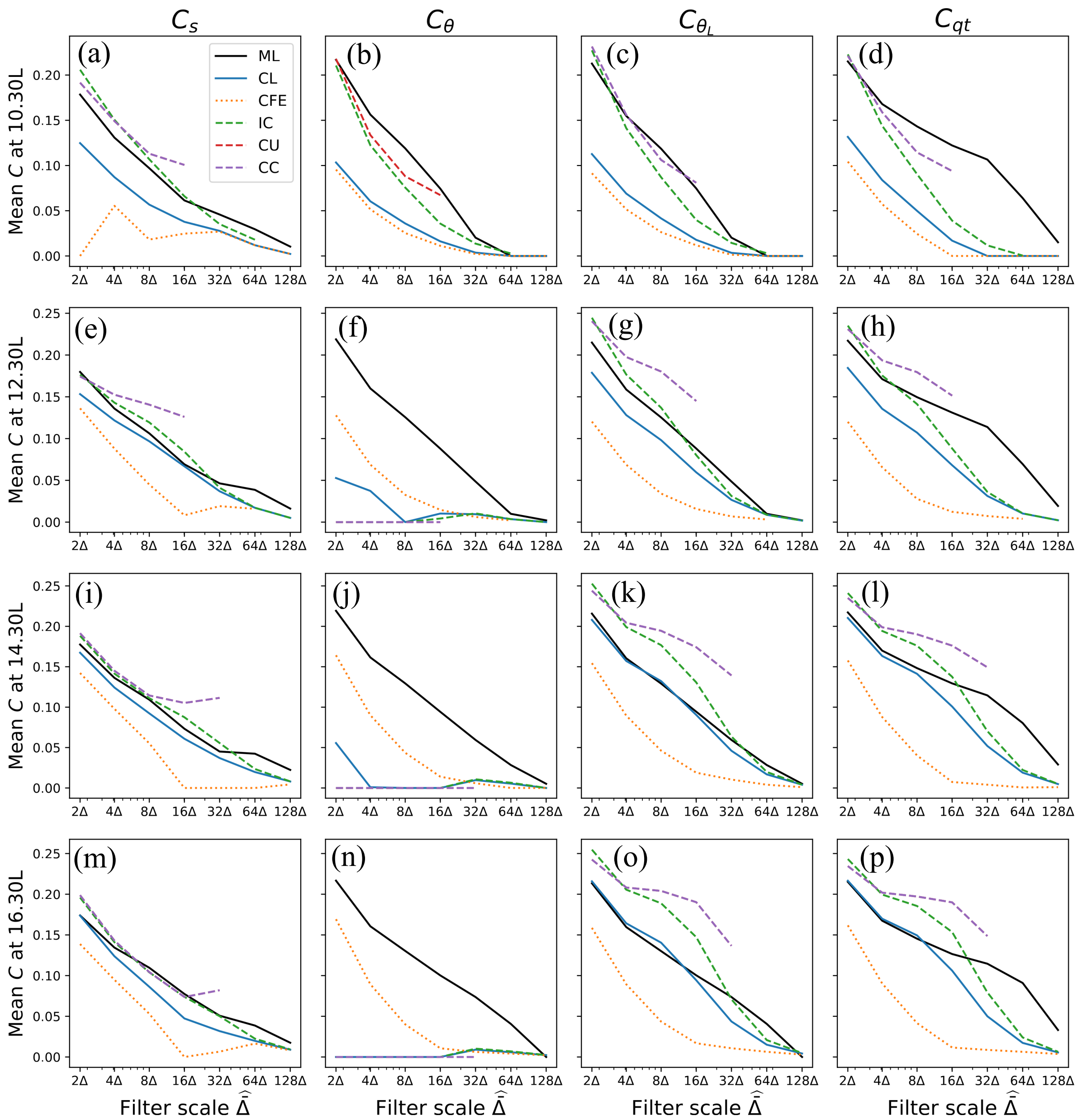

Figure 7.11: *Same as in Figure 7.10, however using data from four different times during the diurnal cycle of the ARM case study. Each row contains data from a distinct time step, while each column illustrates the scale dependence trend of a specific parameter.*

The average Smagorinsky parameter values in the ML for both the thermodynamic variables show very similar trends as $\widehat{\Delta}$ is increased for both the BOMEX data (Figures 7.10 b & c) and the ARM data at each timestep analysed (Figures 7.10 b, c, f, g, j, k, n & o). This is in line with the theory, as the difference in conservation properties does not affect the θ and θ_L variables in the ML due to the lack of condensation occurring here, therefore their behaviour should be identical in this flow regime. Meanwhile, the relationship between C_{q_t} and scale is distinct from the thermodynamic variables in the ML. Figure 7.10d and Figures 7.11 d, h, l & p show the ML of both BOMEX and ARM recording larger C_{q_t} values than C_θ values for the same $\widehat{\Delta}$. This is demonstrated by the gradual slope of the solid black line (ML average) in the C_{q_t} versus $\widehat{\Delta}$

plots at higher resolutions, indicating a weaker scale dependence for the C_{q_t} averages when $\widehat{\overline{\Delta}} \geq 16\Delta$. In contrast, coarser resolutions exhibit a slightly stronger scale dependence. This suggests that the moisture structures in the ML are larger than the heat structures, and therefore the moisture diffusion parameter is less affected by filter scale than the heat parameters, up until the point where the grey-zone begins to impact the q_t circulation, at approximately $\widehat{\overline{\Delta}} \approx 16\Delta$.

The average values of C_s in the CFE exhibit similar trends with filter scale across all cases apart from the morning time ARM case, as shown by the dotted orange line in Figure 7.11 a. The value of $C_s \rightarrow 0$ at $\widehat{\overline{\Delta}}_\beta = 2\Delta$ at 10:30L in the ARM case is an erroneous point resulting from the scale-dependent calculation using β. This method relies on the $\widehat{\overline{\Delta}} = 8\Delta$ data, which experienced a failure of the dynamic method when calculating C_s in the CFE, and therefore contaminated the $\widehat{\overline{\Delta}}_\beta = 2\Delta$ calculation. As a result, this data point is disregarded, and all other cases show the average C_s values in the CFE exhibiting the same trend with scale: Figures 7.10 a and 7.11 e, i & m show that the C_s in the CFE decreases from 0.15 (in the BOMEX case) at $\widehat{\overline{\Delta}}_\beta = 2\Delta$ to zero when $\widehat{\overline{\Delta}} = 16\Delta$. The effective grid-scale at which this plateau occurs aligns with where the grey-zone was observed impacting the C_{q_t} parameter. The behaviour of the Smagorinsky parameter for scalars in the CFE is the same for θ, θ_L, and q_t in each of the cases as $\widehat{\overline{\Delta}}$ increases (see Figures 7.10 b - d for the BOMEX case and Figures 7.11 b - d, f - h, j - l, and n - p for the ARM case). The C_s parameter in the CFE exhibits a gradual decrease in value with scale, plateauing at $C_s = 0$ when $\widehat{\overline{\Delta}} = 16\Delta$. Meanwhile, the C_ψ demonstrates a more rapid decrease in value with scale at high resolutions, but slows to a gradual decrease at coarse resolutions. The steeper slope between $\widehat{\overline{\Delta}}$ and C_s in the CFE confirms that in this region C_s values are, on average, more scale-dependent than those in the ML.

The average values of the in-cloud Smagorinsky parameter for momentum, C_s, also show a stronger scale dependence than the ML averages in the BOMEX case (Figure 7.10 a) and the ARM case during the morning phase of the diurnal cycle (Figure 7.11 a). In these cases the IC values of C_s decrease rapidly with increasing filter, and scale when $\widehat{\overline{\Delta}} \geq 16\Delta$. The later times in the ARM case show the average IC C_s values having similar values to the ML C_s values for the same $\widehat{\overline{\Delta}}$, as presented in Figures 7.11 e, i & m. All cases, except the 10:30L ARM data, show average IC values of $C_\theta = 0$ for all $\widehat{\overline{\Delta}}$. This result was expected, as the non-conservation of θ IC is known to force the corresponding Smagorinsky parameter to zero. The morning time ARM data shows that before the θ parameter has been affected by lack of conservation issues, the C_θ IC values follow a similar trend as the other scalars, decreasing in value as $\widehat{\overline{\Delta}}$ increases.

The BOMEX case shows the in-cloud parameter averages for the conserved scalars sharing behaviours, as presented in Figures 7.10 c & d. The IC averages for C_{θ_L} and C_{q_t} reach higher values than the corresponding ML values, and these parameters also show a different scale dependence in-cloud than the ML averages. This scale dependence is observed to become more severe when $\widehat{\widetilde{\Delta}} \geq 16\Delta$, as C_{θ_L} and C_{q_t} values decrease with scale at a much more rapid rate past this point. The later times in the ARM case, corresponding to the stages of rapid cloud development and the quasi-steady state, also show C_{θ_L} following these same trends in-cloud as the C_{q_t} parameter (see Figures 7.11 k, l, o & p). However, the earlier times in the diurnal cycle demonstrate very different trends, with the C_{θ_L} parameter showing similar average IC values as the ML averages of C_{θ_L} (Figures 7.11 c & g). Meanwhile, the moisture parameter exhibits completely different behaviour, with the IC averages of C_{q_t} decreasing much more rapidly with increasing $\widehat{\widetilde{\Delta}}$ than the ML averages of C_{q_t} for the earlier times in the diurnal cycle (see Figures 7.11 d & h). This implies that the IC C_{q_t} parameter is severely sensitive to scale during the early stages of the ARM case. It might also suggest that distinctions can be made between the Smagorinsky parameter values for scalars, depending on which scalar fluxes are the primary drivers of convection during the cloud initiation and growth stages of the diurnal cycle.

Note that the trend of average values versus scale in the CU and CC regions has also been included in these plots. As these regions often give the same average values, the trend lines completely overlap and only one trend line is visible in the majority of the plots. Once the effective grid spacing is greater than 16Δ, the CU and CC structures can no longer be distinguished within the clouds. This aligns with what would be expected in a grey-zone scale LES, as the CU and CC structures exist on scales that can not be resolved at coarse resolution. This is in line with the overall trend that is evident for many of the Smagorinsky parameters across all the different cases. The impact of the grey-zone takes effect at $\widehat{\widetilde{\Delta}} \approx 16\Delta$ for all flow regimes and Smagorinsky parameters in both the BOMEX and ARM cases.

7.4 Summary of the Parameter Responses to Filter Scale

The overarching result from analysis conducted throughout this chapter is that the impact of filter scale on the value of a Smagorinsky parameter is both regime-dependent and variable-dependent. In the ML, the Smagorinsky parameters for each variable demonstrate similar behaviours at each scale. Negative value C^2 parameters make up approximately 30% to 40% of the ML depending on the case and the filter scale. The distribution of C values in the ML is also reasonably consistent across the range of Smagorinsky parameters at each scale. An increase in filter scale sees a corresponding decrease in the average Smagorinsky parameter value in the ML. This is highlighted by the linear decrease evident in the C versus $\widehat{\widetilde{\Delta}}$ trend graphs. Overall the behaviour of the Smagorinsky parameters for each variable with scale are reasonably consistent within the ML.

However, the effect of scale on the Smagorinsky parameter values in the CFE and IC regions is much less consistent than in the ML. The percentage of negative C^2 values relative to the total area of a specific flow regime show similar values for the IC regions and CFE (if the C_θ data is omitted), with about 45% of both the IC and CFE areas encountering negative C^2 values at high resolution. At coarser resolutions, up to 60% of the IC and CFE areas experience negative C values. The IC and CFE distributions of C are also heavily impacted by changes in filter scale, and are observed to exhibit a stronger scalar dependence compared to the ML values. This scale dependency is evident in the regime-specific profiles of the C parameters. These profiles demonstrate that the average Smagorinsky parameter values in the IC regions and CFE decline to zero significantly faster than those in the ML. Moreover, this decrease in $\langle C \rangle$ does not occur at consistent intervals with increasing filter scale, which is again in contrast to the behaviour observed in the ML data. This rapid decline to zero which is observed in the CL parameter profiles is again highlighted in the C versus $\widehat{\widehat{\Delta}}$ trend graphs. These plots illustrate the IC parameters undergo a steep decline when plotted against the effective grid-scale, leading to a much faster reduction of the C parameters to zero as the filter scale increases compared to the ML parameters.

7.4.1 Responses to the Research Questions

(1) To what extent does the filter scale systematically influence the prevalence of negative C^2 values across the different flow regimes in the CTBL?

Filter scale has a systematic effect on the percentage of the ML experiencing negative Smagorinsky parameter values. As the effective grid spacing increased, the proportion of the ML exhibiting negative C^2 values decreased steadily at reasonably regular intervals, corresponding to the changes in filter scale. However, as filter scale increased the percentage of negative C^2 values in the IC areas and CFE did not show a specific trend. Instead, the change in filter scale primarily affected the cloud cover in the field, with the percentage of cloud increasing with scale. This contaminated the data, as an increasing number of points were classified as IC, which is not the expected outcome for an LES model operating even at coarse scales. This is a limitation of the filtering method when employed in the cloud field. Therefore a systematic effect of the filter scale on the percentage of negative Smagorinsky parameters is evident only in the ML of the CTBL.

(2) How does the variability of the Smagorinsky parameters change with scale in each distinct flow regime?

The effect that filter scale has on the variability of the Smagorinsky parameters is entirely regime-dependent. The ML shows the least scale dependency, though the parameter values still show a clear decrease in variability as filter scale increases. The majority of C values in the ML cluster around the mode at coarse resolutions. However, the distribution becomes severely asymmetric at coarse filter scales as the mode is near-zero and negative values are not permitted. In the IC regions and CFE the variability of Smagorinsky parameters also decreases as filter scale increases, though the primary reason for this seems to be due to the lack of non-zero C^2 values at coarse resolutions. Therefore the overall effect of increasing the effective grid scale is to lower the variability of the Smagorinsky parameters.

(3) What are the key scale dependencies of the Smagorinsky parameters in each of the flow regimes?

The Smagorinsky parameter values in the ML show a consistent decrease in value for all variables as the filter scale increases. However, the Smagorinsky parameters IC and in the CFE show case and variable dependencies, as well as a scale dependency. As filter scale increases, the majority of cases show the C_s parameter decreasing in value more rapidly in-cloud than in the ML. In contrast, most cases show the IC values of C_{θ_L} and C_{q_t} with higher average values than the IC C_s values. Additionally, the decline in value of these scalar parameters as the filter scale increases is less pronounced compared to the trends observed for both IC C_s and all the parameters in the ML. The Smagorinsky parameter values in the CFE follow a different trend to those IC. The C_s values in the CFE exhibit a trend of decreasing linearly with increasing filter scale that is similar to that observed in the ML values, but at consistently lower magnitudes. However the CFE values of C_{θ_L} and C_{q_t} show an exponential decrease with increasing filter scale. These findings suggest that incorporating scale dependencies into a new Smagorinsky parametrization would also require considering regime-specific dependencies, in order to ensure its effectiveness.

8 Accounting for Systematic Dependencies in a New Mixing Length Parametrization

Flow regime and scale have been observed to have a substantial impact on the average Smagorinsky parameter value in the CTBL. These dependencies are as of yet unaccounted for in the standard Smagorinsky models. The work conducted in this chapter aims to derive a parametrization that can account for the key dependencies seen across all cases. The goal is to capture the general behaviour of the flow, without employing the computationally expensive dynamic model.

To achieve this, work from previous chapters has been synthesised to produce the new parametrization. Analysis carried out in Chapter 6 suggests that the planar averaged Smagorinsky parameter values for momentum, heat, and moisture each follow distinct, near-universal mean profiles within specific flow regimes. Within the NC regions (encompassing both the IC areas and CFE) the C values demonstrate a clear height dependency in the CTBL. These profiles follow a distinct shape, which arises as a result of differences in stability regimes between the mixed layer, the capping inversion and cloud base, the cloud layer, and the cloud top. Meanwhile, work detailed in Chapter 7 demonstrates that the Smagorinsky parameter profiles at different resolutions have similar profile shapes, but that these profiles are shifted to lower values as the resolution decreases. Therefore, the values of each Smagorinsky parameter are not only dependent on height but are also strongly dependent on the filter scale of the simulation. Smagorinsky parameter values corresponding to coarser resolutions are consistently lower than those corresponding to higher resolutions. This is also evident in Figures 7.10 and 7.11, which show the average parameter value, for different regions in the flow, decreasing as the filter scale increases. The functional relationships derived between the Smagorinsky parameters and (a) height and (b) scale across different regions of the CTBL are incorporated into the MONC model. The updated model is then tested using the BOMEX and ARM case studies, and the output is compared to the previous simulations produced using the standard Smagorinsky scheme.

This chapter aims to identify the primary dependencies of the Smagorinsky parameters for momentum, heat, and moisture within the CTBL. The overarching aim of this work is to address the following research questions:

- What systematic relationships can be established between the Smagorinsky parameters and height within each flow regime of the CTBL? Can scale dependence be accounted for in these functional relationships?
- What effect does this empirical parametrization have on the modelled fields?

In order to investigate these questions, results from the previous chapters are compiled to produce the necessary universal relations, outlined in Section 8.1 and 8.1.1. The various relations are applied in the model in different configurations, as detailed in Section 8.3. This enables the evaluation of both the individual effects of each relationship and the interactions that occur when multiple relationships are incorporated into the model. The analysis of the outputs from the altered MONC model is contained in section 8.4.

8.1 Variations in the Smagorinsky Parameters Between Distinct Regions in the CTBL

It is important to highlight that partitioning the flow into its distinct flow regimes was considered necessary prior to deriving a new parametrization. This step was required because attempts to establish a relationship between height and the domain-averaged (DA) values of the Smagorinsky parameters, particularly the C_θ parameter, failed in the CL. This is because the majority of the CL is made up of the conditionally stable non-cloudy air, which experiences very low levels of turbulence. This is in sharp contrast to the IC regions, which exhibit significantly enhanced levels of turbulence. Therefore the mixing and dissipation of quantities in these areas are very different, and as such would benefit from distinct parametrization schemes.

The flow within a CTBL can be partitioned into regions with distinct characteristics: the ML, CL, NC, IC, CU, CC. Furthermore, height dependency and scale dependency seem highly influential on the value of a given Smagorinsky parameter in a CTBL simulation. These findings can be used to develop and test various parametrizations. At high resolutions the differences between values in the IC, CU, and CC regions are minimal, as shown in Figures 7.10 and 7.11. It is evident that, at coarser resolutions, a disparity emerges between the average IC values and the average CU and CC values. However, very few grid points would resolve CU or CC regions at such coarse resolutions. Therefore, the new parametrization partitions the flow into just two regions: the cloudy regions, where the Smagorinsky parameters are denoted as C_{IC}, and non-cloudy environment, which includes the ML and the CFE, with the corresponding parameters referred to as C_{env}.

The behaviours of the Smagorinsky parameters in these specific regions are now assessed using the BOMEX and ARM data. The C_{env} profiles were constructed upon the conditional averages over the NC environment, scaled by each case's ML capping inversion height. The C_{IC} parametrization was derived from the profiles of conditional averages within the IC areas, scaled by the corresponding CL depth for each case. This allows the height and stability dependencies of the Smagorinsky parameter values for momentum, heat (both θ and θ_L), and moisture to be parametrized in these flow regimes.

