WAGENINGEN UNIVERSITY
METEOROLOGY AND AIR QUALITY

CLOUD ROOTS

M.Sc. Climate Studies

Specialization: The Physical CLimate System

Master Thesis

Diurnal variability of surface processes, turbulent properties and cloud development in the shallow cumulus regime during the Amazonian dry season

---

by

Sebastiaan de Haas

June 10, 2024

Meteorology & Air Quality,
Wageningen University & Research

Supervisor & First Examiner:
Prof. Dr. Jordi Vilà-Guerau de Arellano

Second Examiner:
Dr. Imme Benedict

Abstract

Cloud formation and the development of the atmospheric boundary layer (BL) in the Amazonia is controlled by a complex, interconnected myriad of processes related to both the surface and atmosphere. Understanding these processes is essential for improving the modelling of interactions between vegetation, atmosphere and cloud formation in weather and climate models. For this master thesis, the surface and turbulent processes underlying the onset of shallow cumulus formation have been investigated. To this end, we have employed a complementary approach by designing a numerical experiment using a large-eddy simulation (LES), supported and constrained by a unique and comprehensive set of atmospheric and surface measurements taken during the Amazonian dry season. Special attention is given to the transition from dry turbulence to moist convection that dictates the onset of cloud formation in the shallow cumulus regime, in addition to the processes underlying this transition. Therefore, we have investigated the asymmetry in the diurnal cycle of the surface turbulent fluxes of latent and sensible heat, which is crucial in determining BL moisture and the production of turbulence. Furthermore, the diurnal variability of turbulent kinetic energy and its distribution between the vertical and horizontal directions have been assessed. These are essential properties that characterize turbulent transport in the BL, which is the process responsible for lifting moist air parcels from the surface that initiate shallow cumulus formation. Additionally, by determining the evolution of potential temperature and specific humidity at various heights within the BL, we have been able to infer how turbulent mixing controls the transition between profiles dominated by large vertical variations and well-mixed conditions. Finally, the explicitness of the LES in simulating turbulence and cloud formation has enabled us to determine the transition from dry turbulence, quantified by the BL height, to moist convection, quantified by the lifting condensation level. We find that this transition occurs at 9:45 local time and is preceded by a rising turbulent kinetic energy that generates efficient vertical transport of moist air parcels, thereby being associated with a rapid increase in cloud cover. Additionally, turbulent transport is able to efficiently mix the potential temperature within the BL, however, this is not reflected in observations of the specific humidity. For future research, the LES case of shallow cumulus conditions designed in this research will be used to assess the sensitivity of the shallow to deep convection transition to surface and atmospheric conditions, in addition to the CO_2 transport by shallow cumulus clouds and the performance of a more detailed canopy scheme.

Acronyms & symbols

LES	Large Eddy Simulation
DALES	Dutch Atmospheric Large Eddy Simulation
ATTO	Amazon Tall Tower Observatory
EC	Eddy Covariance
LT	Local time
BL	Boundary layer (m)
LCL	Lifting condensation level (m)
H	Sensible heat flux (W m^{-2})
LE	Latent heat flux (W m^{-2})
TKE	Turbulent kinetic energy (m^2 s^{-2})
RMSE	Root mean square error
d	Index of agreement (in units of variable)
θ	Potential temperature (K)
q	Specific humidity (g kg^{-1})
q_l	Liquid water specific humidity (g kg^{-1})
Q_{net}	Net available radiation (W m^{-2})
u_*	Friction velocity (m s^{-1})
α_t	Vertical over horizontal anisotropy in wind variance
u	Zonal wind component (m s^{-1})
v	Meridional wind component (m s^{-1})
w	Vertical wind component (m s^{-1})

Contents

Chapter 1

Introduction and motivation

In tropical forests, a large fraction of the boundary layer (BL) moisture is regulated by vegetation through transpiration from the leaf stomata. The vegetation thereby influences cloud formation, as the moisture and also heat from the surface are transported upward in the lower atmosphere by turbulent thermals (Van Heerwaarden and Vilà-Guerau de Arellano, 2008). In turn, clouds and their properties determine the magnitude of incoming radiation and surface turbulent fluxes (latent and sensible heat flux), while also influencing plant photosynthesis and transpiration, thereby affecting atmospheric moisture content and CO_2 concentrations (Freedman et al., 2001). Such interactions between clouds, vegetation and atmosphere are typically not well represented in weather and climate models (Bonan et al., 2021). This is related to the complexity of a combination of chemical, physical and biological processes that operate on various temporal and spatial scales that can differ by orders of magnitude (Suni et al., 2015, see also their figure 7). To illustrate, ∼14% of the standard deviation in mean land warming is caused by misrepresenting stomatal response to environmental conditions in climate models from CMIP6 (6th Coupled Model Intercomparison Project, Park et al., 2021). The CloudRoots project (`https://cloudroots.wur.nl/`) aims to improve the modelling of cloud-atmosphere-vegetation interactions by using various observations and models, and by combining different disciplines and spatial/temporal scales (Vilà-Guerau de Arellano et al., 2024). Additionally, the CloudRoots project investigates how cloud-atmosphere-vegetation interactions differ between boreal, temperate and tropical ecosystems.

The two-way land-atmosphere interactions in the tropics are among the most uncertain in terrestrial climatology (Betts and Silva Dias, 2010, see also their figure 2). This uncertainty is further complicated by deforestation and climate change, which modify the future states of tropical forests and their corresponding climate (Spracklen et al., 2018). Among these tropical ecosystems is the Amazonian rainforest, which is an important carbon sink that is responsible for ∼25% of land CO_2 uptake (Phillips et al., 2009; Brienen et al., 2015). Additionally, the Amazon is a crucial component of the hydrological cycle, recycling ∼28% of precipitation and ∼48% of evaporation in the South American continent (Ent et al., 2010). The Amazon is therefore an important contributor in determining the climate, and in turn, the climate is essential for the state of the Amazonian ecosystem (Cox et al., 2004). However, the processes that dictate the exchange of energy, water and CO_2 between the Amazonian rainforest and atmosphere are complex and not well-understood (Vilà-Guerau de Arellano et al., 2024). These processes affect the regional and global climate, and

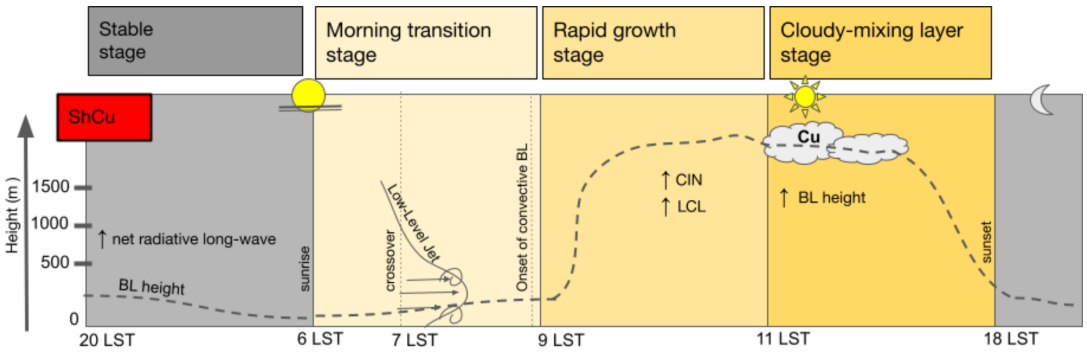

FIGURE 1.1: Typical atmospheric conditions on shallow cumulus days during the Amazonian dry season. Crossover indicates when the sensible heat first becomes positive, BL is short for boundary layer, CIN for convective inhibition, LCL for lifting condensation level, LST for local standard time and ShCu for shallow cumulus. This figure has been taken from Henkes et al. (2021).

an enhanced understanding of vegetation-atmosphere-cloud interactions in the Amazon is therefore crucial for improving climate projections.

An important mechanism for tropical climate dynamics is shallow cumulus convection (Neggers, Neelin, and Stevens, 2007). Shallow cumulus clouds induce shading on the surface, which have been shown to affect e.g. BL dynamics, turbulent transport and subsequently cloud characteristics in the Amazon (Horn et al., 2015). In Henkes et al. (2021), typical conditions of shallow cumulus days during the Amazonian dry season have been identified based on a combination of measurements and conceptual modelling. In Fig. 1.1, a conceptual figure from Henkes et al. (2021) is shown that summarizes these conditions for a 24-hour cycle. Four stages have been defined, the first one being the stable nocturnal stage, which is not considered in this research due to the focus on the diurnal cycle (between 6:00 and 18:00 local time). The morning transition stage, which occurs between 6:00 and 8:00 – 9:00 LT (local time), is characterised by three successive events based on a classification from Angevine, Baltink, and Bosveld (2001). The first is sunrise at 6:00 LT, followed by the time at which the sensible heat flux first becomes positive (crossover), and finally the onset of the convective BL. The latter is defined as the time at which the nocturnal BL has fully eroded and the BL height grows with a rate of > 100 m per hour, usually occurring between 8:00 and 9:00 LT. Subsequently, the rapid growth stage sets in, during which the BL height increases at high pace. The end of this stage is defined by the BL height first decreasing due to the onset of shallow cumulus formation, which is typically around 11:00 LT. Subsequently, the cloudy-mixing stage begins and is characterised by a mixed and deep convective BL with shallow cumulus development. This stage ends around 15:00 LT and is followed by a more rapid decrease in BL height, after which the stable stage again sets in.

Henkes et al. (2021) compare conditions on shallow cumulus days to those on which a transition to deep convection occurs, concluding that shallow cumulus days are typically less moist, warmer and less turbulent compared to their deep convective counterparts. Additionally, shallow cumulus days typically involve a low-level jet, in addition to a high BL height and lifting condensation level (LCL). The latter is defined as the height at which water vapor starts to condensate and form clouds, and depends on the vertical profiles of

temperature and specific humidity. In De Feiter (2023), differences in conditions between shallow and deep convective regimes have also been investigated based on measurements and conceptual modelling. It is concluded that shallow and deep convective days exhibit only subtle differences in conditions such as LCL, potential temperature and specific humidity, and that overall the results are in good agreement with the findings from Henkes et al. (2021).

In this research, we study shallow cumulus development and its underlying processes using a unique and comprehensive data set from the CloudRoots-Amazon22 campaign (see section 2.1), in addition to a turbulence and cloud resolving model (large eddy simulation, see section 2.2) that is especially suitable for simulating Amazonian forest-atmosphere interactions. The combination of this high-resolution model and unique observational data enables a detailed characterisation of the Amazonian diurnal cycle of shallow cumulus conditions during the dry season. We investigate the diurnal variability of the partitioning of latent and sensible heat flux, in addition to the turbulent properties within the BL. Turbulence dictates the degree of BL mixing, which in combination with latent and sensible heat flux largely determine the vertical profiles of potential temperature and specific humidity in the BL. In turn, these profiles shape shallow cumulus development and its characteristics. Combining the comprehensive observational data with the turbulence and cloud resolving model to study the shallow cumulus regime using this systematic approach has, to the best of our knowledge, not yet been done in earlier studies. We thereby formulate the following interconnected research questions:

1) Do the surface turbulent fluxes of latent and sensible heat evolve symmetrically in time during the diurnal cycle?

2) What is the diurnal variability of turbulent kinetic energy and how is this energy distributed between the vertical and horizontal directions?

3) Does turbulent transport mix potential temperature and specific humidity with the same efficiency?

4) When is the transition from clear to cloudy conditions as characterized by the transition from dry turbulence to moist convection?

To answer these research questions, we design a numerical experiment that is representative of shallow cumulus conditions in the turbulence and cloud resolving model, which is constrained by the observations of the CloudRoots-Amazon22 campaign. We note that the design of this numerical experiment is the main goal of this research, as it can be used for a broad spectrum of research applications in the Amazonia, which is discussed in section 4.2. Additionally, the numerical experiment can serve as a validation tool for regional and global models.

In chapter 2, the CloudRoots-Amazon22 campaign, high-resolution model and research strategy is described. Chapter 3 presents and discusses the results by answering each of the research questions, and additionally includes an analysis of model-observations statistics. Finally, in chapter 4, the most important conclusions are summarized and suggestions for future research are formulated. Additional results and a description of observational data pre-processing and model set-up are provided in appendix chapter 5.

Chapter 2

Methodology

For this thesis, a combination of observational data and Large Eddy Simulation (LES) modelling has been used to conduct the research. For the observations, an extensive and comprehensive data set from the CloudRoots-Amazon22 campaign has been used (Vilà-Guerau de Arellano et al., 2024), which is described in section 2.1. A numerical experiment has been designed with the Dutch Atmospheric Large Eddy Simulation (DALES; Heus et al., 2010; Ouwersloot et al., 2017), where a description of the model is provided in section 2.2. Finally, section 2.3 describes the research strategy, which includes the procedure used for designing the numerical experiment in DALES, which is constrained and supported by the comprehensive observational data. As mentioned, the design of the DALES experiment of a representative shallow cumulus day is the main goal of this research, as this experiment can be employed for a wide range of applications in Amazonian atmospheric research (see section 4.2).