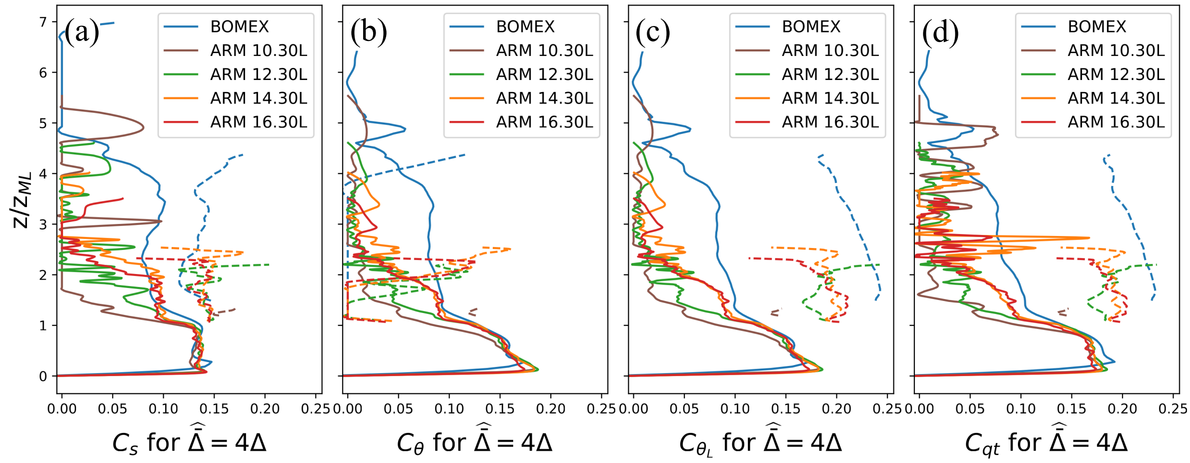

Figure 8.1: *Horizontally averaged profiles of the Smagorinsky parameter for (a) momentum, (b) potential temperature, (c) liquid water potential temperature, and (d) moisture for the BOMEX and ARM case studies. Solid lines show the Smagorinsky parameter profiles in the non-cloudy environment for each case, scaled by the corresponding inversion height. The dashed lines show the profile of the average in-cloud Smagorinsky parameter values for each case, again scaled by inversion height.*

Figure 8.1 shows profiles of the Smagorinsky parameter for momentum, heat, and moisture for both the in-cloud (dashed lines) and non-cloudy (solid lines) areas of the CTBL. The ML values for all the parameters (see plots in Figure 8.1 a - d) show the BOMEX and ARM profiles collapsing along a universal profile, once height has been scaled by the relative mixed layer capping inversions z_{ML}. For each case, the capping inversion height was calculated by finding the z level where the planar averaged vertical eddy heat flux $\overline{\mathrm{w}'\theta'}$ is at a minimum. In the ML the C_s parameter averages to approximately 0.137, while the C_ψ experiences consistently higher values for all scalars ψ. The thermodynamic parameters show a peak value at the surface of $C_\theta \approx C_{\theta_L} \approx 0.19$, while above this peak, $C_\theta \approx C_{\theta_L} \approx 0.17$. The profiles of the moisture parameter C_{q_t} also demonstrate a peak at the surface, though it is less extreme in magnitude. Similar to the heat parameter, at the surface the ARM profiles show $C_{q_t} \approx 0.19$, however above this the moisture parameter maintains a larger value than heat, with $C_{q_t} \approx 0.18$. This indicates that the moisture structures are slightly larger than the heat and momentum structures in the CTBL, which aligns with the results discussed in previous chapters. Furthermore, the BOMEX case shows higher C_{q_t} values than the ARM case, presumably due to the slightly higher ratio of latent heating to sensible heating occurring in the marine (BOMEX) CBL compared to the CBL over land (ARM).

The in-cloud averages for the Smagorinsky parameters show values oscillating around near-constant values with height, though most cases show a slight decrease in the Smagorinsky parameter value with height for the conserved scalar parameters. The noise observed in the IC profiles arises from the limited number of parameter values to sample from in the IC region, as cumulus clouds constitute less than 10% of the CL. The oscillations in the IC profile at the cloud top (CT) reflect the very limited number of cloudy points at the highest levels of the CL. It might also indicate the presence of gravity waves caused by disturbances from the overshooting CTs. In contrast, while the profiles in the CFE exhibit fluctuations around a near-constant value with height, these oscillations have a significantly smaller amplitude compared to those in the IC profiles. This is attributed to the substantially larger number of points in the CFE available to average over. It is believed that if more time-averaging had been possible (currently the BOMEX case has been averaged over 3 time stamps, while the ARM case has been averaged over 2 time stamps), these oscillations would reduce significantly. Therefore, the deviations from the near-constant value are treated as noise and disregarded, resulting in fixed values being proposed for all three different Smagorinsky parameters in the CFE. Above the CL values oscillate around zero. It is believed that the Smagorinsky values only deviate from zero due to the formation of gravity waves.

The values taken by the C_s parameter IC are slightly larger than in the ML, with all cases in this flow regime showing a $C_s \approx 0.15$. However, the values of the Smagorinsky parameter for scalars in the CL, both IC and in the CFE, depict a clear case dependency, whereas their corresponding ML profiles did not. Firstly, looking at the CFE profiles, it is evident from Figures 8.1 b - d that not only do these profiles experience a case dependence, but a slight scalar dependence is also evident in the BOMEX profiles. The CFE profiles from the BOMEX case depict higher C_{θ_L} and C_{q_t} values just above the z_{ML} level compared to the C_θ values at this height. This highlights the role that the increased latent heat release plays in this region, compared to the ARM case which is primarily driven by sensible heat fluxes (recall that BOMEX is mostly driven by latent heat fluxes). The case dependency of the C_ψ parameters within the CFE is clearly evident by the range of values taken by each profile. The earlier times of 10:30L and 12:30L during the ARM diurnal cycle show lower values of C_ψ for all scalars in the CFE. However, once the CL has developed, the later times of 14:30L and 16:30L show C_ψ taking the same value of approximately 0.1 in the CFE. The BOMEX case also takes this value in the C_θ profiles. However, as previously mentioned, the conserved parameters show slightly higher values of $C_{\theta_L} \approx C_{q_t} \approx 0.11$.

	BOMEX	ARM 10:30L	ARM 12:30L	ARM 14:30L	ARM 16:30L
z_{ML}	430 m	795 m	955 m	1095 m	1255 m
z_{CT}	1960 m	1070 m	2200 m	2840 m	3000 m

Table 8.1: *The height levels at which the ML capping inversion (z_{ML}) and cloud top height (z_{CT}) occur for each of the different cases.*

The differences in heights between the cases suggest that a different scaling variable is needed in this layer. While it is not possible to scale the CFE by a different parameter in the current set-up, as it is a part of the NC profile which is already scaled by z_{ML}, it is possible to scale the IC profiles by a length scale which is more appropriate to the CL layer. The heights of the ML capping inversion and CT heights for each case are listed in Table 8.1.

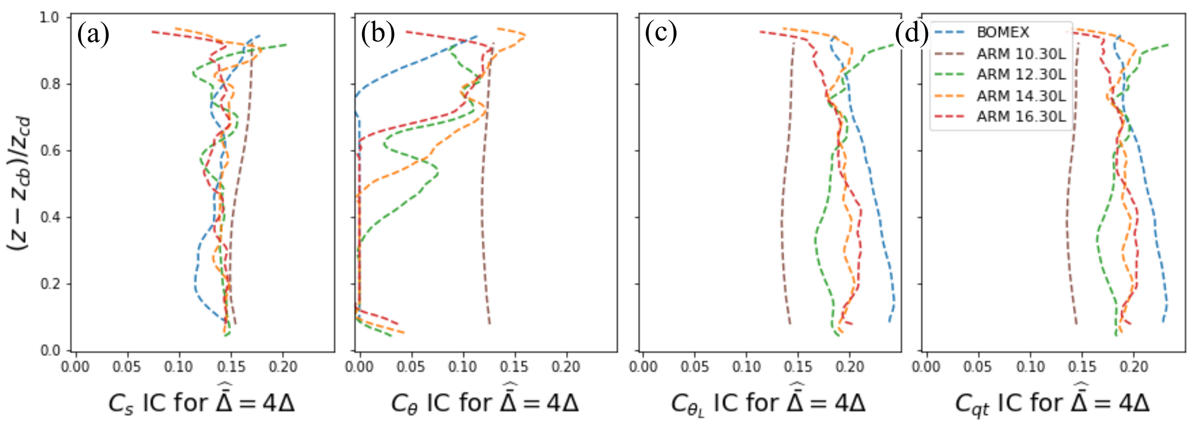

Figure 8.2: *Horizontally averaged profiles of the in-cloud values for (a) $\langle C_s \rangle$, (b) $\langle C_\theta \rangle$, (c) $\langle C_{\theta_L} \rangle$ and (d) $\langle C_{q_t} \rangle$ for the BOMEX and ARM cases. The profiles for each case are scaled by z_{cd}, the cloud layer depth, and re-positioned so that the cloud base is at 0 on the z-axis.*

The horizontally averaged values of the in-cloud parameters for each variable are now analysed. When the IC profiles of these parameters are scaled by their respective CL depths, they demonstrate similar features between the cases. The most notable similarity is observed in the profiles of the $\langle C_s \rangle$ parameter in Figure 8.2 a, which reveals that the average values across all cases remain constant with height at a value of $\langle C_s \rangle \approx 0.15$. In most cases, the $\langle C_\theta \rangle$ profiles converge to zero in the lower levels of the CL, as expected due to the counter-gradient transport of θ occurring in this region. The exception to this behaviour is the morning-time ARM case, which likely deviates because the cloud has only just initiated. As a result, at 10:30L in the ARM case the $\langle C_\theta \rangle$ parameter retains the value associated with the thermal feeding the developing cloud ($\langle C_\theta \rangle \approx 0.13$), as the parameter has not yet adjusted to the counter-gradient heat transport emerging in the cloudy region. However, in the upper levels of the CL, all the other cases show the $\langle C_\theta \rangle$ profiles transitioning from zero toward $\langle C_\theta \rangle, \approx, 0.13$, which is the consistent value observed across the entire height for the 10:30L ARM profile. This is likely due to significant entrainment in the CT region, which introduces large amounts of heat from the warm, dry, and statically stable FT, which overlies the CL. This process enhances the mixing of θ in the CT area, leading to an increase in the $\langle C_\theta \rangle$ value at this level.

In cloudy areas, the Smagorinsky parameters for conserved scalars (θ_L and q_t) exhibit different behaviours compared to the momentum parameters. Figures 8.2 c & d depict the IC $\langle C_{\theta_L} \rangle$ and $\langle C_{q_t} \rangle$ profiles increasing in average value with time, with this trend being particularly evident in the lowest 60% of the CL. However, Figure 8.2 c shows the three ARM IC profiles for times between 12:30L and 16:30L converging to $\langle C_{\theta_L} \rangle \approx 0.2$ in the top 60% to 80% of the CL. Meanwhile, the BOMEX case consistently shows slightly higher $\langle C_{\theta_L} \rangle$ values IC than the ARM cases. In the 10:30L ARM case, the IC $\langle C_{\theta_L} \rangle$ profile remains steady with height at approximately 0.15, with the $\langle C_{q_t} \rangle$ IC parameter exhibiting a similar profile at this time as well. Figure 8.2 d depicts the $\langle C_{q_t} \rangle$ profiles following a similar trend to $\langle C_{\theta_L} \rangle$, however, this figure also shows the BOMEX converging to $\langle C_{q_t} \rangle \approx 0.2$ in the upper 60% to 80% of the CL.

The simple linear relationship of each IC C parameter with height, combined with the limited case dependency, indicates that it is sufficient, as a first approximation, to specify a constant value for the IC Smagorinsky parameters. However, the NC areas of the CTBL did not show such simple dependencies, and therefore a parametrization for these regimes must now be defined.

8.1.1 Parametrization of Systematic Behaviours Resulting from Regime Variations with Height.

In the mixed layer of every case study, Figure 8.1 shows that the C_s, C_θ, C_{θ_L}, and C_{q_t} profiles decrease rapidly to zero as the surface is approached. This behaviour is captured in MONC model by employing Blackadar's method, given in Equation 8.1, when calculating the mixing length value for a given height level z in the LES.

$$\frac{1}{\ell_{\text{mix}}^2} = \frac{1}{(C\Delta)^2} + \frac{1}{(kz)^2} \tag{8.1}$$

This equation sets the mixing length as a function of the Smagorinsky parameter C, grid spacing Δ, height z, and the von Karman constant $k = 0.4$. The new altered model will still use this equation when determining the mixing length. However, the previous "standard Smagorinsky" method of fixing C to a set value will be replaced with a new parametrization for C_{env}, which is derived in the following section. Therefore, the reduction in mixing at the surface does not need to be included in the new C parametrization as it will be accounted for by Equation 8.1.

Above the surface layer, Figure 8.1 shows that for all cases, every parameter remains at a constant value throughout the depth of the mixed layer. At the top of the mixed layer, the capping inversion causes the C_{env} values for all four quantities to reduce. At the ML capping inversion height z_{ML}, most cases for all three parameters show C_{env} rapidly reducing to the value of the Smagorinsky parameter in the CFE, C_{CFE}. The only exception to this is the C_{q_t} parameter which in the CFE shows its average value decreasing with height. This is in contrast with the behaviour of all other parameters, which show a near-constant value with height in the CFE. The Smagorinsky parameter profiles near the CT exhibit a gradual decrease, approaching zero as they reach the level above the CL. The typical shape of the Smagorinsky parameter profile, as it decreases from the ML values to the CFE values, can be captured using the following function, based upon similar reasoning to Blackadar's formula to decrease the mixing length near the surface:

$$f = f(z, z^*_{\mathrm{ML}}) = \frac{\left(\frac{z}{z^*_{\mathrm{ML}}}\right)^{\gamma_f}}{\left(1 - \frac{z}{z^*_{\mathrm{ML}}}\right)^2} \tag{8.2}$$

$$z^*_{\mathrm{ML}} = \alpha_f z_{\mathrm{ML}} \tag{8.3}$$

The decrease in the Smagorinsky parameter profiles at the top of the CL, z_{CT}, can be described by a similar equation:

$$g = g(z, z^*_{\mathrm{CT}}) = \frac{\left(\frac{z}{z^*_{\mathrm{CT}}}\right)^{\gamma_g}}{\left(1 - \frac{z}{z^*_{\mathrm{CT}}}\right)^2} \tag{8.4}$$

$$z^*_{\mathrm{CT}} = \alpha_g z_{\mathrm{CT}} \tag{8.5}$$

where z is the height level, and α_f is a constant used to set the point at which the capping inversion z_{ML} begins to affect turbulence in the levels below it. Similarly, α_g sets the point where the inversion at the CT level, z_{CT}, begins to affect turbulence in the levels below it. The γ_f variable is also set to a constant (which must be a positive integer, typically of order $\mathcal{O}(10)$) used to set the steepness of the slope between the C_{ML} and C_{CFE} values. Meanwhile, γ_g sets the steepness between the Smagorinsky parameters in the CFE and the FT. The exact values of these variables, specific to each parameter, are detailed in Table 8.2 below. This "shape function" can then be used to define a parametrization for C_{env}, which is the generalised profile of a Smagorinsky parameter in the cloud-free environment of a CTBL.

$$C_{\mathrm{env}} = (C_{\mathrm{ML}} - C_{\mathrm{CFE}} - C_{\mathrm{FT}}) \left[\frac{1}{(f+1)}\right] + (C_{\mathrm{CFE}} - C_{\mathrm{FT}}) \left[\frac{1}{(g+1)}\right] + C_{\mathrm{FT}} \tag{8.6}$$

where C_{ML} is the average value of the Smagorinsky parameter in the mixed layer, C_{CFE} is the average value of the Smagorinsky parameter in the cloud-free environment of the cloud layer, and C_{FT} is the average value of the Smagorinsky parameter in the FT above the cloud layer. Equation 8.6 can be rewritten in the form of a weighted function which more clearly represents the transition between the Smagorinsky values in the ML and those in the CFE:

$$C_{\text{env}} = \mathrm{w}_1 C_{\text{ML}} + (\mathrm{w}_2 - \mathrm{w}_1) C_{\text{CFE}} + (1 - \mathrm{w}_1 - \mathrm{w}_2) C_{\text{FT}} \tag{8.7}$$

where the weights, w_1 and w_2 are defined as:

$$\mathrm{w}_1 = \frac{1}{f+1} \tag{8.8}$$

$$\mathrm{w}_2 = \frac{1}{g+1} \tag{8.9}$$

This height dependency based parametrization of C_{env} is a function of $z, z_{\text{ML}}, \alpha_f, \alpha_g, \gamma_f, \gamma_g$, C_{ML}, C_{CFE}, and C_{FT}. Increasing the value of z^*_{ML} shifts the transition between C_{ML} and C_{CFE} upward to higher z levels. Therefore, when approaching the ML capping inversion from the bottom up, the decrease in the C_{env} value begins at a higher z level when $z^*_{\text{ML}} > z_{\text{ML}}$ than would occur if $z^*_{\text{ML}} = z_{\text{ML}}$. As the slope between the ML and CFE does not change, a change in z^*_{ML} merely changes the height at which the transition occurs. Therefore the converse also holds true: when approaching the inversion from the top down, the C_{env} values begin to increase at a higher z level when $z^*_{\text{ML}} > z_{\text{ML}}$ compared to if $z^*_{\text{ML}} = z_{\text{ML}}$. The α_g parameter follows the same principles, but influences the behaviour of the C_{env} profile at the transition zone between the CFE and the OL. Meanwhile, the range of height levels over which this transition zone between C_{ML} and C_{CFE} occurs is governed by the value of the γ_f variable. This is because, when the ratio set out by Equation 8.2 is near unity, the corresponding z levels will have C_{env} values between C_{CFE} and C_{ML}, as per Equation 8.6. Therefore γ_f determines the range of z levels where the numerator and denominator of Equation 8.2 are of the same order of magnitude, with higher powers of γ_f decreasing the number of height levels involved in the transition and therefore steepening the slope of the C_{env} profile at the inversion. Once again, the γ_g variable follows the same principles, but for the transition between the CFE and the overlying OL.

To fit this new parametrization of C_{env} to the case study data, trial and error was used to find suitable values for the input variables. The best-fit values that were assigned to the input variables for Equation 8.6 are detailed in Table 8.2 below and the resulting parametrized profiles of C_{env} for momentum, heat, and moisture are plotted in Figure 8.3 below.

	$C_{s,\,\text{env}}$	$C_{\theta,\,\text{env}}$	$C_{q_t,\,\text{env}}$
α_f	1.4	1.6	1.45
γ_f	16	3	6
α_g	1.2	1.45	1.3
γ_g	16	4	6
C_{ML}	0.137	0.175	0.175
C_{CFE}	0.095	0.09	0.1
C_{FT}	0	0	0

Table 8.2: *Values taken by the various variables input to the C_{env} function in order to parametrize the Smagorinsky profiles for momentum $C_{s,\,env}$, heat $C_{\theta,\,env}$ (note that $C_{\theta_L,\,env}$ uses the same values as $C_{\theta,\,env}$), and moisture $C_{qt,\,env}$ in the non-cloudy environment of a CTBL simulation.*

When these values are substituted into Equation 8.7, the resulting generalised function effectively captures the key influences that the variation in flow regime and stability with height have on the Smagorinsky parameter. These functions, specific to each Smagorinsky parameter focused on in this study, are presented in the following Figures (8.3 a - d).

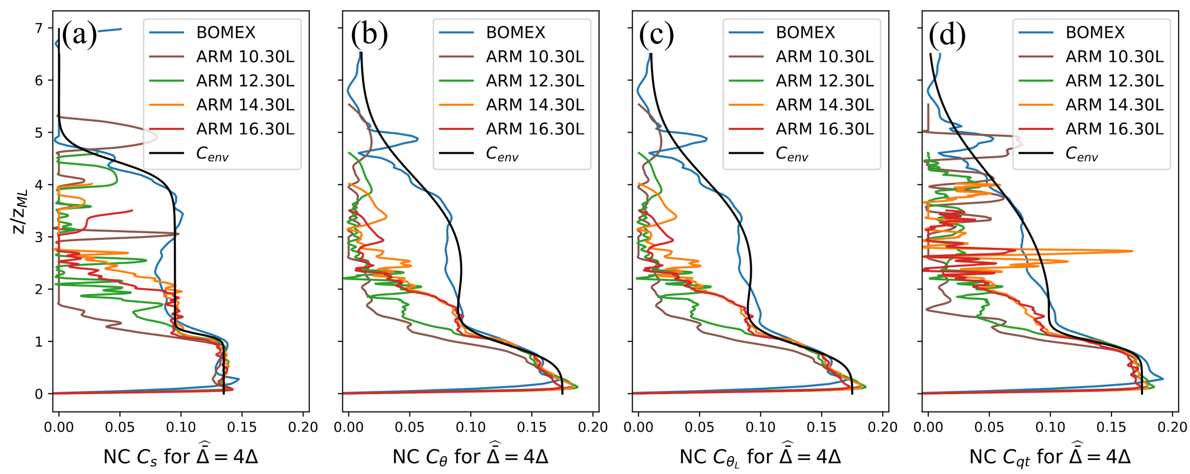

Figure 8.3: *Smagorinsky parameter profiles in the non-cloudy environment for each case in this study, scaled by the corresponding inversion height. The black line shows the parametrization found to account for the effect of height on the value of the Smagorinsky parameter in the non-cloudy environment.*

The C_{env} parametrization follows the BOMEX profile in this example, but the only difference between the BOMEX and ARM profiles here is the ratio of CL depth relative to the ML depth, as the profile has been scaled by z_{ML}. The behaviour of the C_{env} profiles at the CL top remains the same regardless of ML depth, and thus the same parametrization can be applied to both cases, ensuring that the case-specific CT height is accounted for. The other difference evident between the ARM and BOMEX cases in the CFE profiles is the level of noise present in the ARM case. This is a direct result of the averaging process being applied to each case. The BOMEX case data plotted in Figure 8.3 features 3 separate timestamps, 10 minutes apart during this QSS, being used to increase the number of points over which the horizontal averaging procedure is applied, resulting in smoother profiles. The ARM case however is an evolving case, which shows clear differences in fields which are just 10 minutes apart. Since the available data only includes timestamps spaced 10 minutes apart, the horizontal averaging process was only applied to fields from two different time stamps. Consequently, the noise observed in the ARM NC profiles results from the dissimilarity between the timestamps and the limited number of spatial points used in the averaging procedure. This indicates that the noise does not stem from a systematic source, and can therefore be disregarded.

This parametrization is only applied to the CFE. Another parametrization is required for the cloudy regions, as the Smagorinsky parameters in this regime exhibit substantially different behaviour compared to those in the ML. From Figure 8.2, it is evident that the IC values of C remain approximately constant with height. Therefore it was decided to set C_{IC} to a constant value, which was specific to each parameter. The momentum parameter was fixed to $C_{s,IC} = 0.15$. The scalar parameters show different behaviours, depending on the conservation property of the quantity being modelled. Figure 8.2 b shows the mid-CL regions, away from the effects of entrainment at the CB and CT, showing $C_{\theta_L} = 0$ IC. Meanwhile, the IC parameter values for the conserved scalars show very similar behaviour in Figures 8.2 c & d, and therefore the scalar parameters are set to $C_{\theta,\mathrm{IC}} = C_{q_t,\mathrm{IC}} = 0.2$.

8.2 Scale Dependencies of the Smagorinsky Parameters

The parametrization for C described previously in Section 8.1.1 effectively captures the height dependencies. However the scale dependencies are not yet accounted for in this parametrization, as the C_{ML}, C_{CFE}, and C_{IC} are fixed to values based on the $\widehat{\widehat{\Delta}} = 4\Delta$ data. Note that C_{FT} does not vary with scale as it is consistently set to zero. The effect of grid spacing on the C parameters in the other regions can be included in the parametrization by making C_{ML}, C_{CFE}, and C_{IC} functions of the grid spacing, Δ. The resulting parametrization function following on from Equation 8.7 now reads:

$$C_{\mathrm{env}}(\Delta) = \mathrm{w}_1 C_{\mathrm{ML}}(\Delta) + [\mathrm{w}_2 - \mathrm{w}_1] C_{\mathrm{CFE}}(\Delta) + [1 - \mathrm{w}_1 - \mathrm{w}_2] C_{\mathrm{FT}}(\Delta) \tag{8.10}$$

The trends that the C parameter value follows for each case in the ML, CFE, and NC flow regimes are presented in Figures 8.4, 8.5, 8.6 respectively. The ML data has been scaled by the ML depth, while the CFE and IC data have been scaled by the height of the CT. The heights at which these occur for each case are listed in Table 8.1.

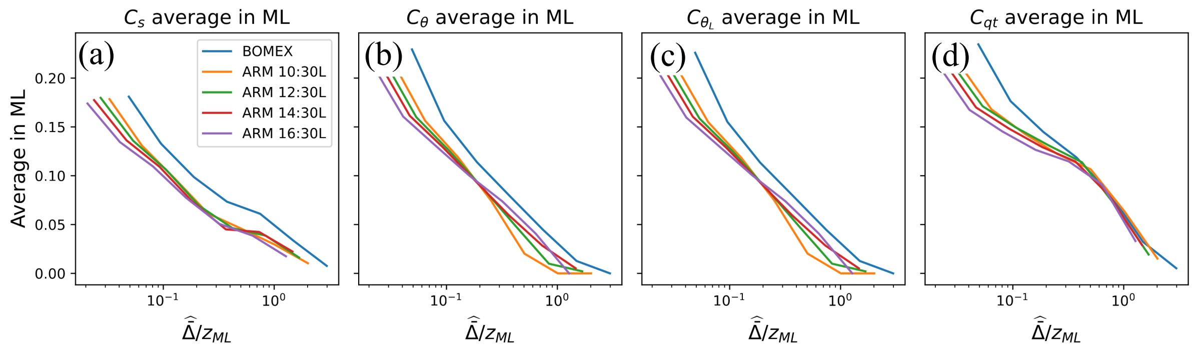

Figure 8.4: *Mean values of the planar averaged Smagorinsky parameter in the mixed layer for the BOMEX and ARM cases, plotted against filter scale* $\widehat{\bar{\Delta}}$ *scaled by the mixed layer inversion height* z_{ML}.

For momentum and heat (both θ and θ_L), Figures 8.4 a, b & c show a log-linear relationship between the Smagorinsky parameter values for these variables and the filter scale. However the C_θ and C_{θ_L} parameters show larger values at high resolution, and decrease to zero more rapidly as the data follows a steeper logarithmic slope than the C_s values. Meanwhile the C_{q_t} values show a similar decrease as C_s at high resolution, however a crest in C_{q_t} values exists at moderate-to-coarse resolutions. This suggests that the grey-zone has a less severe impact on the moisture structures at these scales, and consequently, C_{q_t} declines less rapidly at moderate scales in comparison to the other parameters. This suggests that the C_{ML} values for each variable should be set separately when using Equation 8.10 to parametrize the Smagorinsky parameter behaviour at various scales.

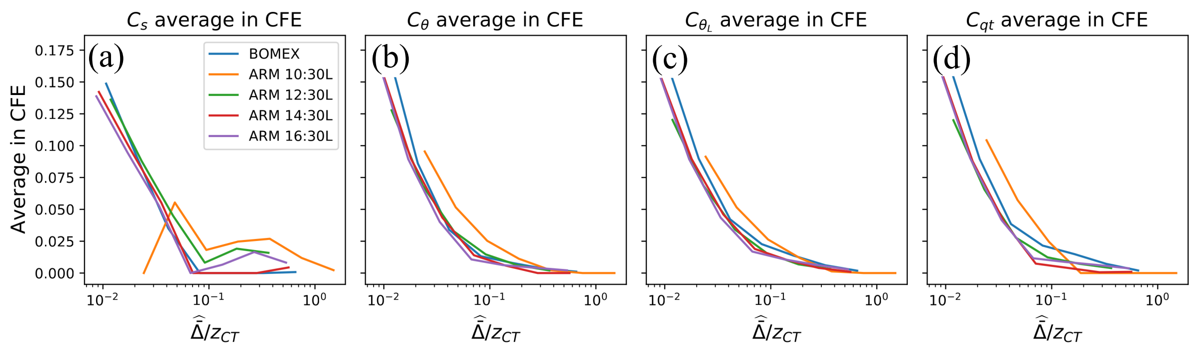

Figure 8.5: *Same as Figure 8.4 but for the* C *values in the CFE, with the effective grid spacing scaled by cloud top height* z_{CT}, *specific to each case.*

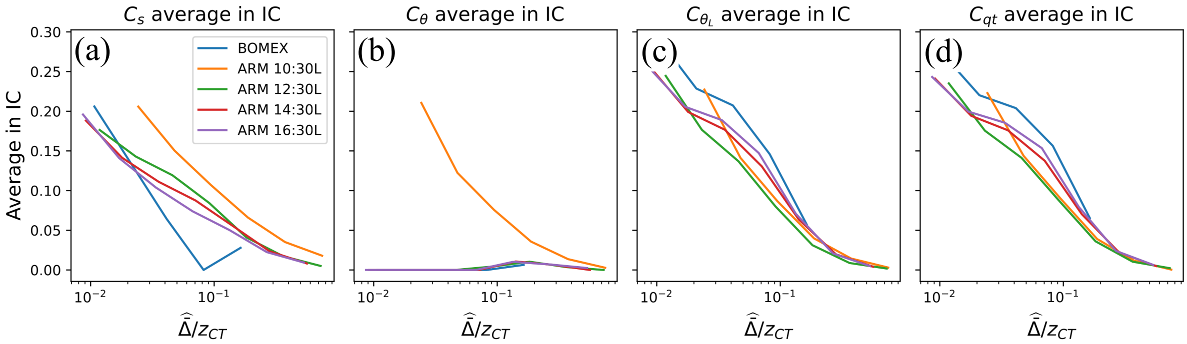

Figure 8.6: *Same as Figure 8.4 but for the C values in the IC, with the effective grid spacing scaled by cloud top height z_{CT}, specific to each case*

Figures 8.5 b & c portray the CFE C values for the thermodynamic parameters experiencing a linear decrease as filter scale increases. The momentum parameters however, show a logarithmic decrease in CFE C_s with $\widehat{\bar{\Delta}}$, more rapid than that observed in the ML. The C_s values in all cases reach zero at higher resolutions than in the ML also. Recall that the highest resolution (2Δ) C_s data shows a zero value due to contamination from the 8Δ data in the scale-dependent calculation. The moisture parameter is shown in Figure 8.5 d to decrease exponentially to zero in the later stages of ARM cases, similar to what is observed in the C_s trends in the CFE. However the BOMEX and earlier stages of the ARM case depict a more linear decrease to zero, similar to the C_θ trends in the CFE. This highlights not only a scalar dependency, but also a case dependency for the Smagorinsky parameter value trends with filter scale within the CFE.

It is in the cloudy regions where the differences between Smagorinsky parameter trends are most clearly seen. Figure 8.6 a shows an exponential decrease in the C_s parameter with increasing scale, though to different extents depending on the case. This illustrates the case dependency of the momentum parameter values when IC, suggesting that the C_s parametrization might benefit from another dependency in this region. This would allow the stability effects, which are the main differences between the cases, to be captured in the parametrization. The C_θ parameter shows all but the 10:30L ARM case showing zero values at all filter scales in-cloud, as expected. The behaviour of the C_{θ_L} and C_{q_t} parameters is near-identical IC. This is in contrast to in the ML, where these parameters showed clear differences in trends, and in the CFE where differences were present but less pronounced. The C_{θ_L} and C_{q_t} parameters show a decrease in value as filter scale increases, though at the moderate filter scales a clear crest is observed, particularly in the BOMEX case and in the ARM at the last two times analysed during the diurnal cycle. This indicates that once the CL is fully developed, the conserved moisture and heat structures in-cloud do not succumb to grey-zone effects as rapidly as in other flow regimes. This highlights the need for the IC parameters to have distinct scaling relationships for each variable being mixed and diffused, independent from parameters in other regions of the CTBL.

8.3 Set-Up of Experiments to Test the New Smagorinsky Parametrization

To test the effects of the parametrizations derived in the previous sections, the MONC model was altered to include the functional relationships. Multiple different configurations were tested during this phase of the investigation. The BOMEX cases, denoted by the subscript "B", were run at two different grid spacings $\Delta_{\mathrm{B}} = 160\,\mathrm{m}$ and $\Delta_{\mathrm{B}} = 320\,\mathrm{m}$ These grid spacings correspond to the filtered BOMEX data at $\widehat{\overline{\Delta}} = 8\Delta$ and $\widehat{\overline{\Delta}} = 16\Delta$ respectively. The ARM cases, denoted by the subscript "A" were run at two different grid spacings for each model configuration: $\Delta_{\mathrm{A}} = 200\,\mathrm{m}$ and $\Delta_{\mathrm{A}} = 400\,\mathrm{m}$, which again correspond to the filtered ARM data at $\widehat{\overline{\Delta}} = 8\Delta$ and $\widehat{\overline{\Delta}} = 16\Delta$ respectively. Results presented in Chapter 7 indicate that a grid spacing of between $\Delta = 160\,\mathrm{m}$ and $200\,\mathrm{m}$ is at the scale where grey-zone impacts begin to take effect. Meanwhile, a grid spacing from $\Delta = 320\,\mathrm{m}$ to $400\,\mathrm{m}$ is well within the traditional grey-zone regime. Some parametrization configurations were also run using the ARM case and a $\Delta_{\mathrm{A}} = 100\,\mathrm{m}$ grid spacing. This is because the scale invariant values presented in Table 8.2 for the C values in each regime were based on the highest resolution filtered data available, which correspond to $\Delta_{\mathrm{A}} = 100\,\mathrm{m}$ for the ARM case. Each combination of modifications to the Smagorinsky method is listed in Table 8.3 below, and further details are given in the following sections.

Name	Δ_{B}, Δ_{A}	Description
Smag 0.23	80 m, 100 m 160 m, 200 m 320 m, 400 m	$C_s = 0.23$, Pr$= 0.7$ throughout entire CTBL flow, as per the convention when running LES of CTBL in MONC. Scale invariant.
Smag 0.137	160 m, 200 m 320 m, 400 m	$C_s = 0.137$, Pr$= 0.7$ throughout entire CTBL flow, using the value determined by the dynamic method at high resolution in the ML. Scale invariant.
Smag 0.11 Smag 0.075	160 m, 200 m 320 m, 400 m	C_s is set according to the values determined by the dynamic method in the ML for the particular grid spacing, Pr$= 0.5$. scale-adaptive.
C_s prof	80 m, 100 m* 160 m, 200 m 320 m, 400 m	$C_s = C_{\mathrm{env}}$ in the CFE, $C_s = C_{\mathrm{IC}}$ IC. Pr$= 0.7$ used to determine the scalar parameters. The C values in each flow regime are set according to the high resolution dynamic data listed in Table 8.2. Scale invariant.
$C_s C_{\theta_L}$ prof	160 m, 200 m 320 m, 400 m	Same as C_s prof, however rather than determining the scalar parameter C_ψ using a Prandtl number, a separate profile is calculated for C_ψ, based upon the high resolution θ_L data in Table 8.2. Scale invariant.
SA C_s prof	80 m, 100 m* 160 m, 200 m 320 m, 400 m	Same as C_s prof, however the values of C_s in each flow regime now uses scale dependent values, depending on Δ, as listed in Table 8.4. Scalar parameters are still determined using Pr$= 0.7$. scale-adaptive.
SA $C_s C_{\theta_L}$ prof	160 m, 200 m 320 m, 400 m	Same as $C_s C_{\theta_L}$ prof, however the values of C_s and C_ψ are determined based on the grid spacing in use, according to the values listed in Table 8.4. scale-adaptive.

Table 8.3: *Summary of the various parametrization set-ups used in this study. Note that the configurations where the Δ is marked with an asterisk symbol * are equivalent set-ups.*

8.3.1 The Control: The Standard Smagorinsky Model

The standard Smagorinsky model, with a fixed C_s and Pr, was used to produce simulations of the BOMEX case and ARM case with grid spacings of $\Delta_{\text{B}} = 80\,\text{m}, 160\,\text{m}\,\&\,320\,\text{m}$ and $\Delta_{\text{A}} = 100\,\text{m}, 200\,\text{m}\,\&\,400\,\text{m}$ respectively. In this model, the Prandtl number was fixed to Pr$=0.7$, and this defined the ratio between momentum dissipation and the diffusion of all scalars, not just heat diffusion. The Smagorinsky parameter was set to $C_s = 0.23$ for both of these resolutions, following the convention previously applied in the high-resolution LES runs that were used to generate the filtered data. These MONC simulations are denoted as the "Smag 0.23" model runs. In order to test the effect of the C_{env} profile in the altered Smagorinsky model (see Equation 8.7), the standard Smagorinsky model was run with $C_s = 0.137$, for both BOMEX and ARM using the same three grid spacings as before. The C_s parameter was fixed to this value as it is the value assumed by $C_{s,\text{ML}}$ in the altered model, as listed in Table 8.2. This model set up is labelled as the "Smag 0.137" simulation.

Constant-Smagorinsky simulations were also performed to facilitate comparisons with the "scale adaptive" simulations. Suitable comparisons were obtained by using the standard model to produce simulations by fixing C_s to the same value as in the "scale-adaptive" altered model. The resulting fields can be used as controls to compare the altered model which used the scale-adaptive C_s profile parametrizations. LES of BOMEX and ARM were produced, but the C_s parameter was fixed to the "scale dependent" values that occur in the ML, while the Prandtl number remained as Pr$=0.7$. The Smagorinsky parameter in the $\Delta_{\text{B}} = 160\,\text{m}$ and $\Delta_{\text{A}} = 200\,\text{m}$ simulations was set to $C_s = 0.11$, and as such was termed the "Smag 0.11" model set-up. Similarly, the simulation with a grid spacing of $\Delta_{\text{B}} = 320\,\text{m}$ and $\Delta_{\text{A}} = 400\,\text{m}$ set $C_s = 0.075$ and was labelled the "Smag 0.075" set-up.