2.1 CloudRoots-Amazon22 campaign data

The observational data from the CloudRoots-Amazon22 campaign (Vilà-Guerau de Arellano et al., 2024) that has been used in this research mostly consists of tower and radiosonde data. The measurements of the campaign were conducted during the late Amazonian dry season between 9 and 18 August 2022. A large fraction of the observational data used in this research and described in this section has been analysed by Vincent de Feiter (De Feiter, 2023). The observational data were collected at and around the 323m-tall Amazon Tall Tower Observatory (ATTO; Andreae et al., 2015). In Fig. 2.1, a map with the location of ATTO is shown. We also use data from a second tower, INSTANT (81 m tall), which is located at ∼80 m from ATTO. Sounding data from weather balloons were recorded several times per day at 6:00, 9:00, 12:00, 15:00 and 18:00 LT (where LT = UTC – 4) and were predominantly launched from the CAMPINA site. This site is located at a distance of ∼5 km from ATTO/INSTANT and is also displayed in Fig. 2.1. The radiosondes that were not launched from CAMPINA (12 out of 66) were released nearby ATTO.

The tower data were predominantly collected using a sonic anemometer with the Eddy Covariance (EC) method in 30-minute intervals. Relative humidity (RH) and temperature (T) were recorded by a thermohygrometer in 1-minute intervals. Additionally, CO_2 mixing ratio data were collected by a picarro, as EC CO_2 mixing ratio data is unreliable and should only be used to investigate trends. Furthermore, the thermohygrometer data consists of improved quality compared to the EC data. Therefore, the fluxes of CO_2, latent

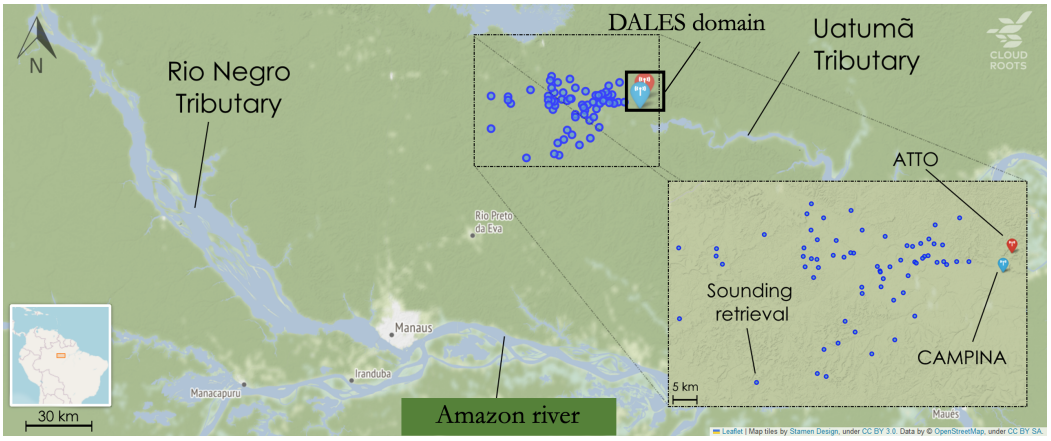

FIGURE 2.1: A map displaying the location of the ATTO (Amazon Tall Tower Observatory, latitude = –2.002°, longitude = –59.000°) and CAMPINA (latitude = –2.182°, longitude –59.022°) measurement sites within the Amazon, in addition to retrieval locations of the radiosondes. The INSTANT tower is not shown in the map, but is located at an ∼80 m distance from ATTO. The observational data used in this research and retrieved from the ATTO and CAMPINA sites were taken during the CloudRoots-Amazon22 campaign (Vilà-Guerau de Arellano et al., 2024). The domain size of DALES (Dutch Atmospheric Large Eddy Simulation, 19 x 19 km domain) has also been included in the map around the ATTO and CAMPINA sites (as the initial conditions of DALES are primarily based on measurements from these sites). This figure was created by Vincent de feiter and adapted for this thesis.

and sensible heat from the EC have been multiplied by a correction factor using thermohygrometer data. Additionally, some variables from EC data have been recalculated from the thermohygrometer data, which includes specific humidity. Further details on the data quality of the thermohygrometer and EC are provided in appendix section 5.1, which also describes the flux corrections and recalculations. Additionally, this section contains a table that lists the most important available tower data from the CloudRoots-Amazon22 campaign, including the heights at which they are available (appendix table 5.1).

As mentioned, the tower data consists of a combination of ATTO and INSTANT. The INSTANT tower collected data at heights of 5, 15, 25, 35, 50 and 81 m. Data from ATTO were recorded at 43, 100, 127, 151, 172, 196, 223, 247, 298 and 316 m. Due to measurement failures throughout August 2022, the thermohygrometer data at 127 m has been excluded from the analysis. The 127, 196 and 316 m EC data have also been omitted due to poor data quality (for unknown reasons, this is discussed in appendix section 5.1).

From the radiosonde data, the lifting condensation level (LCL) has been inferred using the parcel method (Sun and Sun, 2019; De Feiter, 2023). This method is based on an air parcel being lifted dry adiabatically from the canopy top, where the LCL is the height at which the air parcel is saturated. For inferring the LCL, it has additionally been assumed that leaf temperatures at canopy top are 2 degrees Celsius larger than temperatures recorded by instruments of ATTO (Still et al., 2022). This has been assumed due to the air around leaves typically being warmer compared to the ambient air.

Additionally, data from a ceilometer is available that was collected at an altitude of 120 m on the CAMPINA site, conducting measurements in 5-minute intervals. Ceilometer data has been used to obtain the BL height, which was determined based on laser backscattering

(De Feiter, 2023). Finally, data from a pyranometer has been used, which is a measurement device that records shortwave and longwave radiation components. The pyranometer collected data at a height of 75 m.

The ten days of the CloudRoots-Amazon22 campaign (9 – 18 August 2022) have been classified in either the shallow cumulus or shallow-to-deep convective regime by De Feiter (2023). Most importantly, the shallow cumulus regime is characterized by an onset of shallow cumulus formation around 9 – 10 LT and more rapid cloud formation between ∼11 – 14 LT, with a typical cloud base of ∼1.5 km and cloud thickness between ∼0.5 – 2 km. The clouds start to dissipate around 14 LT, and by the end of the day (∼18 LT) most of the clouds have dissolved. A description of the shallow-to-deep convective regime is provided in De Feiter (2023).

During the CloudRoots-Amazon22 campaign, six out of ten days are characterized by the shallow cumulus regime. These six days have been aggregated, meaning that variables at individual times during the diurnal cycle have been averaged over the six shallow cumulus days (De Feiter, 2023). This aggregation allows for an identification of typical shallow cumulus conditions during the Amazonian dry season, rather than relying on only a single or couple of days that may be characterized by anomalous conditions. The shallow cumulus aggregate is essential for designing the numerical experiment, which is explained in section 2.3.

More detailed information on pre-processing of the observational data that has been performed for this thesis is provided in appendix section 5.1, which mostly consists of tower data analysis. Further details on pre-processing and available data of the CloudRoots-Amazon22 campaign data will be published in the future (de Feiter and de Haas, 2024, in preparation).

2.2 The Dutch Atmospheric Large Eddy Simulation

The Dutch Atmospheric Large Eddy Simulation (DALES) is a high-resolution model that has been extensively tested and validated (Heus et al., 2010; Ouwersloot et al., 2017). Similar to other LES models, it is able to resolve up to ∼90% of energy produced by turbulence, while the remaining fraction is parameterized. For turbulence, in addition to cloud formation and convection, DALES and other LES models are less reliant on parameterizations compared to models with coarser resolutions. In general, an LES is used to improve the understanding of BL processes and cloud formation, while aiding in providing parameterizations for weather and climate models.

In this research, DALES version 4.4 is used (in combination with observations) to investigate surface processes, turbulent properties and cloud formation for the shallow cumulus regime during the Amazonian dry season. We have chosen to employ DALES due to the importance of turbulence and cloud resolving resolutions in modelling forest-atmosphere interactions in the Amazon (Vilà-Guerau de Arellano et al., 2020). The initial conditions and various model schemes that have been implemented in DALES for this research are specified in tables in appendix section 5.6, in addition to figures of the initial vertical profiles (where homogeneous initial profiles throughout the domain are prescribed). The horizontal resolution adopted in DALES is 53 x 53 m, and the vertical resolution is 20 m.

The DALES domain spans 19 x 19 km horizontally and 5 km vertically, with no imposed advection or subsidence at the domain boundaries.

The model scheme that simulates CO_2 assimilation and stomatal aperture (and thereby also transpiration) in DALES is A-gs (Ronda, De Bruin, and Holtslag, 2001; Pedruzo-Bagazgoitia et al., 2017), where A is short for CO_2 assimilation and gs for leaf stomatal conductance. We have used A-gs with two big leafs, either sunlit or shaded, where the two big leafs are a simplified representation of all individual leafs in the rainforest. The stomatal aperture in the A-gs scheme responds to environmental conditions such as photosynthetically active radiation and vapor pressure deficit. Furthermore, a bulk canopy scheme has been used in DALES. This means that the canopy constitutes a single layer and so vertical variations in environmental conditions within the canopy are neglected, with the exception of radiation.

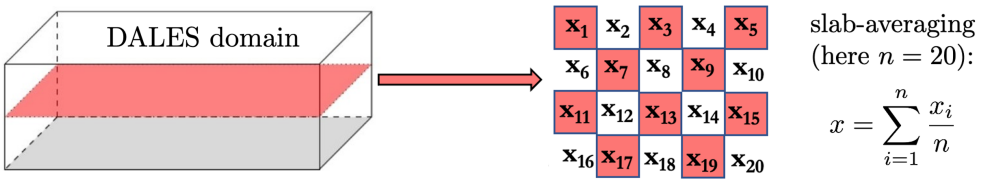

FIGURE 2.2: A schematic representation of slab-averaging in DALES, where a random variable x is averaged over a full horizontal plane within the DALES domain (where x_1, x_2, x_3 up to x_{20} are values of the variable x at the various grid boxes within the horizontal plane). In this example, the number of grid boxes (n) within a single horizontal plane is 20 (not representative of the DALES domain).

An important note is that we use the slab-averaged DALES output, which are horizontal averages over the full 2-dimensional planes in the DALES domain. This is depicted schematically in Fig. 2.2, which shows a random variable x being averaged over a horizontal plane of $n = 20$ grid boxes (this is not representative of the number of horizontal grid points in DALES). Slab-averaging has important implications for the results, for example, variability that may occur for certain variables at individual grid boxes are masked. The advantage of slab-averaging however is that it removes possibly large anomalies for a variable at individual locations within the DALES domain.

2.3 Research strategy

To enhance the understanding of the processes underlying Amazonian shallow cumulus formation during the dry season, a numerical experiment in DALES has been designed and systematically evaluated with the shallow cumulus aggregate from the CloudRoots-Amazon22 campaign data (as described in section 2.1). The observational surface, boundary layer and free tropospheric data from this campaign is essential for constraining the numerical experiment. By using this data in the form of the shallow cumulus aggregate, we have set up a case in DALES that is representative of shallow cumulus conditions, using the procedure described below. Designing the shallow cumulus DALES case (numerical experiment) is the main goal of this research, as this experiment can be used for a large

number of research applications such as the shallow to deep convection transition (see section 4.2).

The initial DALES set-up of this research is based on the DALES case of Vilà-Guerau de Arellano et al. (2020), which in turn is based on observations of the GoAmazon14/15 campaign during September 2014 (which is also during the Amazonian dry season, Martin et al., 2016). The initial conditions of DALES have been modified based on the shallow cumulus aggregate from the CloudRoots-Amazon22 campaign, which have also been used for evaluating the DALES case. This evaluation includes the following atmospheric variables and surface fluxes:

- The diurnal variability of the shortwave up/down and longwave up/down radiation, in addition to the net available radiation (Q_{net}), at an altitude of 75 m.
- The diurnal variability of latent heat, sensible heat and CO_2 flux (net ecosystem exchange) shortly above the canopy at ∼50 m, which is the most representative height above the canopy around ATTO.
- Vertical profiles of θ, q, u and v from the surface up to 5 km height at 6:00, 9:00, 12:00, 15:00 and 18:00 LT.
- The diurnal variability of θ, q, u, v and CO_2 at ∼300 m.
- The evolution of the BL height and LCL between 8:00 and 17:00 LT.
- The diurnal variability of turbulent kinetic energy and friction velocity at 100 and ∼300 m heights.

We have used the following procedure to design the DALES case: (i) the initial conditions in DALES are modified with the goal of improving the reproduction of the shallow cumulus aggregate for one or more variables, (ii) DALES is run with the new initial conditions, and (iii) DALES is evaluated using the above listed variables from the shallow cumulus aggregate. This procedure has been repeated until DALES was able to satisfactory reproduce the above variables, thereby being capable of simulating a representative shallow cumulus diurnal cycle. All the modifications that have been implemented between the DALES runs for designing the shallow cumulus case are listed in section 5.5. Examples of variables that have been modified for the DALES case are the initial vertical profiles of θ, q, u, v and CO_2 between 0 – 5 km and surface properties such as the albedo and soil temperature.