8.3.2 Implementing the C_s Profile into the Parametrization

The Smagorinsky scheme was altered to include a scale-invariant parametrization for the C_s parameter. The flow was first partitioned into IC and NC areas. The cloudy areas had the Smagorinsky parameter for momentum fixed to $C_s = 0.15$, following Figure 8.2. Meanwhile, the C_{env} function applied to determine the C_s parameter value in the NC environment (see Equation 8.7). The variables in this equation were set to the values found using the high-resolution data for C_s, as listed in Table 8.2. The values of the Smagorinsky parameters for scalars were found by computing the ratio of Prandtl number to C_s at each height level, with a dependency on if the point is IC or in the CFE. The Prandtl number remains fixed to Pr$=0.7$ in this model set up. This configuration was used to produce LES of BOMEX and ARM with the same grid spacings as in the control (see Section 8.3.1). This version of the altered Smagorinsky model, with the scale invariant C_s parametrization and fixed value Pr, is referred to as the "C_s prof" MONC set-up.

8.3.3 A Parametrization with Independent Profiles for C_s and C_ψ

The MONC model was then further altered to include a scale-invariant profile for the Smagorinsky parameter for scalars that is completely independent from that for momentum. The MONC model set up is very similar to that which is described in Section 8.3.2, with a $C_s = 0.15$ IC and the C_{env} parametrization used to determine a profile for C_s in the NC areas. However, rather than using a fixed Prandtl number to determine the Smagorinsky parameter for scalars, a separate profile is derived for C_ψ, for all scalars ψ. As it is difficult to separate the diffusion of each parameter in the MONC model, the same parametrization is applied to all scalars. Equation 8.7 is again used to determine the C_ψ profile for the NC areas, but the variables in the function are now set according to the θ_L specific values from high resolution data, as listed in Table 8.2. Meanwhile, the cloudy areas have their Smagorinsky parameter fixed to $C_\psi = 0.2$, based on high resolution data presented in Figure 8.2 for C_{θ_L}. The θ_L variable was chosen as it is similar to C_{q_t} in both the IC areas and CFE, while it is nearly equivalent to C_θ in the ML and CFE. Furthermore, theory suggests that it should be valid to use diffusion based on C_{θ_L} for the θ variables as:

$$\nu_h \nabla \theta_L = \nu_h \left(\nabla \theta - \frac{L}{c_p \Pi} \nabla q_l \right) \tag{8.11}$$

where ν_h is the heat diffusivity, L is the latent heat of vaporization, c_p is the specific heat of air at constant pressure, and Π is the Exner function defined as $\Pi = T/\theta$ with T being the temperature. However, it should be noted that analysis of the data from the dynamic model has clearly exhibited that $C_\theta \rightarrow 0$ IC. However, $C_{\text{IC}} = 0$ was not used in the parametrization as it would force the mixing and dissipation of moisture to zero IC also, as there is not currently a way of isolating this behaviour to just the θ variable in MONC. This scale-invariant model configuration is referred to as the "$C_s C_{\theta_L}$ profs" set-up. This specific parametrization was used to produce LES of both the BOMEX and ARM case studies at the same grid spacings as the control (see section 8.3.1).

8.3.4 Incorporating Scale Dependence into the New Smagorinsky Parametrization

Note that the values of α and γ for both the f and g functions remain set to the same values as used in the previous parametrizations for C_{env}, with the exact values detailed in 8.2. The scale-dependency affects only the values of the C_{ML}, C_{CFE}, and C_{IC} in this parametrization (see Equation 8.10). There are two different scale-adaptive parametrizations which are tested in this study. The first is similar to the C_s prof parametrization which was described in Section 8.3.2. In the scale-adaptive model configurations, the flow is decomposed into the IC and NC regions, however the Smagorinsky parameters which are assigned to these areas are allowed to vary not only with height, but also with grid spacing. The values taken by the C_s parameter in each of the flow regimes at various scales are detailed in Table 8.4. This model configuration sets the value of the Smagorinsky parameter for scalars according to the ratio of Pr$=\ 0.7$ to C_s at each point. This model configuration is labelled the "SA C_s prof" set-up, where SA stands for Scale Aware. LES of both the BOMEX and ARM case studies were produced using this version of the altered Smagorinsky scheme. Grid spacings of $\Delta_{\text{B}}\ = 160\,\text{m}\,\&\,320\,\text{m}$ and $\Delta_{\text{A}}\ = 200\,\text{m}\,\&\,400\,\text{m}$ were used, with the values of $C_{s,\text{ML}}$, $C_{s,\text{CFE}}$, and $C_{s,\text{IC}}$ varying according to scale.

Following on from this, an additional function can be introduced into the previously described model setup to ensure that the scalar parameters are entirely independent from the momentum parameter. As before in the $C_sC_{\theta_L}$ profs configuration, a separate relation between C_ψ and height/flow regime can be included in the Smagorinsky scheme, based upon Equation 8.10 and using the C_{θ_L} specific values outlined in Table 8.4 for each grid scale of interest. A separate scale-adaptive IC C_ψ is also set according to the θ_L values in Table 8.4. The MONC model with a subgrid scheme which was altered to include this new parametrization scheme was used to produce LES of both the BOMEX and ARM case studies. The same grid spacings as before are used, and the values of C_{ML}, C_{CFE}, and C_{IC} were set based on each grid spacing for both the C_s and C_ψ parameters (for all scalars ψ, where the value of C_ψ is determined using the θ_L data in this study). This parametrization set-up is called the "SA $C_sC_{\theta_L}$ profs" configuration.

		$\Delta_{\text{BOMEX}} = 80\,\text{m}$ $\Delta_{\text{ARM}} = 100\,\text{m}$	$\Delta_{\text{BOMEX}} = 160\,\text{m}$ $\Delta_{\text{ARM}} = 200\,\text{m}$	$\Delta_{\text{BOMEX}} = 320\,\text{m}$ $\Delta_{\text{ARM}} = 400\,\text{m}$
C_s:	C_{ML}	0.137	0.11	0.075
	C_{CFE}	0.095	0.05	0.015
	C_{IC}	0.15	0.12	0.09
	C_{FT}	0.0	0.0	0.0
C_{θ_L}:	C_{ML}	0.175	0.13	0.08
	C_{CFE}	0.09	0.05	0.015
	C_{IC}	0.2	0.17	0.13
	C_{FT}	0.0	0.0	0.0
C_{q_t}:	C_{ML}	0.165	0.15	0.13
	C_{CFE}	0.1	0.05	0.01
	C_{IC}	0.2	0.165	0.14
	C_{FT}	0.0	0.0	0.0

Table 8.4: *Values taken by the various variables input to the C_{env} function in order to parametrize the Smagorinsky profiles for momentum $C_{s,\,env}$, heat $C_{\theta_L,\,env}$, and moisture $C_{qt,\,env}$ for a CTBL simulation.*

Table 8.4 includes values specific to the C_{q_t} parameter as the intention was to modify MONC to allow for scalars to be treated individually by the subgrid dissipation scheme. Although this could not be achieved in the remaining time, the specific values are listed here as a reference for any future work that might benefit from this information. The values relating to the C_θ parameter are exactly identical to the C_{θ_L} values in every flow regime apart from IC, where $C_{\theta,\mathrm{IC}} = 0$ for every Δ. The naming convention for each parametrization configuration, along with a brief description of the model has previously been summarised in Table 8.3.

8.4 Analysing the Effect of the New Parametrization on MONC Simulations

The outputs from each of the LES using different parametrization configurations are now compared and contrasted in the following sections. The ARM data is focused on in this section for brevity, as the BOMEX data was observed to exhibit similar results to the ARM case when the CTBL is in QSS at 16:30L. The ARM case also has the added benefit of evolving through time, allowing for the assessment of the parametrizations designed to mitigate grey-zone effects, particularly during transitional periods when these effects are most pronounced.

8.4.1 Cloud Layer Growth

The first step is to investigate the impact of the different parametrization configurations on the developing cloud field in the ARM case. The results of using the C_s prof set-up are analysed first, and compared to the output from the Smag 0.23 model configuration, which is used as a control. The outcome of this comparison is presented in Figure 8.7. Note that high-resolution LES data of the ARM cloud layer growth produced using the Standard Smagorinsky model with a grid spacing of $\Delta = 25\,\mathrm{m}$ was also plotted for comparison, and is used as the "truth" for this analysis.

Including the C_s prof parametrization in the model, rather than setting C_s to a fixed value of 0.23, yields better results for the initiation and growth of the cloud layer, as observed in Figure 8.7. The C_s prof data with $\Delta = 100\,\mathrm{m}$ can be seen initiating at the same time as the LES at $\Delta = 25\,\mathrm{m}$. However, the CT is observed growing to higher elevations past 11:00L. The Smag 0.23 configuration with $\Delta = 100\,\mathrm{m}$ initiates cloud formation approximately 20 minutes later than the $\Delta = 25\,\mathrm{m}$ LES reference. However, the CT height remains relatively accurate until after 12:00L, at which point the model begins to predict the CT at altitudes higher than those observed in the LES truth. That being said, a similar cloud layer development as the reference simulation can be achieved at four times the grid spacing if the C_s prof parametrization is employed in the MONC model. Further to this, the cloud layer growth pattern from the C_s prof configuration at $\Delta = 200\,\text{m}$ is of similar quality to the $\Delta = 100\,\text{m}$ Smag 0.23 control run. The $\Delta = 400\,\text{m}$ C_s prof data shows similarities with the $\Delta = 200\,\text{m}$ Smag 0.23 data, though this resemblance is less pronounced compared to the previously discussed comparisons. That being said the C_s prof with $\Delta = 400\,\text{m}$ initiates cloud formation an hour and a half earlier than the Smag 0.23 data at the same resolution. Additionally, the $\Delta = 400\,\text{m}$ C_s prof output also reaches the correct CT height around 12:30L, whereas the equivalent Smag 0.23 cloud field never attains the correct CT height. This indicates that while the C_s prof parametrization significantly improves the timing of cloud development and CL depth, it remains affected by grey-zone dynamics.

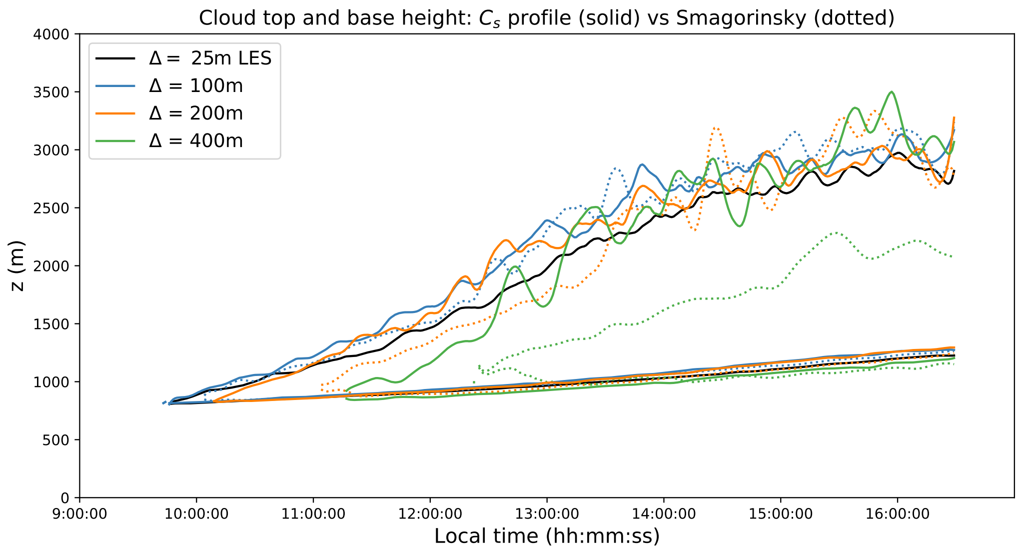

Figure 8.7: *Time series of the development of the ARM CL for simulations at 3 different grid spacings produced using two different MONC model set-ups: (1) the altered model with the C_s prof parametrization (solid lines), and (2) the standard Smagorinsky model with a fixed $C_s = 0.23$ (dotted lines).*

The addition of scale dependency to the parametrization is observed to greatly improve the cloud initiation time at coarse resolution. Figure 8.8 shows the output from the $\Delta = 400\,\text{m}$ SA C_s prof parametrization (dashed green line) with cloud initiating over an hour earlier than the scale-invariant equivalent (solid green line). The parametrization set-up also produces a CT that most closely follows the true CT height as determined by the $\Delta = 25\,\text{m}$ LES. The $\Delta = 400\,\text{m}$ SA C_s prof is very similar to both the $\Delta = 200\,\text{m}$ cloud layers outlined in this plot, with $\Delta = 400\,\text{m}$ SA C_s prof data showing cloud initiating just 20 mins after the $\Delta = 200\,\text{m}$ data. The similarity between both the parametrization set-ups at $\Delta = 200\,\text{m}$ indicates that not as much benefit can be achieved from scale dependency alone at higher resolutions. Recall that the C_s prof is the same as SA C_s prof at $\Delta = 100\,\text{m}$ as the C values in each regime are based partly on the $\Delta = 100\,\text{m}$ ARM data.

It is clear that the C_s prof parametrization set-up improves the ARM simulations at each grid spacing investigated. However, the C_s prof configuration does not account for variations in scale experienced by the Smagorinsky parameter. Therefore the SA C_s prof parametrization is used to test the effect of allowing the C_s parameter to vary based on scale. The resulting cloud growth time series is presented in Figure 8.8.

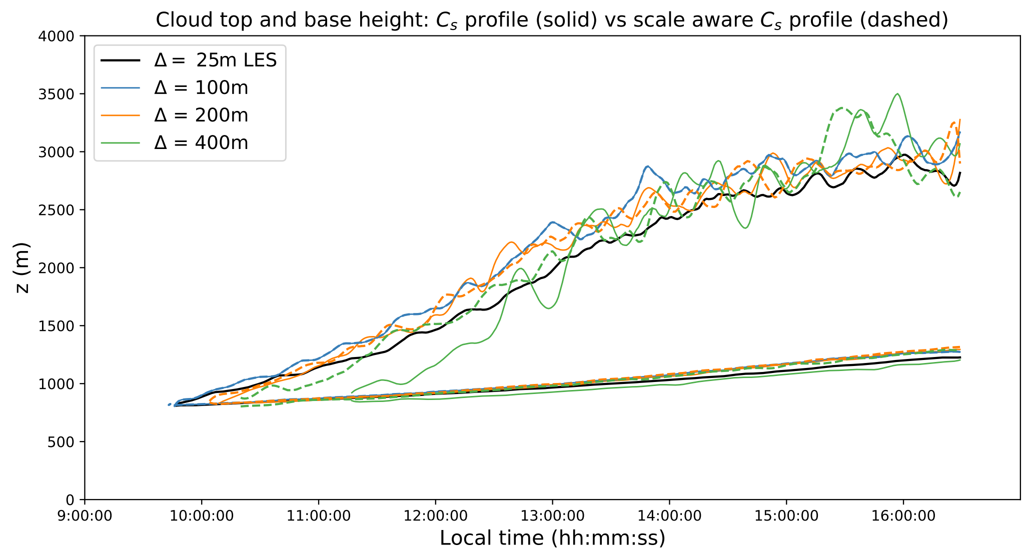

Figure 8.8: *Similar to Figure 8.7, however the two parametrization configurations being compared are as follows: (1) the C_s prof parametrization, shown by the solid line, and (2) the SA C_s prof, with a scale-adaptive C_s profile, shown by the dotted line.*

To investigate the effect that grid spacing Δ has on a given parametrization's ability to predict the cloud initiation time, as well as CT and CB heights, the time series is plotted for each of the parametrization configurations after being applied to the ARM case with grid spacings of $\Delta = 200$ m and $\Delta = 400$ m in Figures 8.9 a & b respectively.

In Figure 8.9 a, the $\Delta = 200$ m data suggests that any subgrid scheme which accounts for scale and/or the variation of C with height/flow regime, is much better able to produce accurate cloud initiation times along with more reliable CL depths. This is evident from the clustering of the vertical lines just after 10:00L, illustrating that subgrid schemes with these considerations can initialise cloud earlier than subgrid schemes without these considerations (the Smag 0.23 set-up, for example). The cloud fields resulting from these parametrization configurations when $\Delta = 200$ m also produce accurate CT heights from the moment of the cloud fields inception. Only Smag 0.23 does not trigger cloud development at 10:00, instead taking until after 11:00L to initiate cloud. The CL depth is not as accurate with this set-up either; the CT does not reach the correct height until after 13:30L in the Smag 0.23 set-up at $\Delta = 200$ m.

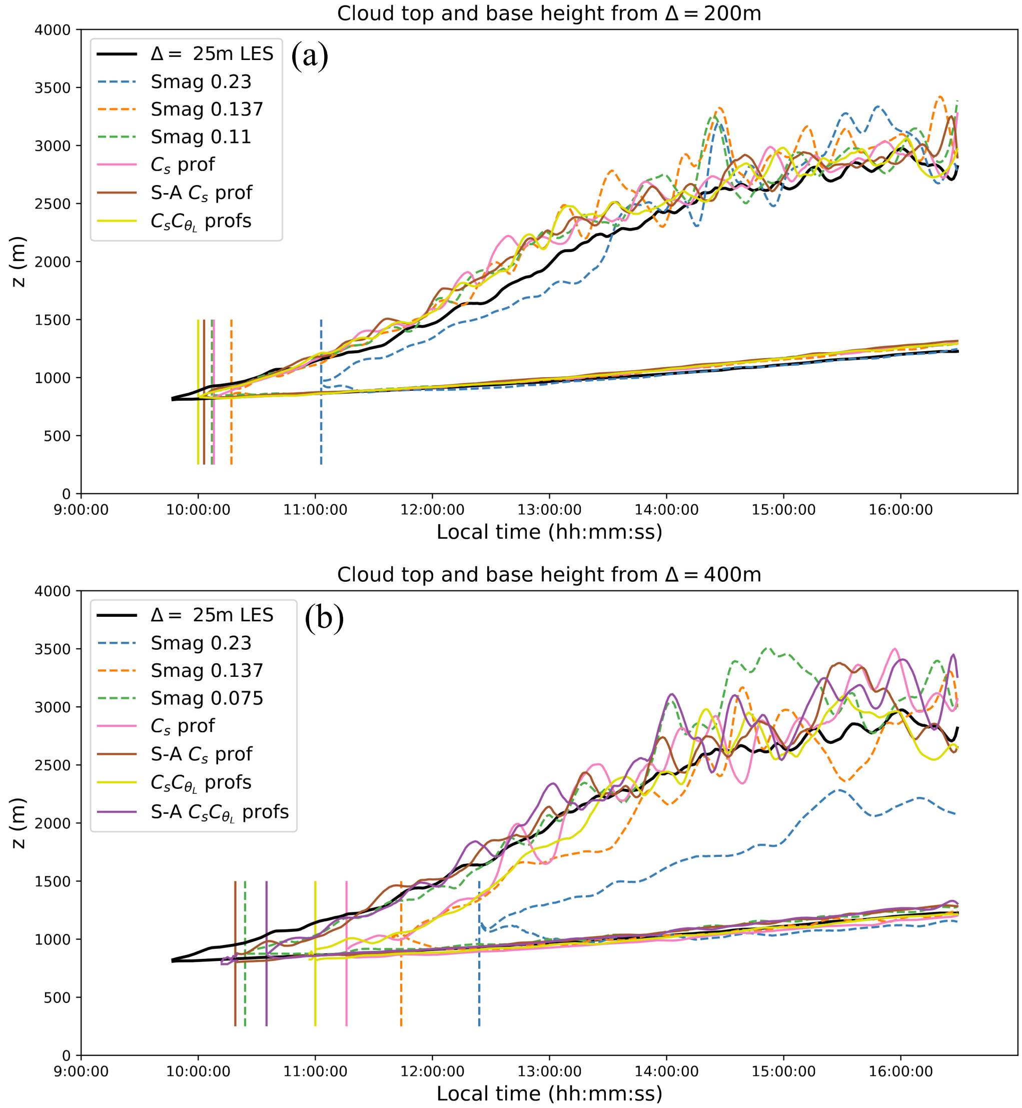

Figure 8.9: *Time series of the CT and CB heights simulated by the MONC model with various different parametrization set-ups for the subgrid scheme, as listed in the legend and described in Table 8.3. The vertical lines show the time at which a continuously growing CL is initiated.*

The time series of the CT and CB for the various parametrization configurations are now grouped into similar set-ups and plotted according to these categories. It is clear from Figure 8.10 a that a subgrid scheme which uses a profile parametrization to determine C_s (the C_s prof set-up), performs better than a subgrid scheme which fixes C_s to a constant (Smag 0.137) when both schemes are scale invariant and have a grid spacing of $\Delta = 400\,$m. This is in spite of Smag 0.137 using the same C_s value as the C_s prof set-up applies to the ML. The C_s prof configuration yields a cloud field which develops earlier, around 11:15L compared to Smag 0.137 which begins to develop at 11:45L. The C_s prof subgrid scheme also produces clouds with more accurate CT heights earlier on in the day than the Smag 0.137 configuration. However the growth of this cloud field is still poor compared to other parametrization set-ups with $\Delta = 400\,$m, as presented in Figure 8.9 b. From this plot, it is clear that accounting for scale dependency in the model greatly improves the resolved cloud field's depth formation.

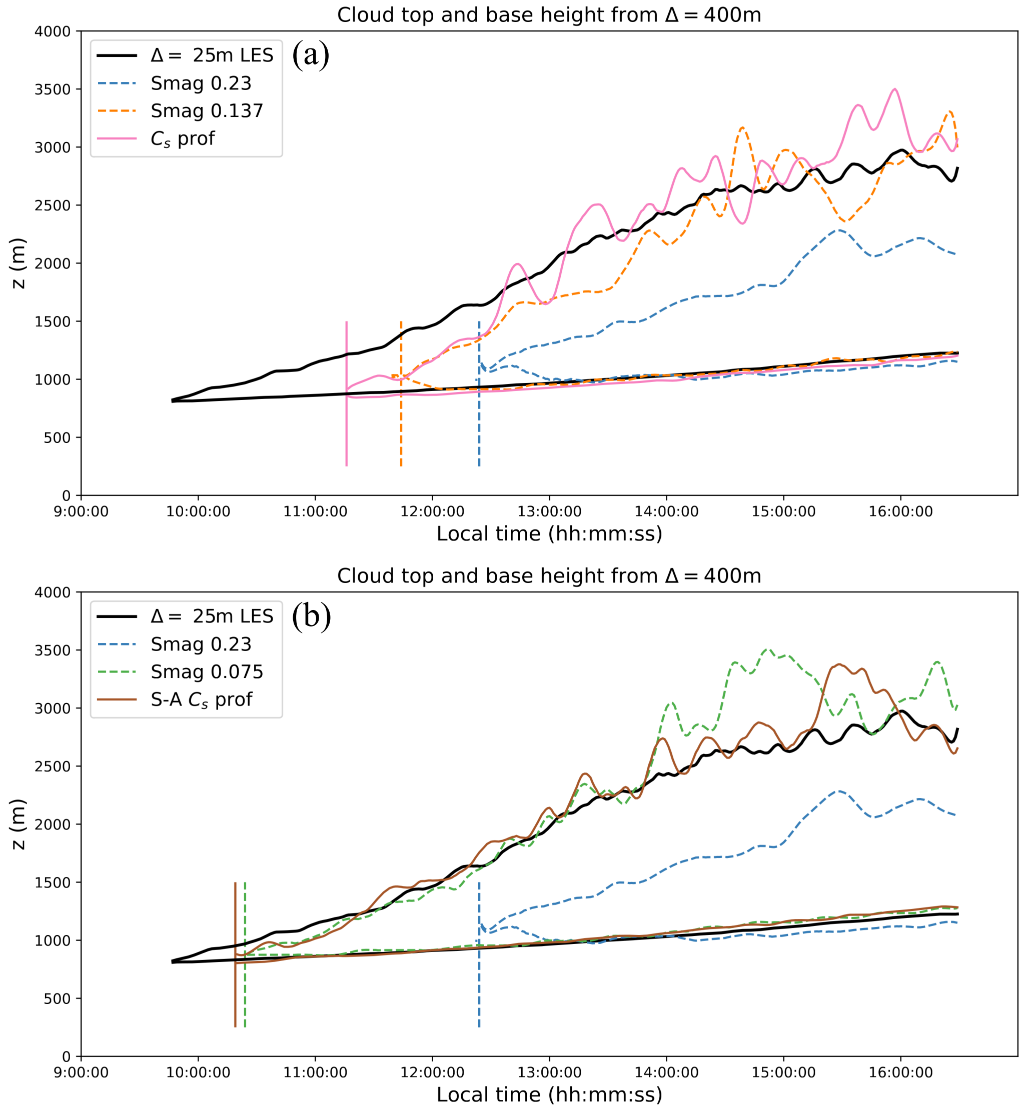

Figure 8.10: *Similar to Figure 8.9 b but with results from certain parameterizations only. The top figure (a) shows the C_s prof data, along with the Smag 0.137 data, which is equivalent to C_s prof, just without a profile function included, as it fixes C_s to the same value as in the ML of the C_s prof set-up. The cloud evolution in the true LES data and Smag 0.23 data are also plotted for comparison. The bottom figure (b) shows the scale-adaptive versions of the parametrizations which were presented in (a), again with the LES data and Smag 0.23 data being plotted as references.*

Figure 8.10 b presents the effect of accounting for the scale dependencies which are evident in the C_s parameter when parametrizing their values in the subgrid scheme. The findings depicted in this figure confirm that scale dependence greatly improves both CL depth and the time of initial cloud formation. There is very little difference in the timing of the first clouds developing between the Smag 0.075 and SA C_s subgrid configurations. This demonstrates that, when the appropriate C value is applied relative to the model grid scale, the timing of cloud initiation will be accurate, irrespective of whether a C profile parametrization is incorporated into the subgrid scheme. However, as observed in Figure 8.10 a, if a C value is used which is incompatible with the scale of the grid spacing, then the C profile parametrization can improve cloud initiation time greatly. The use of a parametrized profile for C_s can be seen to benefit the model later on in the diurnal cycle. In Figure 8.10 b it becomes apparent that the use of the C_{IC} and C_{env} functions enables the CT height to be accurately resolved throughout the majority of the diurnal cycle. Meanwhile, the Smag 0.075 set-up overestimates the height at which the CTs reach later on in the day. Therefore, the scale dependence of C_s appears to be the most significant factor in the subgrid scheme to enable accurate LES of the cloud layer formation and development. However, introducing the regime-specific, height-based functions seems to positively impact the model's ability to determine the CT height more accurately, especially at later times during the diurnal cycle.

So far the effect of varying the C_s parameter values has been investigated, but the results presented in Figures 8.10 a & b force the C_ψ parameter to depend on C_s and Pr, for all scalars ψ. The next phase of this investigation will evaluate the effects of allowing the scalar parameter to vary independently from the C_s parameter in each flow region. The scale invariant and scale aware versions of the subgrid scheme parametrizations with (1) the C_s profile, and (2) the C_s & C_{θ_L} profiles are used to produce the cloud fields from the ARM case. The resulting evolution of these CT and CB heights with time is plotted in Figure 8.11. These simulations were produced using a grid spacing of $\Delta = 400\,\text{m}$ as this scale was found to show the most sensitivity to changes in parametrizations.

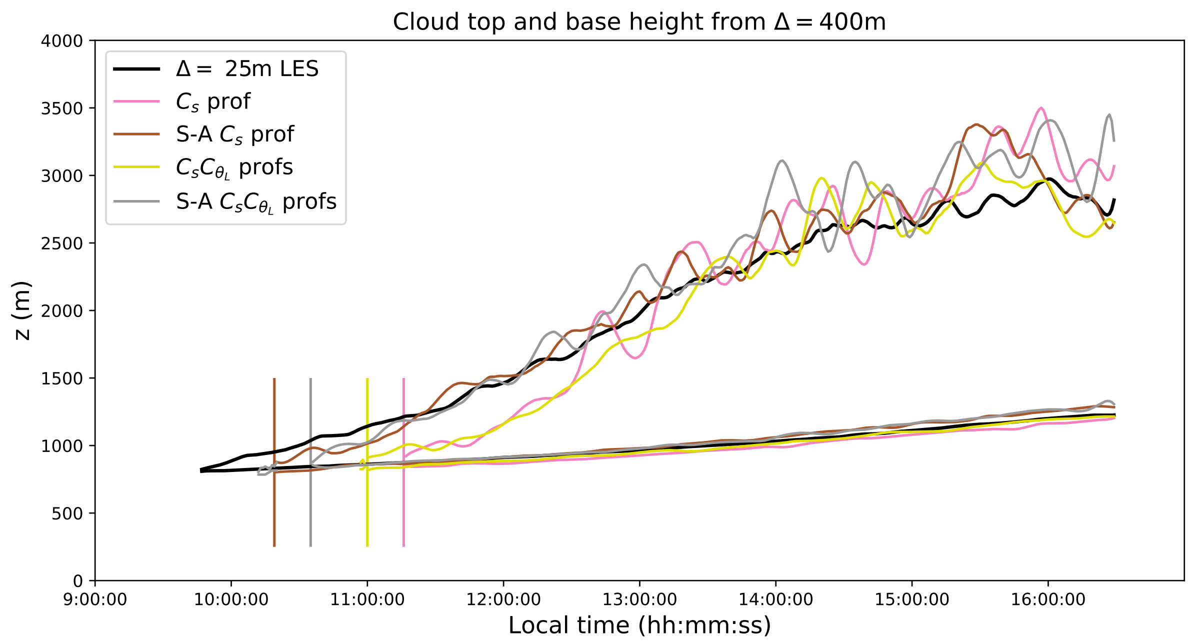

Figure 8.11: *Time series of the CT and CB heights for the ARM case produced using the MONC model with grid spacing $\Delta = 400\,m$ and various different subgrid parametrization schemes, as listed in the figure legend.*

Figure 8.11 highlights that ARM simulations produced with a grid spacing of $\Delta = 400\,\text{m}$ have the most accurate CT heights when using scale-adaptive parametrizations for C. The cloud formation times for the scale invariant parametrizations are later in the day: The $C_sC_{\theta_L}$ prof parametrization begins to form cloud at 11:00L, while C_s prof parametrization triggers cloud development from 11:15L. However the C_s prof set-up begins to resolve the CT height accurately from between 12:30L and 13:00L, meanwhile the $C_sC_{\theta_L}$ prof configuration does not accurately determine the CT height until after 13:30L. Both the C_s prof and $C_sC_{\theta_L}$ profs set-ups struggle to reach the correct CT height, with CT level being underestimated for the first half of the day in these simulations when $\Delta = 400\,\text{m}$. This is because the C parameters are scale-invariant in these configurations, and thus are not being set to the best-fit C values relative to the coarse grid scale. The late stage at which the $C_sC_{\theta_L}$ prof parametrization reaches the correct CT height suggests that the C_{θ_L} profile seems to hinder the height that the CTs can reach. This may be a result of setting the C_{IC} value to $C_{\theta L} = 0.2$ rather than $C_\theta = 0$, and therefore causing an excessive amount of heat diffusion to occur.

However, the additional independent C_{θ_L} profile in the subgrid parametrization appears to help with the initial formation of cloud in the $C_sC_{\theta_L}$ prof case when the set-up is scale invariant. The same cannot be said of the SA $C_sC_{\theta_L}$ prof, which shows cloud developing and persisting after 10:30L, whereas the SA C_s prof set-up shows cloud developing earlier, before 10:15L. This suggests that the C parameter values might require more fine-tuning. However, a small, very shallow layer of cloud is observed forming and dissipating over the course of approximately 5 mins in SA $C_sC_{\Theta_L}$ prof model (outlined in grey), evident just before the brown vertical line at about 10:15L in Figure 8.11. This further suggests that the dissipation IC, or at least in the CL, might be too high in the $C_sC_{\theta_L}$ prof set-ups, both for the scale-invariant and scale-adaptive configurations. It is hypothesized that if C_ψ was based on the θ data rather than the θ_L data, then the first parcel of cloudy air in the that $C_sC_{\theta_L}$ prof set-up might have persisted, or even have been formed earlier.

8.4.2 Mean Profiles of Cloud Cover

The mean profiles depict the total cloud fraction from both the scale-invariant and scale-adaptive set-ups of the C_s prof parametrizations. Analysing these profiles ensures that the identified CT and CB in Figures 8.7 - 8.11 correspond to a well-resolved cloud field, rather than being the outcome of a limited number of spurious cloud formations within an otherwise predominantly cloud-free layer. Cross sections of the cloud field from the modified MONC model were also examined, and while the results appeared reasonable, no obvious systematic differences in the size/structure of the largest clouds were observed. Therefore, the cross-sections are not presented here. The profiles from all analysed time stamps, apart from the very earliest time stamp where the cloud has not yet properly formed, exhibit reasonable cloud fraction profiles consistent with those observed in the LES reference. In regards to the difference between the C_s prof and SA C_s prof model configurations, the largest difference is observed when the grid spacing is $\Delta = 400\,\text{m}$. This is expected as the disparity between the scale-dependent C values and the values assigned to C in the scale-invariant set-up is exacerbated as grid scale increases.

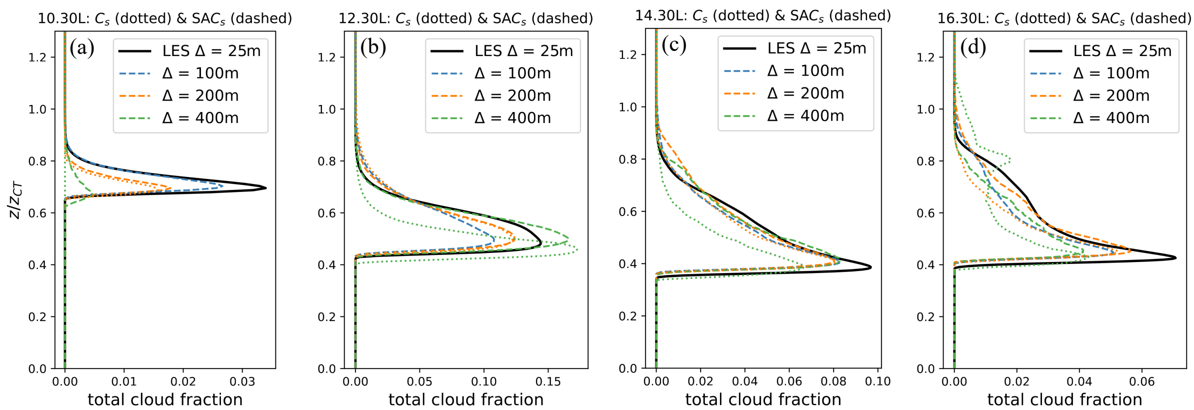

Figure 8.12: *Mean Profiles of the cloud fraction across the four time steps of interest in the ARM case. The dotted lines show results from the C_s prof set-up, while the dashed lines show results from the SA C_s set-up. The black line shows the "truth" from the high-resolution LES. The height axis has been scaled by the CT height.*

At the 12:30 timestamp during the early cloud development phase, Figure 8.12 b shows the coarse grid spacing LES exhibiting the highest percentage of cloud, while the fine grid spacing LES shows the lowest percentage. This could be attributed to the all-or-nothing cloud scheme used in MONC, where each grid cell is classified as either containing cloud or not, depending on if the liquid water content of the grid box exceeds a certain set threshold. This threshold is determined by the saturation vapour pressure in an all-or-nothing scheme. If a similar number of grid cells are classed as cloud for each resolution LES, the LES with larger grid spacing will have a larger area being classed as cloudy compared to the LES with smaller grid spacing. Since the domain size remains constant for each resolution, the result would be a higher cloud formation for coarse resolutions. This behaviour is likely to occur during the initial stages of cloud formation when clouds are still small-scale, and the smallest area that can be classed as a cloud is limited to the size of the grid box. However, as the day progresses and clouds grow larger than the grid spacing, the scheme should converge toward a more accurate cloud field, and therefore cloud fraction, across all resolutions.

The profiles at all other times exhibit the anticipated results for resolved cloud fraction profiles, in relation to the profile's corresponding grid spacing. Figure 8.12 a shows the highest resolution data ($\Delta = 100\,\text{m}$) resolves the cloud fraction most accurately, when the LES reference data is defined as the truth. The $\Delta = 100\,\text{m}$ data exhibit a high degree of similarity between the two parametrization configurations. However, the $\Delta = 400\,\mathrm{m}$ data shows the SA C_s prof configuration beginning to resolve cloud, while the C_s prof remains unable to resolve cloud at this early time in the diurnal cycle. Throughout the time stamps presented here, the C_s profile configuration consistently produces a cloud fraction profile furthest from the LES for all the parametrization set-ups analysed in Figure 8.12. All other configurations show the cloud fraction being reasonably captured compared to the reference data. This demonstrates that including a scale-dependent flow regime based profile parametrization for the C_s parameter can yield accurate simulations of a shallow cumulus cloud field, even when a coarser grid spacing, typically associated with the grey-zone regime, is used.

8.4.3 Computational Efficiency

The new parametrization for C allows the altered model to make use of a larger time step compared to the standard model, while also achieving a higher level of accuracy in the results. This is because the MONC model has adaptive time-stepping, which uses the CFL criteria to determine the largest possible timestep, dt, given the state of the flow. Therefore the lower levels of diffusion in the C_s prof and $C_sC_{\theta_L}$ prof parametrizations allow for a larger time step to be used. This is especially apparent in the scale-adaptive versions of the parametrization, which exhibited the fastest model runs and, when compared to the $\Delta = 25$ m LES reference data, generally produced the most accurate fields of all the parametrization configurations.