Chapter 3

Results & discussion

In this chapter, results from the analysis of both the CloudRoots-Amazon22 campaign data and DALES are presented and discussed. The first four sections are focused around answering and discussing the four research questions. In section 3.5, we discuss the performance of DALES by a statistical comparison to the observations. The chapter concludes with a discussion on the initial vertical profiles of potential temperature and specific humidity implemented in DALES, described in section 3.6.

For the figures presented in this chapter, the observational data shown in red dots are aggregated over six days characterized by the shallow cumulus regime (see section 2.1), where the red shaded area is one standard deviation of the aggregate (see e.g. Fig. 3.1). For DALES, the results are 5-minute intervals (unless otherwise stated) and slab-averaged (see section 2.2). The initial and boundary conditions prescribed in DALES, in addition to the adopted numerical schemes, are listed in appendix section 5.6. Furthermore, the canopy top in the Amazon fluctuates considerably, but is ∼40 m (Gomes Alves et al., 2023). The difference in height for the variables that are compared between DALES and the observations for all figures is then ∼40 m due to the bulk canopy scheme (see section 2.2), meaning the canopy top of the observations corresponds to the 0m-level in DALES. Finally, the roughly first two hours of the DALES run suffers from model spin-up. Therefore, the time between 6:00 LT at which DALES is initiated and ∼8:00 LT may show unphysical trends and/or large discrepancies with the shallow cumulus aggregate.

3.1 Diurnal variability of surface turbulent fluxes

This section addresses the following research question (1):

Do the surface turbulent fluxes of latent and sensible heat evolve symmetrically in time during the diurnal cycle?

To clarify this research question, we are interested in whether the diurnal variability curves of sensible and latent heat flux are symmetrical in time individually, and additionally if they evolve symmetrically in time with respect to each other. This has important implications for turbulence and cloud formation. For example, if the individual sensible heat flux curve is asymmetrical with a rapid rise during the morning while being larger than the latent heat flux, considerable turbulence is produced while the BL is relatively dry. Consequently,

cloud formation would likely be delayed or even inhibited (also depending on the evolution of both fluxes during the afternoon and other conditions) despite efficient vertical transport of surface air parcels by turbulent eddies.

In the left panel of Fig. 3.1, the evolution of net available radiation (Q_{net}) is shown. The observations of Q_{net} are in good agreement with the findings of Vilà-Guerau de Arellano et al. (2020). The Q_{net} is an important variable for the surface turbulent fluxes, as it connects the radiation balance to the surface energy balance. This is especially important in the Amazonia due to the dependence of photosynthesis and transpiration on radiation, and therefore needs to be properly reproduced by DALES. The figure shows that DALES performs well in reproducing Q_{net}, where it evolves similarly compared to the observations. DALES does not exhibit the considerable variability that is observed in Q_{net}, however, this is caused by the slab-averaging of DALES. As this averaging occurs over the full horizontal domain, the variability in Q_{net} at individual grid points that predominantly originates from cloud shading is masked.

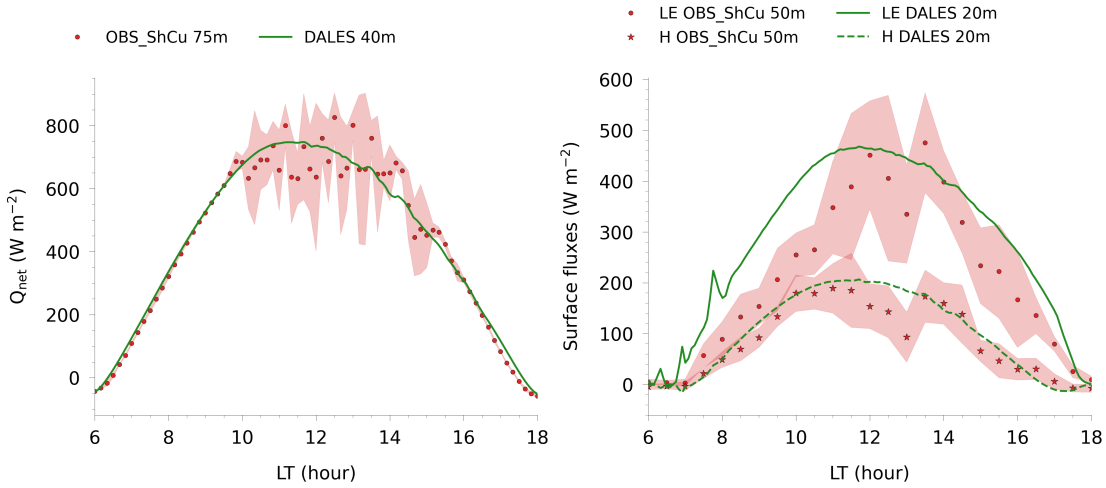

FIGURE 3.1: Diurnal cycle of net available radiation (Q_{net}, left panel), and sensible (H) and latent (LE) heat flux (right panel) for observational data taken during shallow cumulus days and DALES. The ∼40 m difference in height between DALES and the observations is due to the bulk canopy scheme, meaning the canopy top (which is ∼40 m) of the observations corresponds to the 0m-level in DALES. The observational data are 30-minute intervals, and DALES consists of 5-minute intervals. The observational data has been aggregated over six shallow cumulus days (OBS_ShCu), where the red shading represents one standard deviation of the aggregate.

In the right panel of Fig. 3.1, the diurnal variability of sensible (H) and latent heat flux (LE) are shown for the shallow cumulus observations and DALES. The latent heat flux rises more rapidly than the sensible heat flux during the early morning, especially for DALES (although the first two hours may be adversely affected by model spin-up as mentioned). The latent heat flux is especially large compared to the sensible heat flux during the late morning and afternoon for both DALES and the observations, caused by large amounts of evapotranspiration from the Amazonian rainforest. The observations of latent and sensible heat flux agree well with those from Vilà-Guerau de Arellano et al. (2020), who also measured these fluxes shortly above the canopy during the diurnal cycle. To answer the research question, both the sensible and latent heat flux show reasonably similar trends

during the early morning for the observations, during which they evolve roughly symmetrically with respect to each other (although the latent heat flux rises more steeply). This is not reflected in DALES, as the latent heat flux rises earlier and significantly more rapidly compared to the sensible heat flux. During the late morning and afternoon, the sensible and latent heat flux show deviating trends for both DALES and the observations, where the latent heat flux is characterised by considerably larger values. Therefore, we conclude that the sensible and latent heat flux do not evolve symmetrically in time with respect to each other. To answer the research question on whether the latent and sensible heat flux evolve symmetrically in time individually, both fluxes have been integrated in time between 6:00 – 12:00 LT and 12:00 – 18:00 LT, the results of which are provided in table 3.1. If the fluxes are symmetrical in time, the integrated values for the time span 6:00 – 12:00 LT and 12:00 – 18:00 LT should be similar. The time-integrated latent heat flux for the observations, in combination with its evolution shown in the right panel of Fig. 3.1, indicate that the latent heat evolves asymmetrically in time. This is however not reflected in DALES, which shows a roughly symmetrical trend and reasonably similar time-integrated values before and after 12:00 LT (albeit also displaying minor asymmetry due to the peak occuring before 12:00 LT and the decay being slower compared to the rise). The latent heat flux for DALES however shows large discrepancies with the observations, in addition to very different time-integrated values between 6:00 – 12:00 LT, likely suffering from the simplified bulk canopy scheme which is discussed in the next paragraph. Therefore, we conclude based on the observations that the latent heat flux does not evolve symmetrically in time. For sensible heat flux, the observations and DALES agree well for both the diurnal cycle shown in the right panel of Fig. 3.1 and the time-integrated values provided in table 3.1, indicating that the sensible heat flux does evolve symmetrically in time. These results indicate a considerable amount of evapotranspiration, especially during the late morning and early afternoon, due to the large latent heat flux values. This may lead to the formation of a moist BL with possibly favorable conditions for cloud formation, which depends on other factors such as turbulence production and is further discussed in the next sections.

	LE (W m^{-2} h)		H (W m^{-2} h)	
	OBS_ShCu	DALES	OBS_ShCu	DALES
6:00 - 12:00 LT	854.8493	1578.0473	497.9068	636.7611
12:00 - 18:00 LT	1503.8435	1790.5676	479.8998	551.2667

TABLE 3.1: The latent and sensible heat flux integrated over time between 6:00 – 12:00 local time (LT) and 12:00 – 18:00 LT for the observations and DALES (see Fig. 3.1). The observational data has been aggregated over six shallow cumulus days (OBS_ShCu).

As mentioned, the latent heat flux of DALES deviates significantly from the observations, especially during the morning. This may be caused by several factors mostly related to the bulk canopy scheme in DALES, due to the importance of capturing evapotranspiration in the Amazonian rainforest. Firstly, as a bulk canopy scheme is implemented in DALES, vertical variations of in-canopy conditions are neglected. In Pedruzo-Bagazgoitia et al. (2023), it is shown that a multi-layer canopy indeed performs better in simulating latent heat flux, which is further discussed in section 3.2. Additionally, a mechanistic two big-leaf model (either sunlit or shaded) for photosynthesis and stomatal aperture is used (see section 2.2). These are both simplifications of the complex Amazonian rainforest, which is likely to

cause the discrepancy in latent heat flux through a misrepresentation of evapotranspiration in DALES. Another likely contributor to the model-observation deficit is dew, which is an important factor in Amazonian evaporation during the morning and difficult to properly represent in the canopy scheme. This may be big factor in the misrepresentation of latent heat flux for DALES during the morning, but not for the afternoon as the dew has then evaporated and so does not play a role (Zhang et al., 2023). However, considering the complexity of the rainforest and the difficulty in modelling canopy processes that determine the evapotranspiration, we note that the diurnal variability of latent heat flux in DALES is considered satisfactory.

3.2 Diurnal variability of turbulent properties

This section discusses the following research question (2):

2) What is the diurnal variability of turbulent kinetic energy and how is this energy distributed between the vertical and horizontal directions?

The diurnal variability of latent and sensible heat flux are an important contributor in determining turbulent properties in the BL, and these fluxes in combination with turbulent transport of moist air parcels from the surface are in turn essential for the initiation of shallow cumulus formation (which is further discussed in section 3.4). In Fig. 3.2, the diurnal variability of turbulent kinetic energy (TKE) at 100 m (upper left panel) and 298 m (upper right panel) are shown for DALES and the observations. First focusing on TKE at 298 m, we find that DALES performs well in reproducing the observations. At ∼8:00 LT, the TKE rises rapidly for both DALES and the observations, indicating an early production of turbulence at 298 m that may contribute to early development of shallow cumulus clouds (which is further discussed in section 3.4).

At 100 m however, DALES underestimates the TKE for the full diurnal cycle. Turbulence, and thereby TKE, is produced by a combination of mechanical and convective turbulence, where mechanical turbulence is generated by a combination of wind shear and surface friction. Mechanical turbulence in the roughness sub-layer, defined as ∼2-3 times the canopy top height (so this layer would be ∼80-120 m for our focus area in the Amazon), is generated by a combination of wind shear from the logarithmic deceleration in wind towards the surface and atmospheric flow disturbance by the canopy. A good measure of mechanical turbulence in the roughness sub-layer is the friction velocity u_*, which is given by

$$u_*^2 = \sqrt{(\overline{w'u'})^2 + (\overline{w'v'})^2}, \tag{3.1}$$

where $u' = u(t) - \overline{u}$ is the deviation of u from the mean and $\overline{w'u'}$ and $\overline{w'v'}$ are the turbulent momentum fluxes in the zonal and meridional direction, respectively (Vilà-Guerau de Arellano et al., 2015). The diurnal variability of friction velocity is shown in Fig. 3.2 (middle left panel) at 100 m. At an altitude of 100 m, u_* is a good measure of mechanical turbulence as this height is (roughly) within the roughness sub-layer. As DALES underestimates both the TKE and u_* at 100 m, the TKE may be too low for DALES due to