The CPU performance of the MONC model when using different subgrid scheme configurations can be investigated. The BOMEX case can be used as a test case for this analysis, by setting the grids to be uniform between the two configurations: a domain of $16\times16\,\mathrm{km}^2\times 3\,\mathrm{km}$, a horizontal grid spacing of $\Delta = 320\,\mathrm{m}$, and a vertical grid spacing of 10 m. The standard Smag 0.23 configuration requires approximately 56,009 time steps, with a processing time of 42 hours 37 mins to complete the LES run for 4 hours of simulated data. Meanwhile, the model with the SA $C_sC_{\theta_L}$ prof parametrization for the subgrid scheme produces the LES of BOMEX using 15 hours 25 mins of processing time, using 41,290 time steps to simulate the same case. The C_s prof parametrization behaves similarly to the $C_sC_{\theta_L}$ prof parametrization in terms of computational efficiency. This highlights the efficiency of the C_s prof and $C_sC_{\theta_L}$ prof parametrizations, as they allow the LES model to run faster and use less computational resources, while also producing simulations with more accurate cloud initiation and development.

8.5 Summary of Findings on the New C Parametrization for the Smagorinsky Scheme

A new parametrization has been developed to account for systematic variations in the Smagorinsky parameter values with height across different flow regimes. In the non-cloudy areas of the CTBL, this parametrization builds on the Blackadar mixing length formulation, allowing for a smooth transition between C values in the ML, CFE, and OL. Meanwhile, in the cloudy regimes, the C values are fixed to a constant. The new parametrization can be applied to C_s alone, forcing C_ψ to depend on C_s via a fixed Prandtl number Pr, or it can be applied to both C_s and C_ψ for any scalar variable, ψ. In addition to height and variable dependency, the scheme also incorporates grid scale dependency. This scale dependency was found to significantly improve model efficiency, while also improving the formation time and evolution of the simulated cloud fields. The greatest benefit arising from including regime-dependent and/or scale-adaptive C variables in the model is observed at coarse grid scales. Therefore, this parametrization is particularly useful for modulating the onset of the grey-zone in LES of shallow cumulus CTBL.

In conclusion, the newly introduced parametrization for C significantly enhances the performance of the MONC model by enabling the use of larger time steps compared to the standard model, while resolving the CTBL fields to a higher level of accuracy. The greatest improvement in accuracy was observed at coarse resolutions, where the new parametrization schemes enabled the model to more precisely capture the evolution of the diurnal cloud field. This was particularly evident when the scale-adaptive version of the parametrization was used. The new parametrization for C improves the MONC model by allowing larger time steps and greater accuracy. The scale-adaptive versions demonstrated the quickest run times, as well as the highest accuracy when compared to the $\Delta = 25$ m LES reference data. The C_s prof and $C_sC_{\theta_L}$ prof parametrizations significantly reduced both time steps and run time, enhancing computational efficiency while maintaining accurate cloud simulations. Therefore the new parametrizations which have been devised in this chapter enable the model to run quicker while demanding less computational resources and producing more accurate representations of cloud initiation and development at grey-zone resolutions.

8.5.1 Responses to the Research Questions

(1) What systematic relationships can be established between the Smagorinsky parameters and height within each flow regime of the CTBL? Can scale dependence be accounted for in these functional relationships?

By decomposing the CTBL into three main flow regimes, the ML, CFE, and IC regions a height-dependent relationship, based upon a modified Blackadar formula, has been established between the two cloud-free regimes (ML and CFE). Meanwhile, the Smagorinsky parameters in-cloud have their C value fixed to a constant. This parametrization can capture the key effects that height and stability have on the Smagorinsky parameter values in the CTBL.

Scale dependence can easily be added to this scheme by varying the values of C_{ML}, C_{CFE}, and C_{IC} within the defining equations of the parametrization. Accounting for variations in the C value with grid-scale results in marked improvements in the accuracy of the modelled CTBL fields.

(2) What effect does this empirical parametrization have on the modelled fields?

The parametrization which has been derived and tested throughout this chapter has been found to greatly improve the MONC model's ability to resolve cumulus clouds earlier on in the simulated diurnal cycle, improving the accuracy of the timing for the initial cloud formation. Furthermore, the subgrid schemes which included the profile-based parametrizations for C_s and/or C_ψ were found to improve the model's ability to accurately resolve the CT height. These improvements were most pronounced at coarser resolutions, indicating that this parametrization effectively delays the scale at which the grey-zone begins to impact cloud formation and development. Additionally, this parametrization resulted in a more than two-fold increase in the speed of the model during Convective Turbulent Boundary Layer (CTBL) simulations. This increase in computational efficiency was achieved without compromising accuracy; the modified models not only maintained accuracy, but instead demonstrated significant improvements. Therefore this parametrization has a profoundly positive impact on both the accuracy and computational efficiency of the model when producing LES of CTBLs at scales typically considered to be in the grey-zone regime.

9 Conclusions and Discussion

Weather and climate models have had significant improvements in their accuracy in the past few years, particularly through the use of higher-resolution grids that better capture motions in the atmosphere. However, challenges remain when the grid spacing of the model is of the same scale as the dominant coherent structures in the flow that is being modelled. This gives rise to the grey-zone regime, where these turbulent motions are only partially resolved. Effectively capturing the contribution that turbulence has to the overall flow in this regime is crucial for improving the reliability of severe weather forecasts and long-term climate projections. The work detailed in this thesis focuses on the grey-zone of the shallow cumulus topped CBL. However, it is expected that the principle of the main findings can be generalised and applied to other scenarios, such as the importance of including scale adaptivity and flow regime dependence in any subsequent grey-zone parametrizations.

LES has long since been used as a tool, alongside observations, to analyse the behaviour of mixing lengths within the CBL. However, the majority of previous studies focus on the physics and dynamics within either dry CBL's, or the ML of CTBLs. Throughout this thesis, it has been proven that the CL demonstrates a clear distinction in the behaviours of momentum, heat, and moisture fluxes compared to those within the ML. In addition to this, when investigating the key traits of the mixing lengths for each parameter in the CL, distinguishing between the IC and CFE was found to be important for the analysis. Past research in the field has also commonly assumed that heat and moisture fluxes within CTBLs are not only equal to one another, but can also be approximated by depending solely on the momentum flux and a constant Prandtl number. This approach was found to be insufficient for the CTBLs analysed throughout this study. As a result, a new parametrization scheme has been developed to address the current limitations in the existing Smagorinsky scheme. Despite the limitations encountered in this study, the proposed parametrization demonstrates substantial improvements in the model's ability to accurately resolve cloud fields at grey-zone scales, whilst also reducing computational costs. This strongly suggests that further advancements in LES models are possible through additional investigations into the behaviour of fluxes and diffusion within the CL.

9.1 Summary of the Key Findings

This thesis examined the behaviours of the Smagorinsky subgrid scheme, with a focus on the values taken by the Smagorinsky parameters for momentum, heat, and moisture in LES of three different CBL test cases. The overall aim was to identify systematic behaviours of the Smagorinsky parameter values in order to derive a parametrization which would offset the effects of the grey-zone at coarse resolutions. The Smagorinsky subgrid scheme was found to behave as a filter, with characteristics very similar to that of the Gaussian filter. This enabled the Gaussian filter to be employed, along with the offline dynamic Smagorinsky model to calculate fields of flow-dependent Smagorinsky parameter values specific to each variable of interest. These fields have been analysed and an explicit link has been established between the mixing length scale and flow regime within the CTBL, with a scale dependency also clearly evident in the parameter values. A new parametrization scheme was proposed to address limitations in the existing Smagorinsky model, focusing on variable-specific parameters, while accounting for the systematic dependencies observed across specific flow regimes and grid scales. The investigation that led to the identification of this parametrization was divided into four parts, each addressing different research questions and forming a foundation for the subsequent chapters. The major findings from this research are summarized according to chapter in the following sections.

9.1.1 The Smagorinsky Scheme as a Filter

The investigation into the behaviour of the Smagorinsky scheme in spectral space provides significant insights into the functionality of the Smagorinsky subgrid scheme, and establishing it as a grid-scale-based filter within LES models. Following this, the filter scale and filter shape in spectral space were established for the Smagorinsky scheme. This filter was found to maintain a distinct shape across various grid spacings, while also demonstrating a linear scaling with grid spacing. The Smagorinsky filter shape is observed between the scales corresponding to effective resolution R_{eff} and the Nyquist frequency. Between these scales, the Smagorinsky scheme dissipates energy, causing the data to follow a distinct slope. The value of this slope has been approximated by likening the Smagorinsky filter to that of a Gaussian.

The Smagorinsky filter was found to closely resemble the Gaussian filter in spectral space, particularly at high resolutions. This similarity enables the estimation of the filter scale and slope of the Smagorinsky scheme. Furthermore, it confirms the validity of the established relationship between the filter scale σ and the effective grid scale $\widehat{\overline{\Delta}}$. This relationship was derived to relate the Gaussian-filtered data back to a grid-based domain, with it being established as $\widehat{\overline{\Delta}} = 2\sigma$. This relationship was derived by analysing the energy distribution at the smallest scales. Overall, this work provides the basis for understanding the Smagorinsky subgrid scheme and its role as a filter in LES. This establishes a foundation for subsequent investigations into the impact of the grey-zone on the Smagorinsky and Gaussian filters, as well as potential strategies for mitigating these effects.

9.1.2 Flow Regime Dependence of the Smagorinsky Parameter

This section of research is aimed at identifying variations in Smagorinsky parameter values and behaviours across different flow regimes within the CTBL, with a particular focus on the ML, CFE, and IC regions. The IC areas can be further subdivided into the CU and CC regions. While these CU/CC flow regimes do exhibit some differences in C values compared to the more general IC region, no significant or consistent variations were identified across all parameters, scales, or test cases. The main difference observed between the IC and CU/CC regions were in the C_{θ_L} and C_{q_t} parameters at scales between $8\Delta \leq \widehat{\widehat{\Delta}} \leq 16\Delta$. At these coarse scales, it remains undetermined if these discrepancies observed in the data are due to differences in regime, or if it is a result of the expansion of areas being identified as "cloudy" when the resolution of the filtered fields coarsens (this will be discussed further in Section 9.3.1). At coarse grid scales, the LES model is not expected to be able to resolve the small-scale cloud updrafts and cores. Furthermore, even at higher resolutions (where less difference was evident between the IC and CU/CC values was observed), it would be computationally expensive to further divide the cloudy regions into the updraft and core components in the LES model. It would involve imposing conditional checks on the upward velocity and buoyancy fields at each grid point and time step to determine each C value. Therefore, the sub-division of IC areas into CU and CC regions was disregarded, and the focus remained instead on the IC, ML, and CFE.

The ML showed consistent behaviour across the profiles of Smagorinsky parameters for every variable (momentum and scalars), with parameter values at their local maximum in the centre of the ML, while distinct decreases in value are evident at the top and bottom boundaries. The CFE, in contrast, showed more variability across the different Smagorinsky parameters, with momentum parameters continuing to show higher values in certain areas, whereas the scalar parameters took near-zero values across nearly the entire CFE. Each parameter agreed that a reduction in C values occurred as the capping inversion was crossed in the upward direction, with the CFE experiencing lower C values for all parameters than their corresponding values in the ML. This is due to the weakened turbulence in the CFE, as it is a mostly stable layer outside of the cloudy areas. The clear height-dependence in the non-cloudy areas of the CTBL would enable a height-based parametrization scheme for the NC regions in subsequent work.

While a clear distinction exists between the ML and CFE, it is the IC regions where the most pronounced differences are observed. Within the IC regions, the scalar variables (C_θ, C_{θ_L}, and C_{qt}) exhibited marked differences from the momentum parameter C_s, highlighting the distinct nature of scalar transport within clouds. Furthermore, the θ variable is not conserved IC, which results in the cloudy areas being characterized by counter-gradient fluxes of θ, leading to a high prevalence of negative IC C_θ^2 values. This suggests that using a conserved thermodynamic variable, such as θ_L, in the parametrization would be a more suitable approach to reduce the number of negative Smagorinsky coefficients. The Prandtl and Schmidt numbers derived from the scalar parameters indicate that the heat and moisture structures are consistently larger than the momentum structures, particularly IC. The parameters for q_t and θ_L generally exhibited strong agreement in-cloud. A notable characteristic observed across all IC parameters was a near-constant C value maintained throughout the entire depth of the CL.

These regime-dependent trends demonstrate the limitations of using a single, fixed Smagorinsky coefficient for the entire CTBL. For this reason, when running LES in the grey-zone regime, the standard Smagorinsky scheme fails to capture the variations in mixing and dissipation across the different regions of the CTBL. Therefore, the findings detailed in Chapter 6 set the stage for the development of a parametrization scheme that captures the effect of various regimes on the Smagorinsky parameter. This forms one of the core motivations for the work presented in Chapter 8.

9.1.3 Scale-Dependence of the Smagorinsky Parameter

Building on the regime-dependent analysis, Chapter 7 explored the influence of the filter scale on the Smagorinsky parameters within each flow regime. This was necessary to understand how turbulence modelling in LES is affected by varying grid scales, particularly in the transition from fine to coarse resolutions.

The response of Smagorinsky parameters to changes in filter scale is regime-dependent. This underscores the necessity for LES models to be both scale-adaptive and equipped with regime-specific parameters. In the ML, all parameters exhibit a log decrease in magnitude with increasing grid scale, maintaining consistency across all variables. In contrast, the momentum parameter in the CFE shows a rapidly decreasing collapse to zero, while the scalar parameters in the CFE demonstrate a linear decrease in response to increasing filter scale. The response to filter scale changes shown by the IC parameters meanwhile, were observed to be highly dependent on their corresponding variable. Furthermore, a higher proportion of negative IC C values (for all parameters) occurs as the filter scale increases. While this increase in negative C values is not entirely an artificial outcome, it is significantly driven by the substantial rise in regions being identified as IC, rather than solely reflecting the response of the IC Smagorinsky parameters to the filter scale. This strongly suggests that additional research is required into the impact of the grey-zone on the behaviour of fluxes within shallow cumulus clouds. The variability of the Smagorinsky parameter values throughout entire regimes declined significantly with increasing filter scale, implying that the model's capability to distinguish between different flow structures diminishes as the resolution becomes coarser. The largest improvement which was observed across all cases, parametrization configurations, and scales (particularly coarse scales) resulted from including scale-adaptivity in the parametrization. This implies that allowing the C parameters to adapt to the grid scale is key in offsetting the grey-zone impacts at coarse scales.

The findings throughout this chapter directly complement those from Chapter 6 by demonstrating that the regime-dependent behaviours observed in Smagorinsky parameters are not only a function of height and thermodynamic properties but are also significantly influenced by grid scale. This reinforces the need for a parametrization scheme that is both regime-dependent and scale-adaptive, providing further motivation for the new parametrization which has been derived in Chapter 8.

9.1.4 Variable-Dependence of the Smagorinsky Parameter

Chapters 6 consistently highlighted a fundamental characteristic: the Smagorinsky parameter exhibited a strong dependence on the underlying variables. While certain key behaviours were consistently observed across all Smagorinsky parameters, such as decreases in the C value at the ML capping inversion, cloud top, and surface, other characteristics of the Smagorinsky parameters were distinctly linked to specific variables. Furthermore, the Smagorinsky parameters for scalars displayed more variation and nuance depending on the flow regime and filter scale. This strongly suggests that subgrid schemes should allow the scalar parameters to vary independently from the momentum parameter.

The most pronounced variable dependency was observed in the Smagorinsky parameter for non-conserved variables, such as θ and θ_v. Since these thermodynamic variables are not conserved in cloudy regions due to their inability to account for the effects of condensation, counter-gradient transport began to occur IC. As a result, the dynamic Smagorinsky scheme attempted to represent the upscale energy transfer associated with these negative gradients in IC by setting $C_\theta^2 \leq 0$, indicating energy transfer in the opposite direction of dissipation. However, since the Smagorinsky scheme does not account for backscatter or upscale energy transport, the model can only clip these negative values to $C_\theta = 0$ in IC. This effectively halts the mixing and dissipation of θ in the clouds. The behaviour of C_θ in IC, which shows a high percentage of negative C^2 values, leading to $\langle C_\theta \rangle = 0$ in IC, is consistent with the findings from the online dynamic Smagorinsky model, as noted in Shi et al. (2018) and Efstathiou et al. (2024).

The other key difference in the behaviour of the scalar parameters was observed in the ML values of C_{q_t}. While the ML C_{q_t} parameter did show a decrease in value with filter scale, much like the other parameters, this decrease was much more gradual than all other variables analysed throughout this work. Despite C_{q_t} having similar values to the other scalar parameters at high resolutions, such as C_{θ_L}, as the filter scale increases the C_{q_t} values decrease at a much slower rate in the ML than the other Smagorinsky parameters. Furthermore, $C_{q_t}^2$ was the only parameter in the ML observed to show a decline in the number of negative values as filter scale increased, indicating that the number of points experiencing counter-gradient transport of q_t decreases as resolution is coarsened. This suggests that the ML q_t structures are consistently larger and persist to coarser scales than both the heat and momentum structures, and therefore are not as heavily impacted by the grey-zone regime.

9.1.5 A New Parametrization for the Smagorinsky Subgrid Scheme

Chapter 8 culminated in the development of a new Smagorinsky parametrization that addresses some of the limitations of the standard Smagorinsky scheme which have been identified in previous chapters. This parametrization integrates flow regime and grid-scale dependencies in a height-based function to provide a more accurate representation of the behaviour of turbulence within the CTBL.

The systematic responses of the Smagorinsky parameters to flow regime and filter scale was used to derive a new height-base, regime-dependent parametrization for the Smagorinsky parameters. The functional relation with height was derived using a modified Blackadar formula in cloud-free regimes, while a fixed value approach was deemed sufficient to capture the behaviour of parameters IC. This parametrization was further adapted to be scale-dependent, which resulted in a marked improvement in the model's performance and ability to resolve CTBLs across a range of resolutions, particularly in the grey-zone of LES. The empirically derived parametrization scheme was tested by including the relation into the model and simulating the marine CBL BOMEX case and the diurnally evolving ARM case. The parametrization for C was shown to improve both computational efficiency and accuracy in resolving cloud fields, especially at coarser resolutions.

The development of the new parametrization in Chapter 8 is a direct response to the regime-dependent and scale-dependent challenges identified in Chapters 6 and 7, respectively. By incorporating height and scale dependencies, the new scheme effectively extends the LES regime to coarser grid scales, providing a computationally effective approach to modelling CTBL in the grey-zone regime.

9.2 Responses to the Overarching Research Questions

(1) Does the Smagorinsky scheme exhibit a specific filter shape that scales uniformly with grid spacing, and can its filter scale be derived? How does the Smagorinsky filter compare to the Gaussian filter? How can the Gaussian-filtered data be related back to a grid-based scale?