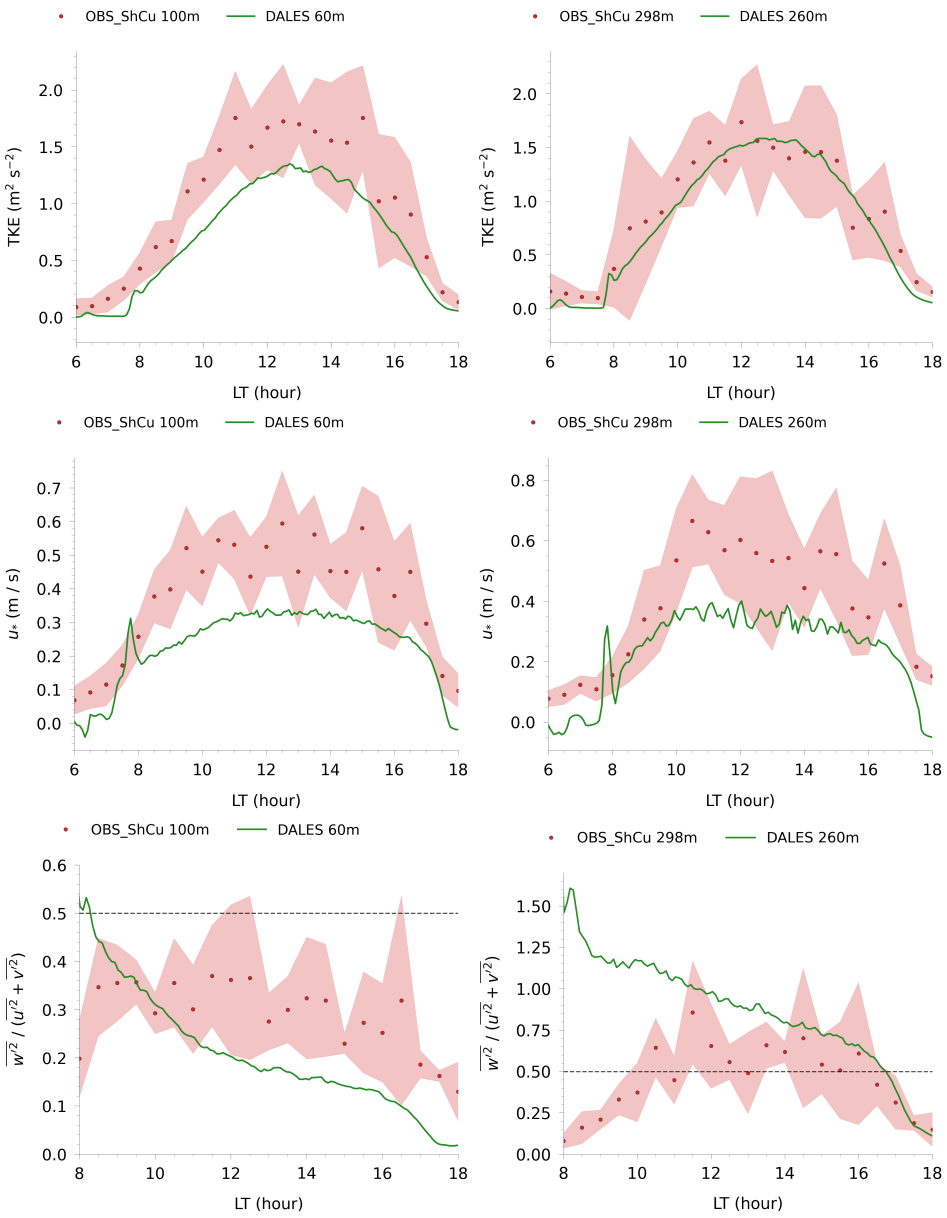

FIGURE 3.2: Diurnal variability of turbulent kinetic energy (TKE), friction velocity (u_*) and fraction of vertical over horizontal wind variance at 100 m (left panels) and 298 m (right panels) for observational data taken during shallow cumulus days and DALES. The ∼40 m difference in height between DALES and the observations is due to the bulk canopy scheme, meaning the canopy top (which is ∼40 m) of the observations corresponds to the 0m-level in DALES. Note that the graph of α_t starts at 8:00 LT, which is due to unphysical trends related to model spin-up between 6:00 and 8:00 LT. The observational data are 30-minute intervals, and DALES consists of 5-minute intervals. The observational data has been aggregated over six shallow cumulus days (OBS_ShCu), where the red shading represents one standard deviation of the aggregate.

misrepresenting u_*. To conclusively determine this would however require an analysis of the contributions of mechanical and convective turbulence to TKE, which is beyond the scope of this thesis.

Similar to the latent heat flux, an important factor that that limits the performance of DALES in reproducing turbulent properties at heights near the surface is the bulk canopy scheme. In Pedruzo-Bagazgoitia et al. (2023), the performance of a bulk and multi-layer canopy representation in DALES have been compared with observations, highlighting that the multi-layer canopy scheme performs significantly better in simulating the diurnal variability of friction velocity, in addition to latent and sensible heat flux (although to a lesser degree, see their figure 8). Furthermore, they conclude that the bulk canopy implementation in DALES is unable to reproduce the turbulence induced by the canopy, thereby missing important canopy-atmosphere features, which likely contributes to the discrepancy between DALES and the observations that we find for u_* and TKE at 100 m.

The friction velocity is also shown at 298 m in Fig. 3.2 (middle right panel). Similar to u_* at 100 m, DALES underestimates the friction velocity. However, as an altitude of 298 m is not within the roughness sub-layer, u_* is not a good measure of mechanical turbulence due to contributions of vertical wind shear. As the TKE is reproduced well by DALES at 298 m while u_* is not, the vertical wind shear and/or convective turbulence is likely also misrepresented in DALES. Importantly however, the TKE is reproduced well at 298 m and performs better with increasing height (as also shown by statistical analysis, see section 3.5). This means that DALES can simulate turbulent properties better when further away from the surface, which is likely caused by the simplified surface scheme. The proper reproduction of TKE at higher altitudes provides confidence in the ability of DALES to simulate shallow cumulus development, due to the importance of turbulent transport of moist air parcels in initiating cloud formation. Finally, we note that u_* at 100 and 298 m show similar values for both the observations and DALES, which is unusual as u_* typically decreases with height, which we are unable to explain.

In Fig. 3.2, the fraction of vertical over horizontal wind variance is shown, again at 100 m (lower left panel) and 298 m (lower right panel). Note that this figures starts at 8:00 LT, as DALES model spin-up causes unrealistic features in this fraction between 6:00 and 8:00 LT. We name this fraction α_t and it is defined as

$$\alpha_\mathrm{t} = \frac{\overline{w'^2}}{\overline{u'^2} + \overline{v'^2}}, \tag{3.2}$$

where $\overline{w'^2}$ is the vertical wind variance and $\overline{u'^2} + \overline{v'^2}$ the horizontal wind variance (Pino et al., 2006). As the TKE is determined by

$$\mathrm{TKE} = \overline{u'^2} + \overline{v'^2} + \overline{w'^2}, \tag{3.3}$$

the fraction α_t represents the anisotropy in turbulent energy of the vertical over the horizontal directions. When $\alpha_\mathrm{t} < 0.5$, the turbulent flow is primarily anisotropic in the horizontal directions, while for $\alpha_\mathrm{t} > 0.5$ the turbulence is predominantly anisotropic in the vertical direction. Therefore, the horizontal dashed lines in the figures indicate when $\alpha_\mathrm{t} = 0.5$. We first only consider the observations of α_t for interpreting the results. At 100 m, the α_t is below 0.5 between 8:00 and 18:00 LT, indicating that the turbulent kinetic energy is primarily generated by wind fluctuations in the horizontal directions. This is likely caused by the proximity to the surface, where wind deceleration towards the surface and surface

roughness disrupt the horizontal atmospheric flow, thereby producing wind fluctuations primarily in the horizontal directions (so larger $\overline{u'^2} + \overline{v'^2}$). At 298 m, the observations of α_t are largely above 0.5 and thereby anisotropic in the vertical direction, which is likely partially caused by a reduced disruption of the horizontal flow at this height. Furthermore, the observations indicate that α_t starts to rise later at 298 m compared to 100 m. This could be caused by the delay in buoyancy reaching higher altitudes in the BL, which is an important contributor to vertical wind variance. As buoyancy decreases towards the end of the afternoon, and thereby also the vertical wind variance, the α_t declines at the end of the diurnal cycle for both 100 and 298 m.

Comparing α_t between the observations and DALES at both 100 and 298 m during the early morning, α_t is relatively large for DALES, especially at 298 m. This is presumably related to the unstable (well-mixed potential temperature) profile that is used to initialize DALES at 6:00 LT. This causes an early onset of convection in DALES that may drive up $\overline{w'^2}$, whereas in reality the BL is stable at 6:00 LT (a discussion on the initial vertical profiles implemented in DALES is provided in section 3.6). This likely causes the overestimation of α_t in DALES during part of the morning, especially at 298 m due to the significantly earlier onset of convection in DALES at this height compared to the observations. At 100 m, DALES likely shows poor agreement with the observations for α_t due to the bulk canopy scheme in DALES. We expected the α_t to show improved performance at 298 m compared to 100 m when excluding the morning as discussed, as this height is located outside of the roughness sub-layer (and thereby the surface should have a relatively small influence on turbulent properties). However, α_t at 298 m is too large for DALES throughout most of the diurnal cycle, also when excluding the morning. This indicates that while DALES is able to reproduce the TKE at 298 m, it is not able to capture the anisotropy in turbulent energy between the vertical and horizontal directions. During the afternoon, this could be related to the large discrepancy between vertical profiles of u and v between DALES and the observations (see Fig. 5.10). This discrepancy in wind profiles may also contribute to the deficit between DALES and the observations for α_t at 100 m, in addition TKE at 100 m and u_* at both heights. Additionally, the domain size in DALES may be too small for properly capturing wind. In Sikma et al. (2018), DALES with a domain of 48 x 48 km can properly capture the wind structure (see their figure 2), however, we use a 19 x 19 km domain which could be insufficient. If the wind patterns are not properly captured, the vertical wind shear may be misrepresented, which may contribute to the discrepancies of turbulent properties in Fig. 3.2. Finally, we note that Siebesma et al. (2003) who simulate classical shallow cumulus convection in the trade wind with 10 different LES models also find that the TKE at ∼300 m has a dominant contribution from the vertical wind variance (see their figure 5). However, at 100 m their findings are not in line with the DALES α_t of Fig. 3.2, but this is likely related to the different surface type in their LES models (ocean instead of rainforest).

3.3 Turbulent mixing of potential temperature and specific humidity

This section discusses the following research question (3):

Does turbulent transport mix potential temperature and specific humidity with the same efficiency?

As discussed in the previous section, the turbulent properties outside of the roughness sub-layer are simulated reasonably well by DALES, although issues related to friction velocity and anisotropy in turbulent energy remain. In this section, we discuss turbulent mixing of potential temperature and specific humidity within the BL for both DALES and the observations, which provides an indication on whether these issues are propagated to turbulent transport. To address the first part of the above research question, to what extent the state variables become well-mixed, the evolution of potential temperature (θ) and specific humidity (q) from the observational tower data at six different heights are shown in Fig. 3.3 for a full 24-hour cycle. Note that this is data from a single shallow cumulus day (2022-08-10), so not the aggregate. In the panels, sunrise and sunset are marked with a black dashed line, while the onset of the convective BL and rapid growth stage as identified by Henkes et al. (2021) are indicated in blue and yellow, respectively (see chapter 1). The six plotted heights are all above 100 m, as this is the approximate altitude at which the BL becomes well-mixed for θ. The left panel of Fig. 3.3 clearly shows that from ∼9 LT and for heights above ∼100 m, the difference in potential temperature between the various heights is extremely small, indicating well-mixed conditions. However, for specific humidity, the boundary layer is not or only weakly mixed throughout the full 24-hour cycle, as shown in the right panel of Fig. 3.3. Additionally, similar specific humidity figures for different days (both shallow and deep convective, not shown here) show that the degree of mixing for q varies significantly per day. Furthermore, for the u and v wind components, it also depends substantially per day to what extent the variables mix within the BL (see appendix Fig. 5.3 and Fig. 5.4).

As the potential temperature is well-mixed, the mixed layer θ of DALES is compared with the mean θ of the tower data between 100 and 298 m in the left panel of Fig. 3.4. The mixed layer value of a random variable ψ, denoted as $\langle\psi\rangle$, is defined as

$$\langle\psi\rangle = \frac{1}{h}\int_{z_0}^{h}\psi(z)dz \tag{3.4}$$

where h is the BL height (determined here as the height at which the buoyancy flux is minimum), z is the vertical component of the cartesian grid and z_0 is the reference height of the variable at or near the surface (Vilà-Guerau de Arellano et al., 2015). For the observations, the mean θ of the tower data between 100 and 298 m is taken as an approximation of the mixed layer value (the radiosonde data is not included due to a discrepancy between sonde and tower data, see Fig. 3.5). For DALES, the mixed layer value is only plotted between 8 and 16 LT, as the BL height of DALES that is inferred from the minimum buoyancy flux is not well-defined outside this time range. The left

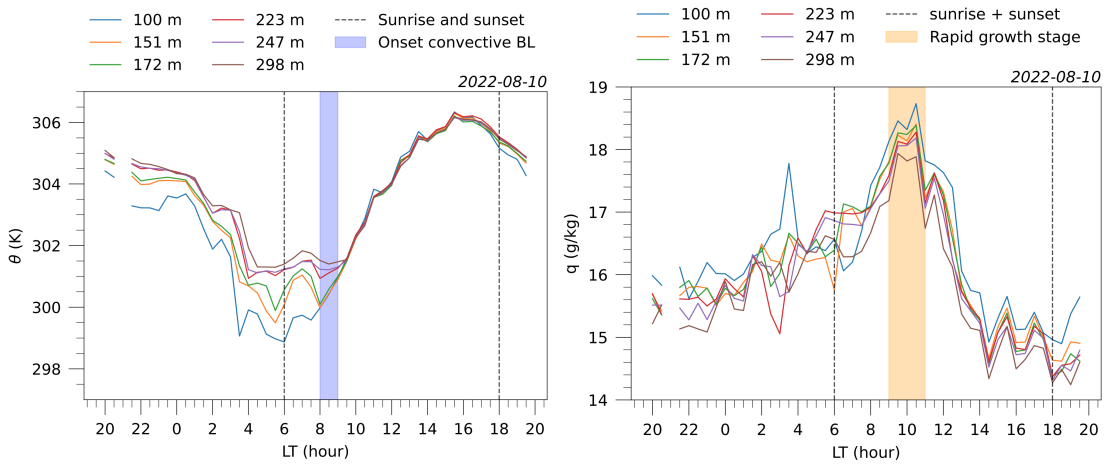

FIGURE 3.3: The 24-hour evolution of potential temperature (θ, left panel) and specific humidity (q, right panel) for tower data from six different heights of a single shallow cumulus day (10-08-2022). The onset of the convective BL and rapid growth stage that are indicated in the figure are based on Henkes et al. (2021), which is discussed in chapter 1.

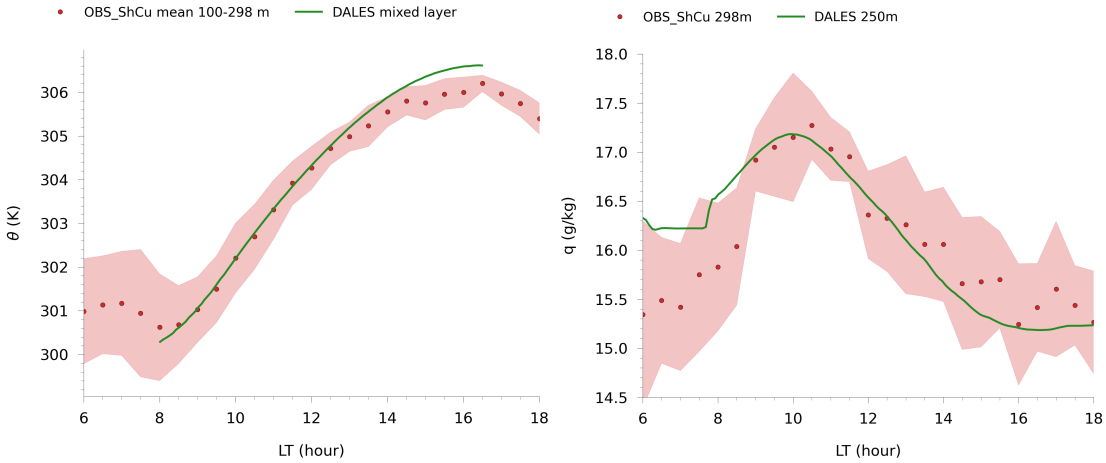

FIGURE 3.4: **Left panel:** Evolution of the mixed layer potential temperature (θ) of DALES and the mean θ of tower data between 100 and 298 m (above the roughness sub-layer, which is ∼2–3 times the canopy height), taken as an approximation of the observational mixed layer value. The tower data is an aggregate of six shallow cumulus days (OBS_ShCu), where the red shading represents one standard deviation of the aggregate. The observational data has been 30-minute averaged (over thermohygrometer data consisting of 1-minute intervals), while DALES consists of 5-minute intervals (also for the right panel). **Right panel:** Diurnal variability of specific humidity (q) for the observations and DALES at 298 m. The ∼40 m difference in height between DALES and the observations is due to the bulk canopy scheme, meaning the canopy top of the observations corresponds to the 0m-level in DALES.

panel of Fig. 3.4 shows excellent agreement for the mixed-layer θ between DALES and the observations, indicating that the remaining issues in simulating turbulent properties do not adversely affect the BL mixing of θ in DALES (which is also reflected in vertical profiles of θ as shown in the upper panels of Fig. 3.5). Furthermore, the left panel of Fig. 3.4 shows that θ starts to rise around 8 LT for both DALES and the observations due to

entrainment of free tropospheric air and heating of the surface after sunrise, showing a peak value around 16:30 LT.

The right panel of Fig. 3.4 displays a comparison of specific humidity between DALES and the observations at 298 m. The mixed-layer value is not compared for this variable, as q only shows weak mixing at best. The q rises steeply in the early morning due to evapotranspiration from the surface, while dropping around 10 LT due to the entrainment of dry air from the free troposphere. The timing of the q peak then represents the onset of entrainment dominance over evapotranspiration from the surface, as these are the two main drivers of the evolution of specific humidity. Additionally, we note that from ∼9 LT, after model spin-up, DALES and the observations show good agreement, although after ∼13 LT DALES slightly underestimates q. The reproduction of the diurnal variability of q in DALES is however satisfactory, especially considering the complexity in simulating evapotranspiration from the rainforest as discussed in section 3.1.

In Fig. 3.5, vertical profiles of potential temperature (upper panels) and specific humidity (lower panels) at 9:00, 12:00 and 15:00 LT are shown for the observations and DALES. The vertical profiles are shown up to an altitude of 2500 m, as the degree of BL mixing is then more clearly visible compared to figures of the profiles up to 5000 m (the full vertical domain). The red points are the shallow cumulus aggregate as before, while the grey profiles represent the individual profiles within the aggregate. For the observations, 40 m has been subtracted from the height (as the canopy top of the observations corresponds to the 0m-level in DALES due to the bulk canopy scheme, as mentioned), meaning the in-canopy observations are excluded. Furthermore, the lower 260 m consists of tower data, while the observations above 260 m are from radiosondes (which causes the discontinuity at this height for the observational profiles). For DALES, the vertical profiles are 30-minute averaged. The BL height, inferred from a ceilometer, is shown with a dashed black line. When examining the profiles of θ below the BL height, only minimal variation with altitude is shown for both DALES and the observations, except for radiosonde data at 9:00 LT. This is in line with Fig. 3.3, as at ∼9 LT the BL only just starts to become well-mixed for the θ tower data, meaning the radiosonde profile of θ is still height-dependent due to delayed deepening of the convective BL at this time. The vertical profiles of θ at 12:00 and 15:00 LT show excellent mixing throughout the full BL, indicating the formation of a deep convective BL, similar to the findings of Henkes et al. (2021). For the vertical profiles of q, the observations indicate weak BL mixing, contrary to DALES which shows almost constant q with altitude within the BL. This could be rooted in the misrepresentation of some of the turbulent properties as discussed in section 3.2, however, the BL mixing of θ for DALES and the observations do show good agreement as mentioned. Therefore, we hypothesize that the disagreement in turbulent mixing between DALES and the observations for q is likely caused by the absence of moisture advection in DALES. Additionally, we note that the weak mixing of q for the observations is not in line with Henkes et al. (2021), who do infer well-mixed BL conditions for q in the shallow cumulus regime based on radiosonde data.

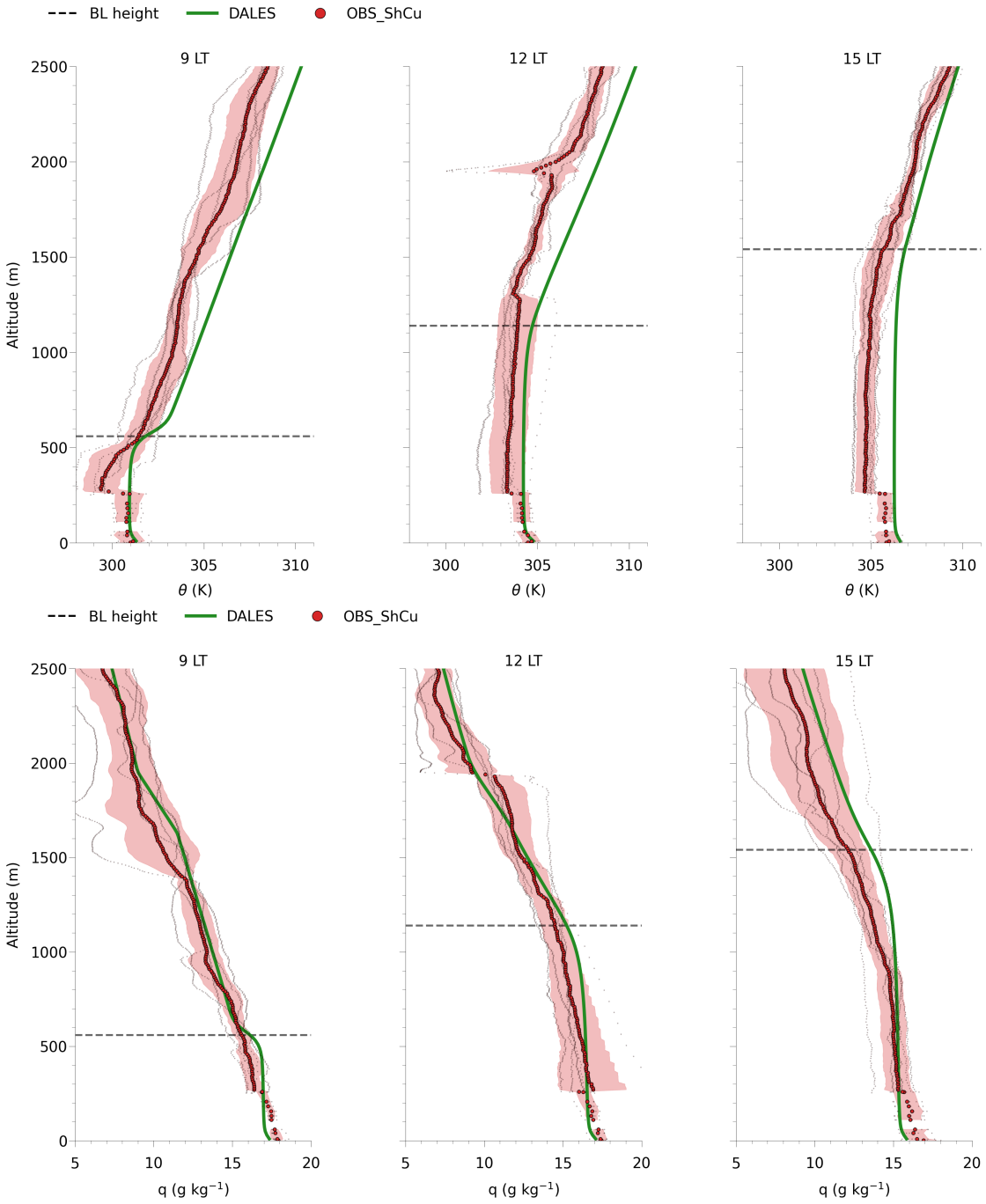

FIGURE 3.5: Vertical profiles of DALES and the observations up to 2500 m of potential temperature (upper panels) and specific humidity (lower panels) at 9:00, 12:00 and 15:00 local time (LT). The lower 260 m of the observational profiles consists of tower data (30-minute intervals), while above 260 m the data is from radiosondes. The observational data shown in red is an aggregate of six shallow cumulus days (OBS_ShCu), where the red shading represents one standard deviation of the aggregate. The grey points are the individual profiles from the aggregate. The BL heights are inferred from a ceilometer. The DALES profiles are 30-minute averaged (over six 5-minute intervals).

3.4 Transition from clear to cloudy conditions driven by moist convection

This section addresses the following research question (4):

When is the transition from clear to cloudy conditions as characterized by the transition from dry turbulence to moist convection?

The transition from dry turbulence to moist convection, and whether and to what degree this leads to shallow cumulus development, depends on the vertical profiles of potential temperature and specific humidity and thereby on the degree of turbulent mixing of these variables, as discussed in the previous section. We define the transition from dry turbulence to moist convection as the first time during the diurnal cycle at which the LCL is lower than the BL height at some location of the considered region. Therefore, this transition only occurs locally, as the LCL in the shallow cumulus regime is generally expected to remain above the BL height in regions without cloud development. Conversely, in areas where shallow cumulus clouds form, the LCL is expected to be below the BL height. For the results, it is important to keep in mind that the transition from dry turbulence to moist convection is therefore regarded as the timing when the LCL in some subregion is first located below the BL height, even if the LCL is above the BL height everywhere else.

As the transition from dry turbulence to moist convection is defined by a local feature, it poses a problem for determining it from observational data. Namely, as the LCL is inferred from radiosondes and the BL height from a ceilometer, they represent values at single trajectories within a large region. For determining the time at which dry turbulence transitions to moist convection, data of the BL height and LCL throughout the full considered region is needed. Therefore, the observational data are insufficient for answering this research question, and we require an atmospheric model as the LCL and BL height can be determined throughout the full model domain. An LES is then the ideal tool to answer this research question due to its capacity to largely resolve turbulence and cloud formation, which are crucial for accurately simulating the BL height and LCL.