The findings from Chapter 5 provide significant insights into the behaviour of the Smagorinsky subgrid scheme. The Smagorinsky scheme has been observed behaving as a grid-scale based filter, demonstrating a distinct and consistent shape in spectral space that scales linearly with grid spacing. This subgrid scheme was found to closely resemble the Gaussian filter in spectral space, particularly at high resolutions. This similarity enables the estimation of filter scales for the Smagorinsky scheme. Meanwhile, the Gaussian filter scale can be related back to a grid-based scale, defined as the effective grid spacing $\widehat{\overline{\Delta}}$, using the relationship $\widehat{\overline{\Delta}} = 2\sigma$. This allows for the direct comparison between the Smagorinsky-filtered and Gaussian-filtered datasets.

Overall, this chapter provides an understanding of the Smagorinsky subgrid scheme's role as a filtering mechanism in LES modelling, and the Gaussian filters behaviour relative to grid-based domains. The work carried out in this chapter found critical insights into the behaviour of these filters, which are essential for justifying the use of Gaussian filters as substitutes for the Smagorinsky scheme in the offline dynamic model. The relationships established between grid scale and filter scale provide a foundation for further investigations into the behaviour of the Smagorinsky scheme at coarser, grey-zone grid-scales. Therefore, by addressing these research questions, the work in this chapter lays the groundwork for further developments in parametrization strategies which can be employed within the Smagorinsky subgrid scheme to mitigate grey-zone effects.

(2) Do the Smagorinsky parameters exhibit systematic patterns based on flow regimes, and if so, are there significant differences in the transport and diffusion characteristics of momentum, heat, and moisture within each of these regimes?

The Smagorinsky parameters take on distinct characteristics depending on the flow regime that they are in, namely the ML, CFE, and IC regions. The IC regime can be further subdivided into CU and CC regions. With the exception of certain parameters showing slightly higher C values in the CU/CC areas, there was little evidence of differing systematic behaviours between the IC, CU, and CC regions. Therefore the main differences in systematic responses to the regime were found to occur in the ML, CFE, and IC regions. The non-cloudy areas of the CTBL, that is the ML and CFE regimes, follow predictable patterns with height decreasing from the ML C value as the flow transitions across the inversion layer to the CFE. Therefore a height-dependent parametrization is employed to effectively capture this height-based regime dependence. Meanwhile, the IC parameters are roughly constant throughout the depth of the cloud layer, leading to C_{IC} being approximated by a fixed constant.

The ML Smagorinsky parameters showed consistent, parabolic-shaped profiles for all variables in the ML, with max values occurring in the mid-ML levels and distinct decreases in value as the boundaries are approached. Above the ML, the C values for each variable show a marked decrease within the CFE. In the CFE close to the ML capping inversion, both Prandtl and Schmidt numbers for all the scalars are near unity, suggesting that the heat and moisture structures are of the same scale as the momentum structures. However, as the top of the CL is approached, the Pr and Sc values in the CFE increase above 1, indicating that the moisture and heat structures are now smaller than the momentum structures — an unusual outcome for the CTBL. The IC parameter values meanwhile, exhibit distinct differences between the scalar variables (C_θ, C_{θ_L}, & C_{qt}) and the momentum parameter C_s, highlighting the distinct nature of scalar transport within clouds. The average C_s values IC were approximately constant with height and only slightly larger than in the ML. The parameters for conserved scalars, however, exhibited much larger values than the ML, but also remained approximately constant throughout the CL depth. Meanwhile, the Smagorinsky parameter for non-conserved θ scalar collapsed to zero IC.

Regarding the variations between the Smagorinsky parameters themselves, independent of flow regime, the following traits were noted. The momentum field experienced high levels of noise, and lacked a clear structure. Differences in the mean C_s values between regions were evident, albeit minimal. Meanwhile, the parameter fields for conserved variables exhibited much less noise and a higher degree of structure was evident in each of the flow regimes. There was also clear distinctions between each flow regime in the average Smagorinsky parameter values for these conserved variables. The non-conserved variables were observed to behave much the same as the conserved variables in all flow regimes but one - the IC region. In this regime, the non-conserved scalars experience counter gradient transport resulting in unrealistic C^2 values which are clipped to zero. Consequently, it was concluded that it is important to allow the parametrization to distinguish between both the variables and the flow regime.

(3) What is the impact of filter scale on the various Smagorinsky parameters across the three flow regimes of interest?

The impact of filter scale on the value of a Smagorinsky parameter is both regime-dependent and variable-dependent. The results show that the Smagorinsky parameter values across all flow regimes generally decrease as filter scale increases, albeit at different rates. In the ML, the C values decrease logarithmically as filter scale increases, indicating that turbulence in this region remains well-represented even as the grid scale increases. The distribution of C values also remains normal for all variables as filter scale increases. In contrast to this, the effect that filter scale has on the Smagorinsky parameters IC and in the CFE is much less consistent than what was observed in the ML. Both of these CL regimes exhibit a more rapid decrease in parameter values with increasing filter scale, particularly for scalar variables. In addition to this, the decrease does not follow a consistent response to changes in filter scale between each parameter, rather it behaves more erratically than trends observed in the ML C values. The distributions of C for all variables also demonstrate a strong scale dependence, becoming more and more skewed as the mode shifts to near zero when the filter scale increases. In the ML, the parameters demonstrate a regular decline with increased grid scale, while the IC and CFE regions show less consistent behaviour, most likely due to changes in cloud coverage and the model's representation of turbulence at coarser scales. This distinct difference between the behaviour of C in the ML compared to IC and in the CFE, suggests it might be a result of the limitations of filtering the CL to very coarse scales. This will be discussed further in Section 9.3.1. If this is the underlying cause, then it can be assumed that the behaviour of C IC determined at high resolution remains valid, though further investigation is necessary to understand its behaviour at coarser scales.

(4) Can a scale-adaptive relationship be established between height and the Smagorinsky parameters within each flow regime? Does this empirical parametrization improve LES of CTBLs in the grey-zone?

If the CTBL is partitioned into three distinct flow regimes consisting of the ML, the IC regions, and the CFE, then it is possible to derive a parametrization which captures the height and scale dependencies of the Smagorinsky parameters. The height-based component of this new parametrization is based upon a similar function to the Blackadar mixing length formulation. It is applied to the non-cloudy areas of the CTBL, enabling a smooth transition in Smagorinsky values across the ML, CFE, and OL. Fixing the various IC C parameters to a constant value throughout the entire depth of the CL was found to be a sufficient representation of the typical Smagorinsky parameter behaviour. Additionally, the parametrization incorporates grid scale dependency by adjusting the Smagorinsky values as a function of the filter scale. This adaptive approach ensures that the model retains high levels of accuracy across a range of resolutions, particularly in the grey-zone of LES. This new parametrization can be applied selectively to specific Smagorinsky parameters. If applied to C_s only, then the scalar parameters remain reliant on the Prandtl number along with the momentum parameter. This parametrization can also be applied to both C_s and C_ψ, using different variable-specific dependencies, for any scalar ψ. The flexibility of this parametrization allows it to perform effectively across a range of scenarios.

The proposed parametrization was implemented and tested in the MONC model using various different configurations. This allowed for the contributions from the addition of the height (regime) dependency, scale dependency, and variable dependency to be analysed. All parametrization configurations showed substantial improvements in both accuracy at grey-zone scales and computational efficiency compared to both the standard and dynamic Smagorinsky models. At coarse resolutions, the height-based element of the parametrization scheme was noted to delay the onset of the grey-zone by better capturing the dynamics of cloud formation and development. This led to more accurate simulations of the diurnal cloud cycle, particularly in terms of timing and vertical structure. Furthermore, the addition of the profile in the Smagorinsky parametrization resulted in quicker run times with fewer time steps required, compared to the standard model for the same test case.

The inclusion of scale-adaptive Smagorinsky values further improved both model accuracy and computational efficiency. By employing scale-adaptive parameters, the model was able to use larger time steps without sacrificing stability, leading to faster simulations. A two-fold increase in computational efficiency was achieved when the height-dependent, scale-aware parametrization scheme was implemented in MONC, compared to the standard model. This gain in efficiency did not compromise accuracy, as substantial improvements in the simulation of cloud field formation and evolution were observed at grey-zone resolutions when the parametrization was implemented. The most notable improvements in accuracy occurred at the coarsest scales, indicating that this parametrization is particularly advantageous at grey-zone resolutions. This implies that this parametrization for the Smagorinsky subgrid scheme, is particularly beneficial at grey-zone resolutions. The combined benefits of increased computational efficiency and improved accuracy make this new parametrization a valuable tool for future LES studies of CTBL.

9.3 Wider Implications, Limitations, and Future Work

The results presented in this thesis highlight the benefits of including a parametrization scheme for C in LES subgrid models. The parametrization which has been developed allows the C values to adapt to the grid scale of the model, as well as to the various flow regimes within the CTBL. This is achieved by integrating both height and scale dependencies, providing a straightforward and computationally efficient approach to mitigating the impacts of the grey-zone. By using a regime-dependent, height-based approach, the parametrization better captures the unique turbulence structures associated with cloud formation, leading to more accurate simulations of cloud fields. The scale-adaptive nature of the new parametrization allows for larger time steps, reducing computational costs while maintaining accuracy. When this scheme was employed in the MONC model a clear improvement was evident in the representation of cloud initiation and development in LES models. Therefore, this parametrization improves simulations of CTBLs at coarse resolutions, without incurring the computational expense of an online dynamic scheme which needs to continuously update the C_s value at each grid point and time step.

The research conducted in this thesis demonstrates that this novel approach to parametrizing C, using relatively simple modifications to the existing Smagorinsky subgrid scheme, can yield significant improvements in the simulation of CTBL. These improvements in both model accuracy and computational efficiency, are applicable across a range of grid scales, however, it is simulations in the grey-zone which experience the most substantial improvements from this parametrization. The simplicity and versatility of the new scale-adaptive, regime-dependent parametrization makes it suitable for incorporation into other models. With minimal tuning of the α and γ parameters in Equations 8.2, 8.3, 8.4, and 8.5, it is reasonable to expect that the parametrization (Equations 8.7 and/or 8.10) could be adopted by the broader modelling community and deliver immediate benefits.

One of the key contributions of this work is the clear demonstration that the modified Smagorinsky scheme provides a computationally efficient alternative to the more resource-intensive dynamic models. It is well known that the online dynamic model can improve results through flow-dependent adjustments of the Smagorinsky parameter at each grid point and time step. However, the analysis and testing carried out during this study pinpoint which aspects of the dynamic treatment of C lead to better outcomes. It was found that by introducing a height-dependent, scale-aware Smagorinsky parametrization, the benefits of improved accuracy and stability are achieved without the substantial computational cost associated with running dynamic models online. Furthermore, the new parametrization not only outperforms the standard Smagorinsky model in computational efficiency, but also provides more accurate timings of initial cloud formation along with improved representation of the cloud layer evolution, especially at grey-zone resolutions. This balance of efficiency and precision makes it a valuable tool for use in LES modelling of CTBLs.

An important outcome of this work is its potential to shape future research within the ParaChute project. ParaChute is a four year collaboration project between the UK Met Office and NERC funded researchers at multiple different universities. The main objective of this project is to improve the representation of turbulence in kilometre and sub-kilometre scale models, with the aim of improving the accuracy of forecasts for extreme weather events. The insights gained from the findings presented throughout this thesis are expected to inform decisions made during the ParaChute project and potentially influence the design of this ongoing research. Therefore, this PhD research will contribute both to the theoretical understanding of turbulence modelling at grey-zone scales and to practical advancements in the field.

9.3.1 Constraints and Limitations of this Study

While this study offers valuable insights into turbulence behaviour and parametrization methods within the grey-zone of the CTBL, certain limitations must be acknowledged. Firstly, the analysis primarily focuses on two case studies, BOMEX and ARM, which represent a quasi-steady marine CTBL and an evolving diurnal CTBL over land, respectively. Although these cases capture different boundary layer dynamics, they may not fully reflect the full range of complexities found in real-world CTBLs. Additionally, the filtering of cloud fields to coarse resolutions presents a challenge. It is well known that at coarse filter scales, the q_{cl} content becomes widely dispersed, causing the entire cloud layer to be identified as IC. This raises concerns about the validity of conclusions drawn about turbulent behaviour in the CL at the coarsest resolutions. However, trends from higher resolutions were extrapolated by eye for comparison and generally yielded consistent results with the coarse resolution data.

The framework of the MONC model posed a further limitation, as it assumes uniform diffusion for all scalars, based solely on heat diffusivity. Therefore, significant alterations to the model code would have been required to allow independent variation of scalar diffusivities. Consequently, the scalar-dependent parametrization was based on one parameter only. Following the analysis of multiple thermodynamic variables, the C_{θ_L} parameter was chosen, rather than C_θ, despite MONC using θ as the basis of scalar diffusivity. This is because the non-conserved nature of θ in-cloud leads to counter-gradient fluxes of θ in these regions, resulting in the diffusivity being set to zero. Therefore, if the parametrization had used a diffusion based on C_θ, then the in-cloud diffusion would have been set to zero for all parameters, due to the non-conserved nature of θ IC, resulting in counter gradient θ fluxes. In contrast, the cloud-conserved variables C_{θ_L} and q_t demonstrated similar values that were consistently non-zero IC, making C_{θ_L} a more appropriate and general choice for the diffusion parameter. Further to this, it is valid to diffuse θ with a C_{θ_L}, as discussed in Section 8.3.3, therefore C_{θ_L} was chosen as the diffusion parameter. Despite these limitations, the study's approach is valid, however, future work should aim to expand its scope and validate the findings across a broader range of CTBLs.

9.3.2 Future Research Directions

Further Testing and Optimization: The parametrization scheme should be tested in a broader range of convective cases to identify its limitations and potential for optimization. This includes applications to different shallow cloud types, such as stratocumulus, to assess the model's ability to adapt to different cloud regimes. Tests can be carried out to investigate if improvements can be achieved when simulating deep convection, by using this parametrization when simulating the preceding shallow cumulus environment in the lead-up to deep convection. The LBA case, based on observations from 23rd February 1999 over the Amazon, serves as a well-studied example detailing convective development and a transition from shallow to deep convection over land (Grabowski et al., 2006). This case could be used as a test case for this investigation, as it transitions through the same phases as the ARM case, before continuing to develop to a deep convection state. The fields obtained from running the LES model with the new parametrization at grey-zone resolutions could then be compared to high-resolution fields of the same case, following the methodology used for other cases in this thesis. Adapting and testing the new scheme within other LES models could demonstrate its potential as a standardized approach for parametrizing Smagorinsky coefficients in CTBL simulations.

Scalar Dependence: Analyses conducted in this thesis provides evidence that different values of Smagorinsky parameters might be beneficial for individual scalars in the cloudy regions of CTBLs. A significant area for advancement is the development of a new routine within the MONC model that differentiates between the variables being transported and diffused, such as heat and moisture. This would enable the model to assign scalar-specific Smagorinsky parameters, with the intention of improving the representation of each variable's distinct turbulent characteristics. Furthermore, it is recommended to carry out a test to evaluate a parametrization configuration where the IC $C_\theta = 0$, to assess whether this adjustment would improve cloud simulations, as suggested by the findings in this thesis. Extending this approach, another possible avenue to investigate would be to apply the dynamic Smagorinsky analysis to a passive tracer. The results could be insightful, as it would allow for a comparison between the diffusion of a non-interactive variable and that of heat and moisture. This comparison would yield deeper insights into the scalar-specific transport and diffusion processes within the CTBL. Overall, this work would further investigate if employing individual Smagorinsky parameters for each variable is beneficial, or if certain parameters have a more pronounced influence than others.

Time Dependence: Future research should focus on incorporating the time-dependent aspects of C into the parametrization. The C_{θ_L} and C_{q_t} parameters demonstrate their IC values increasing with time as the diurnal cycle progresses (see Figure 8.2). This remains, as of yet, unaccounted for in the proposed parametrization scheme. Additionally, the development of thermals as the ML evolves from a nocturnal shear-driven layer to a daytime convective layer warrants further investigation. Analyzing the developing ML during the early stages of the diurnal cycle may reveal significant time-dependent behaviour in ML C values until the convective circulation is fully established. Understanding how this transition influences the Smagorinsky parameters could provide valuable insights. Further to this, previous analyses in Section 6.1, indicates that the Smagorinsky parameters for scalars display their most extreme ML values within the thermals. To this end, the ML region could be further partitioned into thermals and their surrounding environments, to determine if this decomposition affects the average C values. It is hypothesized that diffusion within thermals in the ML may exhibit a temporal behaviours similar to that observed in-cloud during the development stages of the cloud. Specifically, we would expect the scalar Smagorinsky parameter values within the thermals to increase as the thermal circulation strengthens, ultimately reaching the critical threshold. By accounting for these temporal variations in the parametrization, it is anticipated that the model might better simulate the initial growth of the ML, which in turn may improve the initiation and growth of clouds.

9.4 Concluding Remarks

This thesis has made significant progress in advancing the understanding of turbulent length scale behaviour within CTBL simulations. The analysis has highlighted that Smagorinsky parameters, which determine these length scales, exhibit strong sensitivity to various factors such as flow regime, filter scale, and the conservation properties of the variables being mixed and dissipated. A new parametrization scheme has been proposed to account for these dependencies within the Smagorinsky model, and it has been found to improve both model performance and accuracy. Furthermore, this work has laid the foundation for addressing some of the key challenges associated with the grey-zone problem in LES modelling of turbulent boundary layers, where turbulent processes fall between resolved and unresolved scales. The findings suggest that by refining how sub-grid turbulence is handled, particularly through scalar-specific treatment of diffusion, we can achieve more reliable predictions in grey-zone simulations. The new parametrization scheme marks an important step forward, offering a pathway for future research to build upon. This work paves the way for future advancements in LES modelling of turbulent boundary layers, bringing the community closer to mitigating the grey-zone problem in CTBL simulations.

References

Ahlgrimm, M., R. M. Forbes, J.-J. Morcrette, and R. A. J. Neggers, 2016: ARM's impact on numerical weather prediction at ECMWF. *Atmospheric Radiation Measurement (ARM) Program: The first 20 years*, D. D. Turner, and R. G. Ellingson, Eds., Vol. 57, Amer Meteorological Society, Boston, doi:10.1175/AMSMONOGRAPHS-D-15-0032.1.

Beare, R. J., 2014: A length scale defining partially-resolved boundary-layer turbulence simulations. *Boundary-Layer Meteorology*, **151 (1)**, 39–55, doi:10.1007/s10546-013-9881-3.

Betts, A. K., 1973: Non-precipitating cumulus convection and its parameterization. *Quarterly Journal of the Royal Meteorological Society*, **99 (419)**, 178–196, doi:10.1002/qj.49709941915.

Blackadar, A., 1962: The vertical distribution of wind and turbulent exchange in a neutral atmosphere. *Journal of Geophysical Research*, **67 (8)**, 3095–3102, doi:10.1029/JZ067i008p03095.

Bony, S., and J.-L. Dufresne, 2005: Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. *Geophysical Research Letters*, **32 (20)**, doi:10.1029/2005GL023851.