In the left panel of Fig. 3.6, the evolution of the BL height and LCL from DALES and the observations are shown. Despite being unsuitable for inferring the timing of the transition from dry turbulence to moist convection, the observations are included for determining how well DALES is able to reproduce the BL height and LCL. As indicated in the figure legend, the LCL for the observations has been determined with the parcel method (see section 2.1), while for DALES it has been inferred based on the height at which the liquid water specific humidity, q_l, first becomes larger than 0. Consequently, even if only a single grid box within a horizontal slice of the DALES domain has a $q_l > 0$, this is defined as the LCL (if the horizontal slice is vertically the first one with $q_l > 0$). The figure shows a large discrepancy in LCL between DALES and the observations, especially at 9:00 LT. This is however likely caused by the LCL from DALES being based on the shallow cumulus cloud with the lowest LCL by definition, while for the observations it is dependent on the trajectory of the radiosondes in the shallow cumulus aggregate and how many of them pass through clouds (as these are characterised by lower LCLs compared to areas without cloud formation). For the BL height, the observations are inferred from a ceilometer

based on laser backscattering, while for DALES it is determined based on the minimum buoyancy flux. The figure shows that the BL height in DALES is in good agreement with the observations, falling within the standard deviation of the shallow cumulus aggregate throughout the displayed evolution.

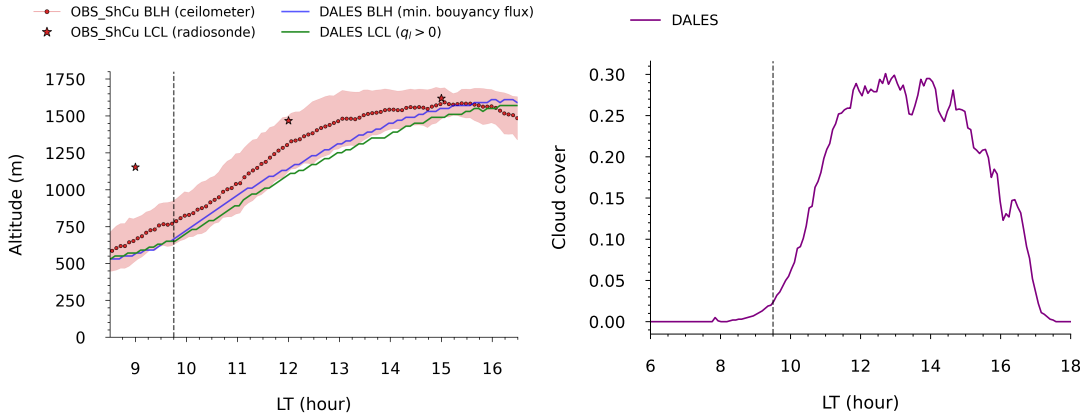

FIGURE 3.6: **Left panel:** Evolution of the BL height (BLH) and LCL from DALES and the observations. The BL height from the observations is determined using a ceilometer, while the one from DALES is inferred using the minimum buoyancy flux. The LCL from DALES is the first height at which the liquid water specific humidity is larger than 0, while the observational LCL is inferred from radiosondes using the parcel method (with +2 K at canopy top, see section 2.1). The observational data is an aggregate of six-day shallow cumulus days (OBS_ShCu), where the red shading represents one standard deviation of the aggregate. The black dashed line indicates the time at which the LCL first drops below the BL height in DALES. The observational data of BL height and DALES consist of 5-minute intervals, while the LCL from observations are 3-hour intervals. **Right panel:** Cloud cover evolution in DALES. The black dashed line again represents the time at which the LCL first drops below the BL height in DALES.

The black dashed line in Fig. 3.6 represents the timing of the transition from dry turbulence to moist convection in DALES, which occurs at 9:45 LT. Albeit being defined by a local feature, the transition from dry turbulence to moist convection marks an important turning point. Since the LCL of DALES is defined as locally the first height where $q_l > 0$, it indicates the formation of a deep-convective BL that is able to vertically transport moist air parcels from the surface to the top of the BL. The transition additionally represents the time at which shallow cumulus clouds start to develop more rapidly within the DALES domain. This is shown in the right panel of Fig. 3.6 that displays the cloud cover evolution in DALES, where the cloud cover rises steeply after the transition to moist convection (again indicated by the black dashed line). This figure additionally indicates that the shallow cumulus clouds start to dissipate around 15:00 LT, and are largely dissolved at 17:00 LT.

As discussed, DALES simulates well-mixed q conditions that are not reflected in the observations. Additionally, the vertical profile of θ at 9:00 LT, shortly before the transition to moist convection, shows a deep convective BL for DALES which is not reflected in the observations. As these features indicate earlier transport of moist air parcels to the top of the BL in DALES compared to the observations, they should lead to an earlier transition to moist convection and thereby shallow cumulus formation in DALES. Therefore, we note

that the actual transition from dry turbulence to moist convection is likely to be later than 9:45 LT. On the other hand, De Feiter (2023) finds that cloud formation in the shallow cumulus regime already starts to develop between 9 – 10 LT.

3.5 Model-observation statistics

In this subsection, the performance of DALES is compared to the observations using the index of agreement, d, and root mean square error (RMSE) statistics. The RMSE reflects the mean difference between the model and observations with the same unit as the corresponding variable, and the index of agreement d is a bounded and dimensionless measure that is suitable for cross-comparing different models (Willmott, 1982). The index of agreement is determined by

$$d = 1 - \left(\sum_{i=1}^{N} (P_i - O_i)^2 \Big/ \sum_{i=1}^{N} (|P_i'| - |O_i'|)^2 \right), \tag{3.5}$$

where N is the number of data points, P_i is the model data at time i, O_i is the observational data, $P_i' = P_i - \bar{O}$ where $\bar{O}$ is the mean of the observations and $O_i' = O_i - \bar{O}$ (Willmott, 1982). The index of agreement is always between 0 and 1, where a value of 1 reflects perfect agreement between model and observations. For the RMSE, a lower value indicates better model performance.

The index of agreement and RMSE for both DALES and the observations are listed in table 3.2 for θ, q, H and TKE at five different heights. The latent heat flux is not included in the table, as reliable data from this variable is only available at 50 and 81 m (see appendix section 5.1). For both DALES and the observations, the RMSE and d are calculated over the diurnal cycle with 30-minute intervals, however excluding the first two hours (between 6:00 and 8:00 LT) due to model spin-up.

The statistics indicate that DALES performs well in simulating all the selected variables. This is most evident from the index of agreement, which is < 0.8 for only a single instance (0.694 for TKE at 43 m). The RMSE generally decreases with height for q and TKE, indicating that DALES performs better at higher altitudes for these variables, which is also reflected in the index of agreement. Furthermore, for q and TKE, the RMSE and index of agreement indicate that DALES performs considerably better when outside of the roughness-sublayer. These results indicate that, similar to Pedruzo-Bagazgoitia et al. (2023), the bulk canopy representation in DALES is insufficient for properly simulating forest-atmosphere interactions in the Amazonia, instead requiring a multi-layer canopy scheme (which is a topic of future research). Only for the sensible heat flux do the statistics indicate better model performance when closer to the surface. The potential temperature statistics do not show a clear trend with altitude. However, for θ at all five heights, the index of agreement is very close to 1 and the RMSE is low, indicating an excellent reproduction of θ in DALES.

Height (m)		θ (K)		q (g/kg)		H (W m^{-2})		TKE (m^2 s^{-2})	
OBS	DALES	RMSE	d	RMSE	d	RMSE	d	RMSE	d
43	40	0.507	0.985	0.650	0.811	27.8	0.954	0.581	0.694
100	100	0.333	0.994	0.702	0.809	25.8	0.965	0.396	0.835
151	160	0.317	0.994	0.503	0.884	34.4	0.931	0.292	0.898
223	220	0.313	0.994	0.465	0.896	35.9	0.918	0.192	0.954
298	300	0.463	0.986	0.320	0.945	33.4	0.908	0.192	0.955

TABLE 3.2: Root mean square error (RMSE) and index of agreement (d) between DALES and the observations (OBS) for potential temperature (θ), specific humidity (q), sensible heat flux (H) and turbulent kinetic energy (TKE) at five different heights. Equation 3.5 has been used to determine d. The RMSE and d are determined based on 30-minute interval data of DALES and the observations during the diurnal cycle, however excluding the first two hours (between 6 LT and 8 LT) due to model spin-up. The statistics have been determined for five heights of the ATTO tower (43, 100, 151, 223 and 298m) and the closest corresponding DALES height is selected for calculating the statistics (40, 100, 160, 220 and 300m). Note that for each of these heights the actual selected DALES height is the one that's 40 m lower due to the bulk canopy scheme of DALES.

3.6 Discussion on the DALES initial vertical profiles

The initial vertical profile of θ that has been implemented in DALES does not resemble the observational profile based on radiosondes. This is contrary to the initial vertical profile of specific humidity, which does closely follow the observations. The initial θ profile is not based on the observations for two reasons. The first and most important one is cloud development, which showed unphysical features when using the observational θ profile to initiate DALES. This was most pronounced in the cloud cover evolution, which rapidly increased from ∼35% to ∼100% in a time span of $\lesssim$ 15 minutes at ∼13:30 LT. Additionally, the liquid water path was unrealistically large, reaching values of up to 250 g m^{-2} at the end of the afternoon. These unphysical characteristics did not originate from the initial specific humidity profile, as reducing the q throughout the vertical domain did not resolve the issue. Instead, modifying the initial θ profile to reduced values prevented the unphysical features. Furthermore, we chose to initialize DALES with an idealized θ profile with stable conditions in the lower ∼500 m, an inversion jump of ∼2.5 K and a constant θ lapse rate profile between ∼500 – 5000 m (see Fig. 5.12). The argumentation behind this profile is explained in the next paragraph, however, we would like to stress that the cloud cover and liquid water path issues were very sensitive to then initial vertical profile of θ. When reducing the θ values in this initial profile, the issues in many cases returned (see appendix section 5.5).

The idealized θ profile (see Fig. 5.12) has been chosen due to the timing of the onset of entrainment into the BL from the free troposphere. This timing is essential to the evolution of the BL height, which was not well simulated by DALES when implementing an initial stable θ profile in the lower atmosphere. When modifying this to an unstable profile, the diurnal cycle of BL height in DALES was in improved agreement with the observations, especially during the morning. Additionally, the diurnal variability of q was better reproduced by DALES when adopting an initial unstable profile. Similar to the evolution of BL height, this is rooted in the onset of entrainment of dry air from the free troposphere. This largely dictates the diurnal cycle of q due to the significantly

lower moisture content in the free troposphere compared to the BL, driving down the q around 10:30 LT (see Fig. 3.4). When implementing an initial stable θ profile, the onset of entrainment in DALES occurred too early, resulting in a large discrepancy in q between DALES and the observations. Finally, we note that the initial profiles of the wind components u and v also do not resemble the observational data from radiosondes, instead consisting of an idealized profile similar to θ (see appendix Fig. 5.13). This is further discussed in appendix section 5.3.

Chapter 4

Conclusions and future research

4.1 Conclusions

In this research, the diurnal cycle of surface processes, turbulent characteristics and cloud formation in the shallow cumulus regime have been investigated. The research has been conducted by combining a unique and comprehensive set of observational data from the CloudRoots-Amazon22 campaign (Vilà-Guerau de Arellano et al., 2024) with the Dutch Atmospheric Large-Eddy Simulation (DALES), which is a high-resolution model capable of largely resolving turbulence and cloud formation. Guided by the observations from the campaign, we design a numerical experiment in DALES that is representative of a typical shallow cumulus diurnal cycle.

When comparing DALES to the observations, DALES performs satisfactory in reproducing the diurnal cycle of net available radiation, sensible heat flux and latent heat flux shortly above the canopy, despite an overestimation in DALES of the latter. This is likely related to the simplified bulk canopy scheme that has been adopted in DALES, whereas the processes underlying evapotranspiration from the canopy are complex and likely require a more detailed canopy representation (Pedruzo-Bagazgoitia et al., 2023, see also section 3.1 and 3.2). DALES performs less well in simulating turbulent properties in the BL, especially close to the surface, but is able to reproduce turbulent kinetic energy outside the roughness sub-layer (see section 3.2 and section 3.5). Furthermore, DALES is able to capture the diurnal variability of potential temperature and specific humidity in the BL (see section 3.3 and section 3.5). Finally, DALES is able to largely reproduce the diurnal cycle of BL height, albeit not being capable of reproducing the LCL (see section 3.4). The large discrepancy in LCL between DALES and the observations is however likely caused by different methodologies in inferring the LCL. Finally, statistics between DALES and the observations for potential temperature, specific humidity, turbulent kinetic energy and sensible heat flux at five tower heights show that DALES performs well in reproducing the observations (see section 3.5).

Using both DALES and the observations, we have found that the sensible heat flux shows a roughly symmetrical trend during the diurnal cycle, whereas latent heat flux evolves asymmetrically (see section 3.1). These fluxes do not show symmetrical evolutions with respect to each other, as the latent heat flux is significantly larger, indicating the formation of a moist BL. Furthermore, the turbulent kinetic energy and friction velocity show similar evolutions at 100 and 300 m, where it is especially unexpected that the friction velocity at these heights show similar values (see section 3.2). The turbulent energy is anisotropic

in the horizontal direction at 100 m and anisotropic in the vertical direction at 298 m, which is likely largely related to surface roughness disrupting the horizontal flow at 100 m. Additionally, turbulence is able to mix the potential temperature efficiently in the BL from ∼9:00 LT, however, this is not reflected in the specific humidity (see section 3.3). Finally, the transition from clear to cloudy conditions, as determined by the the transition from dry turbulence to moist convection in DALES, occurs at 9:45 LT and is accompanied by a rapid increase in cloud cover (see section 3.4).

4.2 Future outlook

An important aspect of the Amazonian climate is the transition from shallow to deep convection, causing precipitation and potentially convective storm formation. Investigating the potential drivers of the shallow to deep convection transition using the DALES case designed in this research is the main future research goal. Several variables may drive, or contribute in driving the transition from shallow to deep convection during the Amazonian dry season. Deep convection during daytime is often triggered by local processes (Ghate and Kollias, 2016), thereby being directly influenced by surface properties such as surface turbulent flux partitioning. For the atmospheric variables, the 0–3 km and 0–6 km vertical wind shears seem important contributors in the transition during the dry season (but have little or no impact during the wet or transition seasons; Zhuang et al., 2017). Also, substantial boundary layer and mid-tropospheric moisture content is required for deep convection (Zhuang et al., 2017), and thereby moisture advection is a likely important contributor to deep convection. In De Feiter (2023), both vertical wind shear and moisture advection have been identified as important drivers of the shallow to deep convection transition based on the same observational data of the CloudRoots-Amazon22 campaign of this research, in addition to conceptual modelling. The next step, as also indicated in De Feiter (2023), is to use a LES case of a representative shallow cumulus day to identify the drivers of the transition to deep convection using sensitivity analysis. As discussed, the sensitivity analysis should then include vertical wind shear, moisture advection, and surface properties such as surface turbulent flux partitioning, and can be performed using the DALES case designed in this research. Furthermore, this DALES case can be employed for more applications of Amazonian research. For example, it is currently being used to investigate CO_2 transport by shallow cumulus clouds, and will additionally be used as a basis for assessing the difference in performance when using a bulk and multi-layer canopy scheme in DALES. Finally, the DALES case can be used as a validation tool for larger-scale atmospheric models with coarser resolutions.

Acknowledgements

I would like to thank my supervisor, Prof. Jordi-Vilà-Guerau de Arellano, for his extensive help with this MSc thesis and frequent meetings to discuss my progress and further steps, despite his busy schedule. Furthermore, I would like to thank Vincent de Feiter for his help with analysing the data from the CloudRoots-Amazon22 campaign, which he pioneered and for which he conducted the bulk of work. Additionally, I would like to thank Stephan de Roode for discussing the results from the DALES experiment and suggesting additional steps for improving its performance. Also, I would like to thank my supervisor and Stephan de Roode for the opportunity to present the results from this thesis, in addition to the work of Vincent de Feiter, at the Buys Ballot Research School - Boundary layer & Turbulence symposium at TU Delft. Finally, I would like to thank Guido Giovanelli-Haytzmann, Raquel González Armas and Daniel Rikkers for useful discussions on the research.

Chapter 5

Appendix

5.1 Tower data pre-processing

This section describes a large part of the pre-processing of tower data from the CloudRoots-Amazon22 campaign (Vilà-Guerau de Arellano et al., 2024). The remaining parts, in addition to the pre-processing of vertical profiles and cloud & BL height data, was conducted by Vincent de Feiter. Further details on the pre-processing and data availability of the CloudRoots-Amazon22 campaign will be published at some point in the future (de Feiter and de Haas, 2024, in preparation). As mentioned in the methodology (section 2.1), a thermohygrometer recorded temperature (T) and relative humidity (RH) data, which we have used to replace the T and RH data from the EC. This is due to the improved data quality of the thermohygrometer, an elaborate discussion on the argumentation behind this will be provided in (de Feiter and de Haas, 2024, in preparation). In short, the most important reason is the potential temperature mixing in the BL, which is considerably better for the thermohygrometer data compared to the EC, indicating better data quality.

We have excluded EC data at heights of 127, 196 and 316 m. This data has been omitted due to the θ and q from this data deviating considerably from the other tower heights. This was evident from both vertical profiles and time evolutions of these variables, where the discrepancy with data at other heights was unrealistically large. The cause behind the poor data quality at these heights is unclear.

For the tower data of the CloudRoots-Amazon22 campaign, some variables are only available for specific tower heights, while others are available at (almost) all altitudes on which data was collected. In table 5.1, some frequently used variables are listed, where a cross ($\times$) indicates that data of the listed variables is available at the corresponding height. The T and RH humidity have been collected with the thermohygrometer data, while the rest of the variables have been obtained from the EC. In the table, the "I" after the tower heights indicates that the data is from INSTANT, while data that has been collected with ATTO is marked with an "A". The tower heights that suffer from poor data quality are indicated with E (excluded), although note that this is not the case for T and RH, as these have been collected with the thermohygrometer rather than the EC. Note that more tower data is available, a description of which will be included in de Feiter and de Haas (2024, in preparation).

Height (m)	T	RH	CO_2	CO_2 flux	LE	H	P	ρ	e	e_s	u	v	TKE	u_*
5 (I)	×		×	×	×	×		×	×	×	×	×	×	×
15 (I)	×					×	×	×		×	×	×	×	×
25 (I)	×	×	×	×	×	×		×	×	×	×	×	×	×
35 (I)	×	×				×		×		×	×	×	×	×
43 (A)	×	×				×	×	×		×	×	×	×	×
50 (I)	×	×	×	×	×	×		×	×	×	×	×	×	×
81 (I)	×	×	×	×	×	×	×	×	×	×	×	×	×	×
100 (A)	×	×				×		×		×	×	×	×	×
127 (A, E)	×	×				×		×		×	×	×	×	×
151 (A)	×	×				×		×		×	×	×	×	×
172 (A)	×	×				×		×		×	×	×	×	×
196 (A, E)	×	×	×	×	×	×		×	×	×	×	×	×	×
223 (A)	×	×				×		×		×	×	×	×	×
247 (A)	×	×				×		×		×	×	×	×	×
298 (A)	×	×				×		×		×	×	×	×	×
316 (A, E)	×	×	×	×	×	×	×	×	×	×	×	×	×	×

TABLE 5.1: Availability of some important directly measured variables (so not inferred from other variables) at all tower heights, where a cross (×) indicates the variable is available at the corresponding height. The T and RH were collected with a thermohygrometer, while the remaining variables were obtained from the EC. The (I) after tower heights indicates the data is from the INSTANT tower, while (A) heights are data recorded by ATTO. The E (excluded) means that the EC data suffers from poor data quality and has therefore been excluded, so this includes all variables except T and RH. Please note that more tower data is available, this is just a sample of some of the more frequently used variables. The ρ is air density, e the vapor pressure and e_s the saturation vapor pressure, the other symbols are provided in the table at the beginning of this thesis (page iv).

The CO_2 data in table 5.1 has been collected using the Eddy-covariance (EC) method. For CO_2 specifically, this means the values are inaccurate and cannot be trusted. Instead, this data should only be used to investigate CO_2 trends. Therefore, picarro data has instead been used in this research (see Fig. 5.8), which consists of reliable CO_2 mixing ratios.

Pressure interpolation

For the tower data, pressure is only available at 15, 43, 81 and 316 m. For the conversion from temperature to potential temperature, the pressure at the remaining heights is required (see equation 5.1). Therefore, the pressure is interpolated by using a linear fit for every time step (30 minute intervals), based on the four heights where pressure is available. Two examples of the linear fits at two random time steps are shown in figure 5.1, which indicate that the interpolation is suitable for inferring pressures at additional heights.

To confirm the validity of the pressure interpolation, Δp is compared between the hydrostatic equation and interpolated pressure for heights where tower data is available, which is shown in figure 5.2. The hydrostatic equation used in the figure is given by $\Delta p = -g\rho\Delta z$, where g is the gravitational constant, ρ the density and Δz the vertical difference between two subsequent tower heights. The 127 m data for the hydrostatic equation is not included, as the data at this height has been omitted due to data quality issues, as mentioned. The

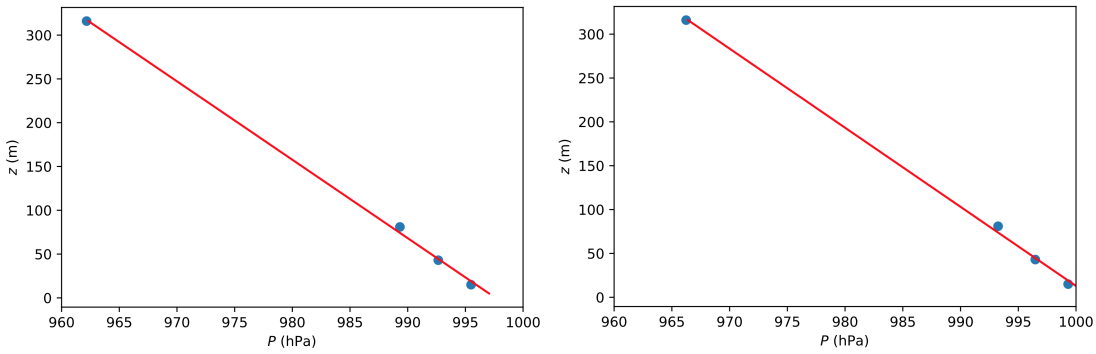

FIGURE 5.1: Two randomly selected linear pressure (P) fits for two separate days. The left fit is based on pressure data taken at 6 local time and the right panel fit corresponds to 11 local time.

figure shows excellent agreement between the hydrostatic equation and Δp from the interpolated pressure for the two random time steps that are shown (and for other random time steps that are not displayed here), confirming the validity of the pressure interpolation method. Using the interpolated pressure, the potential temperature (θ) is then calculated using

$$\theta = T \cdot (p_0/p)^{R_d/c_p} \tag{5.1}$$

where p_0 is the standard reference pressure (typically 1000 hPa), R_d is the specific gas constant for air and c_p specific heat of dry air at constant pressure. An alternative equation for calculating θ with height instead of pressure has also been tested, and the two methods have been compared. From the comparison, it has been concluded that the two methods only show minor differences for the tower data. An elaborate description of this comparison will be included in de Feiter and de Haas (2024, in preparation).

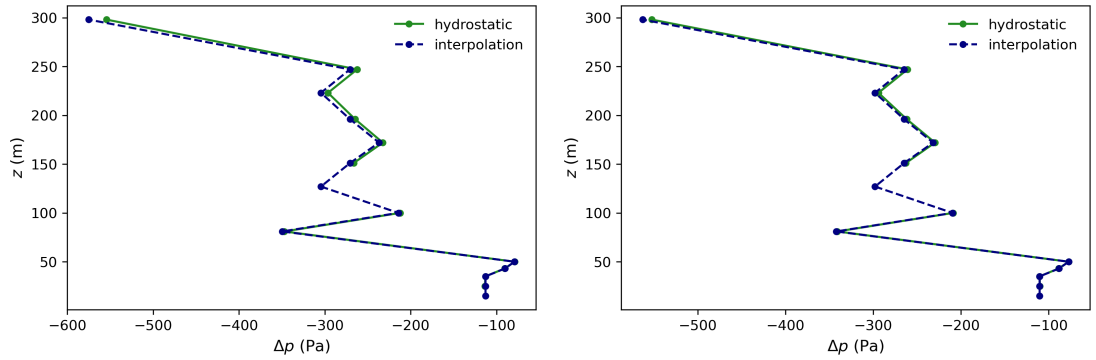

FIGURE 5.2: Two randomly selected Δp vertical profiles for the interpolated pressure and hydrostatic equation, where the 127 m height has been excluded due to data quality issues (see text). The two profiles show excellent agreement, confirming the validity of the pressure interpolation method.

Recalculations from thermohygrometer data

Due to the improved data quality of the thermohygrometer compared to the EC, the thermohygrometer T and RH have been used to infer other variables, rather than using the available EC data of these variables. This is done in a step-wise process for each height, where we initially calculate the saturation vapor pressure using

$$e_s = 611.2 \cdot \exp\left(\frac{17.62(T - 273.15)}{(-30.03 + T)}\right), \tag{5.2}$$

from which we determine the vapor pressure

$$e = \frac{\mathrm{RH}}{100} \cdot e_s \tag{5.3}$$

where RH is in %. Consequently, we infer the water vapor density

$$\rho_v = \frac{e}{R_v T} \tag{5.4}$$

and the density of dry air

$$\rho_d = \frac{e - p}{R_d T}, \tag{5.5}$$

from which we calculate the new air density

$$\rho_{\mathrm{th}} = \rho_v + \rho_d, \tag{5.6}$$

which apart from the pressure is inferred only from the thermohygrometer data, hence the subscript th. Subsequently, specific humidity is determined by

$$q = \frac{\rho_v}{\rho_d}, \tag{5.7}$$

which is the q that is used in Fig. 3.4 and 3.5. Additionally, the fraction of thermohygrometer air density over EC air density, given by

$$\mathrm{cf} = \frac{\rho_{\mathrm{th}}}{\rho_{\mathrm{EC}}}, \tag{5.8}$$

is used as a correction factor (cf) for the EC fluxes such as latent heat by

$$\mathrm{LE}_{\mathrm{th}} = \frac{\rho_{\mathrm{th}}}{\rho_{\mathrm{EC}}} \mathrm{LE}_{\mathrm{EC}}. \tag{5.9}$$

The density fraction of equation 5.8 has additionally been used for correcting the sensible heat flux and CO_2 flux.