Bony, S., B. Stevens, F. Ament, S. Bigorre, P. Chazette, S. Crewell, J. Delanoë, K. Emanuel, D. Farrell, C. Flamant, S. Gross, L. Hirsch, J. Karstensen, B. Mayer, L. Nuijens, J. H. Ruppert, I. Sandu, P. Siebesma, S. Speich, F. Szczap, J. Totems, R. Vogel, M. Wendisch, and M. Wirth, 2017: EUREC4A: A Field Campaign to Elucidate the Couplings Between Clouds, Convection and Circulation. *Surveys in Geophysics*, **38 (6)**, 1529–1568, doi:10.1007/s10712-017-9428-0.

Bou-Zeid, E., C. Meneveau, and M. Parlange, 2005: A scale-dependent Lagrangian dynamic model for large eddy simulation of complex turbulent flows. *Physics of Fluids*, **17 (2)**, 025 105, doi:10.1063/1.1839152.

Bretherton, C. S., J. R. McCaa, and H. Grenier, 2004: A new parameterization for shallow cumulus convection and its application to marine subtropical cloud-topped boundary layers. Part I: Description and 1D results. *Monthly Weather Review*, **132 (4)**, 864–882, doi:10.1175/1520-0493(2004)132⟨0864:ANPFSC⟩2.0.CO;2.

Brient, F., and T. Schneider, 2016: Constraints on Climate Sensitivity from Space-Based Measurements of Low-Cloud Reflection. doi:10.1175/JCLI-D-15-0897.1.

Brown, A. R., R. T. Cederwall, A. Chlond, P. G. Duynkerke, J.-C. Golaz, M. Khairoutdinov, D. C. Lewellen, A. P. Lock, M. K. MacVean, C.-H. Moeng, R. a. J. Neggers, A. P. Siebesma, and B. Stevens, 2002: Large-eddy simulation of the diurnal cycle of shallow cumulus convection over land. *Quarterly Journal of the Royal Meteorological Society*, **128 (582)**, 1075–1093, doi:10.1256/003590002320373210.

Brown, N., M. Weiland, A. Hill, B. Shipway, C. Maynard, T. Allen, and M. Rezny, 2014: A highly scalable Met Office NERC Cloud model.

Browning, K., A. Betts, P. Jonas, R. Kershaw, M. Manton, P. Mason, M. Miller, M. Moncrieff, H. Sundqvist, W. Tao, M. Tiedtke, P. Hobbs, J. Mitchell, E. Raschke, R. Stewart, and J. Simpson, 1993: The GEWEX cloud system study (GCSS). *Bulletin of the American Meteorological Society*, **74 (3)**, 65–80, doi:https://doi.org/10.1175/1520-0477(1993)074⟨0387:TGCSS⟩2.0.CO;2.

Chlond, A., O. Böhringer, T. Auerswald, and F. Müller, 2014: The effect of soil moisture and atmospheric conditions on the development of shallow cumulus convection: A coupled large-eddy simulation–land surface model study. *Meteorologische Zeitschrift*, 491–510, doi:10.1127/metz/2014/0576.

Cuijpers, J. W. M., 1994: Large eddy simulation of cumulus convection. PhD Thesis, Delft University of Technology, available at https://repository.tudelft.nl/record/uuid:52dcce07-aebd-41df-8ab6-084457ef2acd.

Dairay, T., E. Lamballais, S. Laizet, and J. C. Vassilicos, 2017: Numerical dissipation vs. subgrid-scale modelling for large eddy simulation. *Journal of Computational Physics*, **337**, 252–274, doi:10.1016/j.jcp.2017.02.035.

Davidson, B., 1968: The Barbados oceanographic and meteorological experiment. *Bull. Am. Meteorol. Soc.*, **49 (9)**, 928–935.

de Roode, S. R., H. J. J. Jonker, B. J. H. van de Wiel, V. Vertregt, and V. Perrin, 2017: A diagnosis of excessive mixing in Smagorinsky subfilter-scale turbulent kinetic energy models. *Journal of the Atmospheric Sciences*, **74 (5)**, 1495–1511, doi:10.1175/JAS-D-16-0212.1.

Deardorff, J. W., 1970: A numerical study of three-dimensional turbulent channel flow at large Reynolds numbers. *Journal of Fluid Mechanics*, **41 (2)**, 453–480, doi:10.1017/S0022112070000691.

Deardorff, J. W., 1971: On the magnitude of the subgrid scale eddy coefficient. *Journal of Computational Physics*, **7 (1)**, 120–133, doi:10.1016/0021-9991(71)90053-2.

Deardorff, J. W., 1980: Stratocumulus-capped mixed layers derived from a three-dimensional model. *Boundary-Layer Meteorology*, **18 (4)**, 495–527, doi:10.1007/BF00119502.

Durran, D., J. A. Weyn, and M. Q. Menchaca, 2017: Practical considerations for computing dimensional spectra from gridded data. *Monthly Weather Review*, **145 (9)**, 3901–3910, doi: 10.1175/MWR-D-17-0056.1.

Efstathiou, G. A., 2023: Dynamic Subgrid Turbulence Modeling for Shallow Cumulus Convection Simulations beyond LES Resolutions. *Journal of the Atmospheric Sciences*, **80 (6)**, 1519–1545, doi:10.1175/JAS-D-22-0132.1.

Efstathiou, G. A., and R. J. Beare, 2015: Quantifying and improving sub-grid diffusion in the boundary-layer grey zone. *Quarterly Journal of the Royal Meteorological Society*, **141 (693)**, 3006–3017, doi:10.1002/qj.2585.

Efstathiou, G. A., and R. S. Plant, 2019: A dynamic extension of the pragmatic blending scheme for scale-dependent sub-grid mixing. *Quarterly Journal of the Royal Meteorological Society*, **145 (719)**, 884–892, doi:10.1002/qj.3445.

Efstathiou, G. A., R. S. Plant, and M.-J. M. Bopape, 2018: Simulation of an Evolving Convective Boundary Layer Using a Scale-Dependent Dynamic Smagorinsky Model at Near-Gray-Zone Resolutions. *Journal of Applied Meteorology and Climatology*, **57 (9)**, 2197–2214, doi:10.1175/JAMC-D-17-0318.1.

Efstathiou, G. A., R. S. Plant, and F. K. Chow, 2024: Grey-zone simulations of shallow-to-deep convection transition using dynamic subgrid-scale turbulence models. *Quarterly Journal of the Royal Meteorological Society*, **n/a (n/a)**, doi:10.1002/qj.4817.

Emanuel, K., 1994: *Atmospheric Convection*. Oxford University Press.

Garratt, J., 1994: *The Atmospheric Boundary Layer*. Cambridge Atmospheric and Space Science Series, Cambridge University Press.

Germano, M., U. Piomelli, P. Moin, and W. H. Cabot, 1991: A dynamic subgrid-scale eddy viscosity model. *Physics of Fluids A: Fluid Dynamics*, **3 (7)**, 1760–1765, doi:10.1063/1.857955.

Goldberg, U. C., S. Palaniswamy, P. Batten, and V. Gupta, 2010: Variable turbulent Schmidt and Prandtl number modeling. *Engineering Applications of Computational Fluid Mechanics*, **4 (4)**, 511–520, doi:10.1080/19942060.2010.11015337.

Grabowski, W. W., P. Bechtold, A. Cheng, R. Forbes, C. Halliwell, M. Khairoutdinov, S. Lang, T. Nasuno, J. Petch, W.-K. Tao, R. Wong, X. Wu, and K.-M. Xu, 2006: Daytime convective development over land: A model intercomparison based on LBA observations. *Quarterly Journal of the Royal Meteorological Society*, **132 (615)**, 317–344, doi:10.1256/qj.04.147.

Grachev, A. A., E. L. Andreas, C. W. Fairall, P. S. Guest, and P. O. G. Persson, 2007: On the turbulent Prandtl number in the stable atmospheric boundary layer. *Boundary-Layer Meteorology*, **125 (2)**, 329–341, doi:10.1007/s10546-007-9192-7.

Gray, M. E. B., J. Petch, S. H. Derbyshire, A. R. Brown, A. P. Lock, H. A. Swann, and P. R. A. Brown, 2004: *Version 2.3 Of The Met Office Large Eddy Model: Part II. Scientific Documentation*.

Gualtieri, C., A. Angeloudis, F. Bombardelli, S. Jha, and T. Stoesser, 2017: On the values for the turbulent Schmidt number in environmental flows. *Fluids*, **2 (2)**, 17, doi:10.3390/fluids2020017.

Hill, A., N. Brown, and B. Shipway, 2018: *Met Office/NERC Cloud Model (MONC): User Documentation.*

Holland, J. Z., and E. M. Rasmusson, 1973: Measurements of the atmospheric mass, energy, and momentum budgets over a 500-kilometer square of tropical ocean. *Monthly Weather Review*, **101 (1)**, 44–55, doi:10.1175/1520-0493(1973)101⟨0044:MOTAME⟩2.3.CO;2.

Honnert, R., G. A. Efstathiou, R. J. Beare, J. Ito, A. Lock, R. Neggers, R. S. Plant, H. H. Shin, L. Tomassini, and B. Zhou, 2020: The atmospheric boundary layer and the "gray zone" of turbulence: A critical review. *Journal of Geophysical Research*, 26.

Honnert, R., V. Masson, and F. Couvreux, 2011: A diagnostic for evaluating the representation of turbulence in atmospheric models at the kilometric scale. *Journal of the Atmospheric Sciences*, **68 (12)**, 3112–3131, doi:10.1175/JAS-D-11-061.1.

Houze, R. A., 2014: *Cloud Dynamics*. 2nd ed., International Geophysics Series Volume 104.

Ito, J., H. Niino, M. Nakanishi, and C.-H. Moeng, 2015: An extension of the Mellor–Yamada model to the terra incognita zone for dry convective mixed layers in the free convection regime. *Boundary-Layer Meteorology*, **157 (1)**, 23–43, doi:10.1007/s10546-015-0045-5.

Johnson, R. H., 1976: The role of convective-scale precipitation downdrafts in cumulus and synoptic-scale interactions. *Journal of the Atmospheric Sciences*, **33 (10)**, 1890–1910, doi:10.1175/1520-0469(1976)033⟨1890:TROCSP⟩2.0.CO;2.

Kaimal, J. C., and J. J. Finnigan, 1994: *Atmospheric Boundary Layer Flows: Their Structure and Measurement*. Oxford University Press, New York.

Kaimal, J. C., J. C. Wyngaard, Y. Izumi, and O. R. Coté, 1972: Spectral characteristics of surface-layer turbulence. *Quarterly Journal of the Royal Meteorological Society*, **98 (417)**, 563–589, doi:10.1002/qj.49709841707.

Katul, G. G., D. Li, H. Liu, and S. Assouline, 2016: Deviations from unity of the ratio of the turbulent Schmidt to Prandtl numbers in stratified atmospheric flows over water surfaces. *Physical Review Fluids*, **1 (3)**, 034 401, doi:10.1103/PhysRevFluids.1.034401.

Kealy, J. C., G. A. Efstathiou, and R. J. Beare, 2019: The onset of resolved boundary-layer turbulence at grey-zone resolutions. *Boundary-Layer Meteorology*, **171 (1)**, 31–52, doi:10.1007/s10546-018-0420-0.

Leonard, B., M. MacVean, and A. Lock, 1993: Positivity-preserving numerical schemes for multidimensional advection. *NASA STI/Recon Technical Report N*, **93**, 27 091.

Li, D., 2016: Revisiting the subgrid-scale Prandtl number for large-eddy simulation. *Journal of Fluid Mechanics*, **802**, R2, doi:10.1017/jfm.2016.472.

Li, D., 2019: Turbulent Prandtl number in the atmospheric boundary layer - where are we now? *Atmospheric Research*, **216**, 86–105, doi:10.1016/j.atmosres.2018.09.015.

Lilly, D., 1966: On the application of the eddy viscosity concept in the inertial sub-range of turbulence. Tech. Rep. 123, UCAR/NCAR. doi:10.5065/D67H1GGQ.

Lilly, D. K., 1962: On the numerical simulation of buoyant convection. *Tellus*, **14 (2)**, 148–172, doi:10.1111/j.2153-3490.1962.tb00128.x.

Lilly, D. K., 1992: A proposed modification of the Germano subgrid-scale closure method. *Physics of Fluids A: Fluid Dynamics*, **4 (3)**, 633–635, doi:10.1063/1.858280.

Mason, P. J., 1994: Large-eddy simulation: A critical review of the technique. *Quarterly Journal of the Royal Meteorological Society*, **120 (515)**, 1–26, doi:10.1002/qj.49712051503.

Mason, P. J., and A. R. Brown, 1999: On subgrid models and filter operations in large eddy simulations. *Journal of the Atmospheric Sciences*, **56 (13)**, 2101–2114, doi:10.1175/1520-0469(1999)056⟨2101:OSMAFO⟩2.0.CO;2.

Mason, P. J., and N. S. Callen, 1986: On the magnitude of the subgrid-scale eddy coefficient in large-eddy simulations of turbulent channel flow. *Journal of Fluid Mechanics*, **162**, 439–462, doi:10.1017/S0022112086002112.

McMillan, O., J. Ferziger, and R. Rogallo, 1980: Tests of subgrid-scale models in strained turbulence. *13th Fluid and PlasmaDynamics Conference*, American Institute of Aeronautics and Astronautics, doi:10.2514/6.1980-1339.

Meneveau, C., T. S. Lund, and W. H. Cabot, 1996: A Lagrangian dynamic subgrid-scale model of turbulence. *Journal of Fluid Mechanics*, **319 (-1)**, 353, doi:10.1017/S0022112096007379.

Moeng, C. H., 1984: A large-eddy-simulation model for the study of planetary boundary-layer turbulence.

Moeng, C.-H., and J. C. Wyngaard, 1988: Spectral Analysis of Large-Eddy Simulations of the Convective Boundary Layer.

Moraes, O. L., D. R. Fitzjarrald, O. C. Acevedo, R. K. Sakai, M. J. Czikowsky, and G. A. Degrazia, 2008: Comparing spectra and cospectra of turbulence over different surface boundary conditions. *Physica A: Statistical Mechanics and its Applications*, **387 (19-20)**, 4927–4939, doi:10.1016/j.physa.2008.04.007.

Neggers, R. a. J., P. G. Duynkerke, and S. M. A. Rodts, 2003: Shallow cumulus convection: a validation of large-eddy simulation against aircraft and Landsat observations. *Quarterly Journal of the Royal Meteorological Society*, **129 (593)**, 2671–2696, doi:10.1256/qj.02.93.

Nitta, T., and S. Esbensen, 1974: Heat and Moisture Budget Analyses Using BOMEX Data. *Monthly Weather Review*, **102 (1)**, 17–28, doi:10.1175/1520-0493(1974)102⟨0017:HAMBAU⟩2.0.CO;2.

Norris, J. R., 1999: On trends and possible artifacts in global ocean cloud cover between 1952 and 1995. *Journal of Climate*, **12 (6)**, 1864–1870, doi:10.1175/1520-0442(1999)012⟨1864:OTAPAI⟩2.0.CO;2.

Nuijens, L., B. Medeiros, I. Sandu, and M. Ahlgrimm, 2015: The behavior of trade-wind cloudiness in observations and models: The major cloud components and their variability. *Journal of Advances in Modeling Earth Systems*, **7 (2)**, 600–616, doi:10.1002/2014MS000390.

Orlanski, I., 1975: A rational subdivision of scales for atmospheric processes. *Bulletin of the American Meteorological Society*, **56 (5)**, 527–530.

Piacsek, S. A., and G. P. Williams, 1970: Conservation properties of convection difference schemes. *Journal of Computational Physics*, **6 (3)**, 392–405, doi:https://doi.org/10.1016/0021-9991(70)90038-0, URL https://www.sciencedirect.com/science/article/pii/0021999170900380.

Piomelli, U., P. Moin, and J. H. Ferziger, 1988: Model consistency in large eddy simulation of turbulent channel flows. *The Physics of Fluids*, **31 (7)**, 1884–1891, doi:10.1063/1.866635.

Pope, S. B., 2000: *Turbulent Flows*.

Porté-Agel, F., C. Meneveau, and M. B. Parlange, 2000: A scale-dependent dynamic model for large-eddy simulation: Application to a neutral atmospheric boundary layer. *Journal of Fluid Mechanics*, **415**, 261–284, doi:10.1017/S0022112000008776.

Prandtl, L., 1925: Report on investigation of developed turbulence, translation of bericht fiber untersuchungen zur ausgebildeten turbulent. *Zeitschrift fuer Angewandte Matematik und Mechanik*, **5 (2)**.

Sakradzija, M., A. Seifert, and A. Dipankar, 2016: A stochastic scale-aware parameterization of shallow cumulus convection across the convective gray zone. *Journal of Advances in Modeling Earth Systems*, **8 (2)**, 786–812, doi:10.1002/2016MS000634.

Shi, X., F. K. Chow, R. L. Street, and G. H. Bryan, 2018: An evaluation of LES turbulence models for scalar mixing in the stratocumulus-capped boundary layer. *Journal of the Atmospheric Sciences*, **75 (5)**, 1499–1507, doi:10.1175/JAS-D-17-0392.1.

Shin, H. H., and S.-Y. Hong, 2015: Representation of the Subgrid-Scale Turbulent Transport in Convective Boundary Layers at Gray-Zone Resolutions. doi:10.1175/MWR-D-14-00116.1.

Siebesma, A. P., C. S. Bretherton, A. Brown, A. Chlond, J. Cuxart, P. G. Duynkerke, H. Jiang, M. Khairoutdinov, D. Lewellen, C.-H. Moeng, E. Sanchez, B. Stevens, and D. E. Stevens, 2003: A large eddy simulation intercomparison study of shallow cumulus convection.

Siebesma, A. P., and J. W. M. Cuijpers, 1995: Evaluation of parametric assumptions for shallow cumulus convection. *Journal of the Atmospheric Sciences*, **52 (6)**, 650–666, doi:10.1175/1520-0469(1995)052⟨0650:EOPAFS⟩2.0.CO;2.

Smagorinsky, J., 1963: General circulation experiments with the primitive equations: I. The basic experiment. *Monthly Weather Review*, **91 (3)**, 99–164, doi:10.1175/1520-0493(1963)091⟨0099:GCEWTP⟩2.3.CO;2.

Stull, 2009: *An Introduction to Boundary Layer Meteorology*.

Warhaft, Z., 1976: Heat and moisture flux in the stratified boundary layer. *Quarterly Journal of the Royal Meteorological Society*, **102 (433)**, 703–707, doi:10.1002/qj.49710243315.

Wyngaard, J. C., 2004: Toward Numerical Modeling in the "Terra Incognita". *Journal of the Atmospheric Sciences*, **61**, 11.

Wyngaard, J. C., and O. R. Coté, 1971: The budgets of turbulent kinetic energy and temperature variance in the atmospheric surface layer. *Journal of the Atmospheric Sciences*, **28 (2)**, 190–201, doi:10.1175/1520-0469(1971)028⟨0190:TBOTKE⟩2.0.CO;2.

Zhou, B., J. S. Simon, and F. K. Chow, 2014: The convective boundary layer in the terra incognita. *Journal of the Atmospheric Sciences*, **71 (7)**, 2545–2563, doi:10.1175/JAS-D-13-0356.1.