5.2 Collapse of wind u- and v-components

In Fig. 5.3, the zonal wind component u is plotted at six different heights from tower data for a full 24-hour cycle on two different shallow cumulus days (2022-08-10; left panel and 2022-08-17; right panel). The figure highlights the difference in degree of BL mixing of u between the two days, showing well-mixed conditions on 2022-08-10, but only weak mixing on 2022-08-17. This is also reflected in the mixing of the meridional wind component v on the same shallow cumulus days, shown in Fig. 5.4. Therefore, one shallow cumulus day shows well-mixed BL conditions for both u and v, while the other is characterized by weak mixing. These u and v figures, along with similar ones for four other shallow cumulus days (not shown in this thesis), highlight that the degree of BL mixing for u and v varies substantially per day.

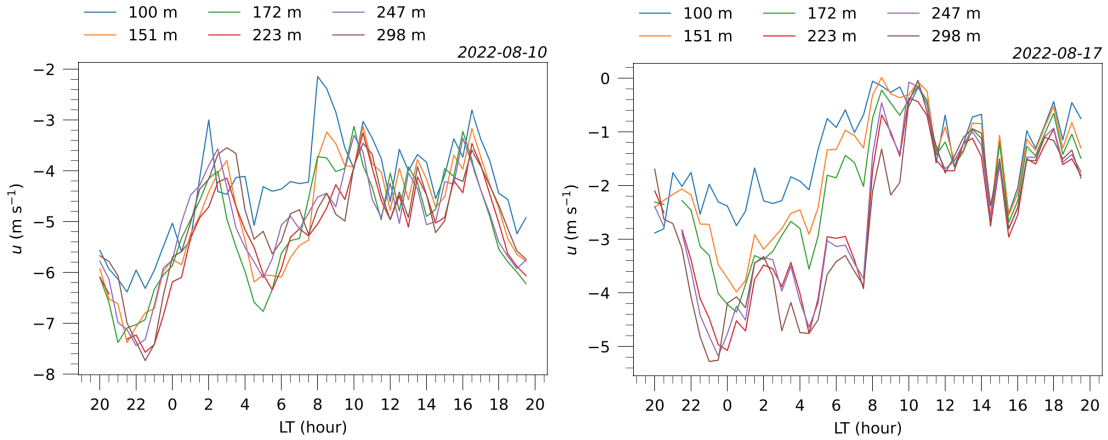

FIGURE 5.3: The 24-hour cycle of zonal wind component u for tower data from six different heights for two different shallow cumulus days (2022-08-10; left panel and 2022-08-17; right panel).

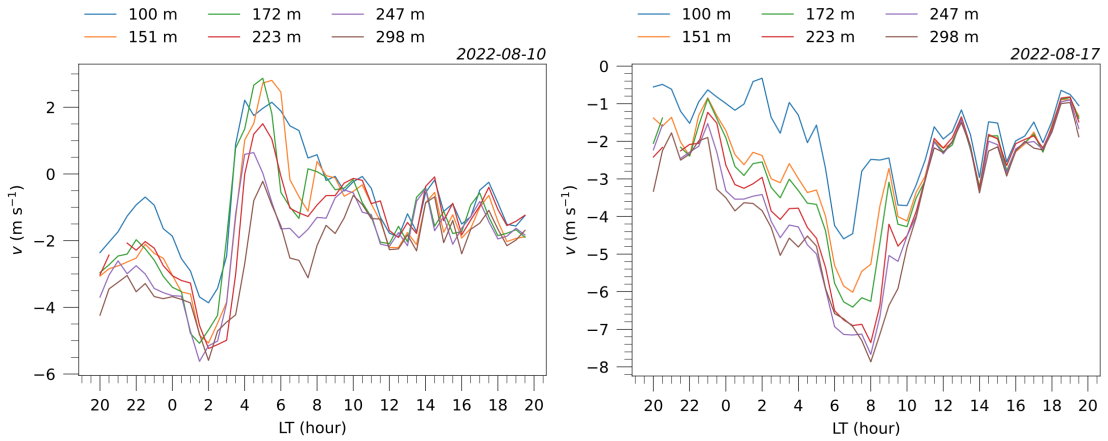

FIGURE 5.4: The 24-hour cycle of meridional wind component v for tower data from six different heights for two different shallow cumulus days (2022-08-10; left panel and 2022-08-17; right panel).

5.3 Diurnal variability of wind, CO_2 and radiation components

Radiation components

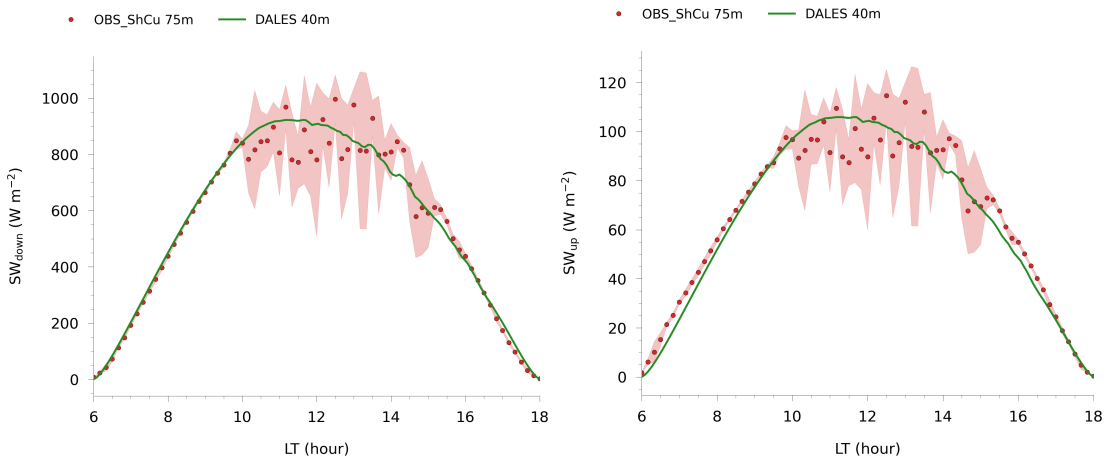

FIGURE 5.5: Diurnal cycle of the shortwave down component (SW_{down}, left panel) and shortwave up component (SW_{up}, right panel) for observational data taken during shallow cumulus days and DALES. The ∼40 m difference in height between DALES and the observations is due to the bulk canopy scheme, meaning the canopy top (which is ∼40 m) of the observations corresponds to the 0m-level in DALES. The observational data are 10-minute intervals, while DALES consists of 5-minute intervals. The observational data is aggregated over six shallow cumulus days (OBS_ShCu), where the red shading represents one standard deviation of the aggregate.

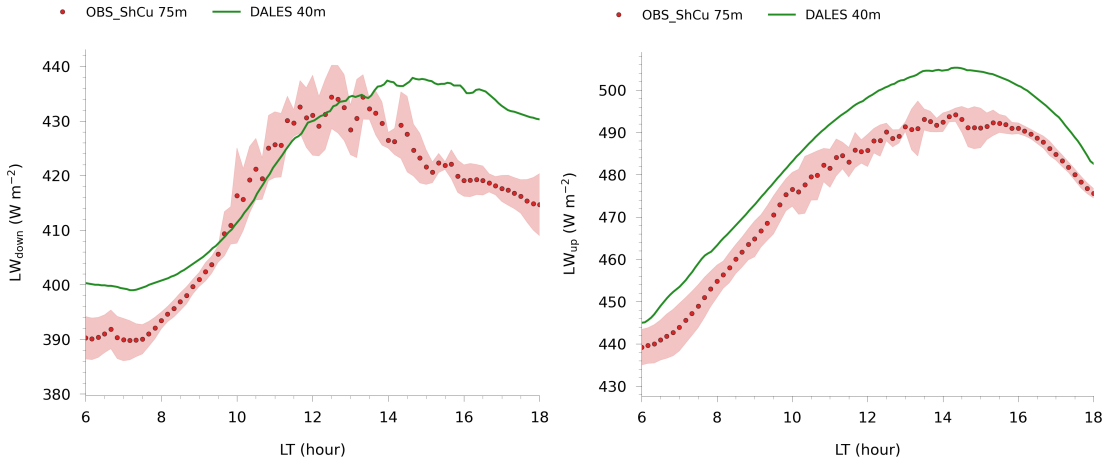

FIGURE 5.6: Diurnal cycle of the longwave down radiation component (LW_{down}, left panel) and longwave up component (LW_{up}, right panel) for observational data taken during shallow cumulus days and DALES. The ∼40 m difference in height between DALES and the observations is due to the bulk canopy scheme, meaning the canopy top (which is ∼40 m) of the observations corresponds to the 0m-level in DALES. The observational data are 10-minute intervals, while DALES consists of 5-minute intervals. The observational data is aggregated over six shallow cumulus days (OBS_ShCu), where the red shading represents one standard deviation of the aggregate.

In Fig. 5.5, the radiation components of shortwave up (left panel) and shortwave down (right panel) are shown for DALES and the shallow cumulus observations during the diurnal cycle. The figure highlights the excellent agreement between DALES and the observations for both the shortwave up and down components. The variability in the observations of radiation components around mid-day is not reflected in DALES, but this due to slab-averaging (see section 2.2). In Fig. 5.6, the diurnal variability of the longwave up (left panel) and longwave down (right panel) components are shown, again for DALES and the shallow cumulus observations. The observations of longwave radiation components are not well-captured by DALES, showing significant discrepancies for both longwave up and down during the diurnal cycle. This could be caused by a discrepancy in vertical profiles of q between DALES and the observations (see Fig. 3.5) and/or inaccuracies in DALES surface properties.

Carbon dioxide

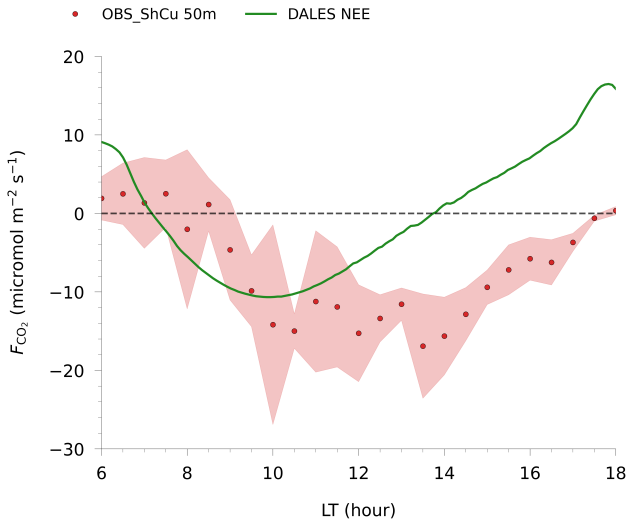

FIGURE 5.7: Diurnal variability of the CO_2 flux (F_{CO_2}) for observational data taken during shallow cumulus days and DALES. The DALES CO_2 flux (NEE; net ecosystem exchange) is from the 0m-level in the model. The ∼40 m difference in height between DALES and the observations is due to the bulk canopy scheme, meaning the canopy top (which is ∼40 m) of the observations corresponds to the 0m-level in DALES. The observational data are 30-minute intervals, while DALES consists of 5-minute intervals. The observational data is aggregated over six shallow cumulus days (OBS_ShCu), where the red shading represents one standard deviation of the aggregate. The diurnal cycle of F_{CO_2} is clearly misrepresented by DALES.

The diurnal cycle of CO_2 flux (or NEE; net ecosystem exchange) is shown in Fig. 5.7 for DALES and the observations. The figure highlights the misrepresentation of the diurnal variability of CO_2 flux by DALES, especially during the afternoon. This is likely rooted in the simplified bulk canopy scheme used in DALES, and may therefore require a more complex canopy representation to attain a reproduction of CO_2 flux. This is a topic of future research, where the implementation of a multi-layer canopy in DALES will be investigated for the Amazonia. In Fig. 5.8, the diurnal variability of CO_2 mixing ratio is shown for DALES and the observations at 53 m (left panel) and 79 m (right panel). The

figures show that DALES does perform reasonably well in reproducing the observations of CO_2 mixing ratio when excluding model spin-up (after ∼8:00 LT).

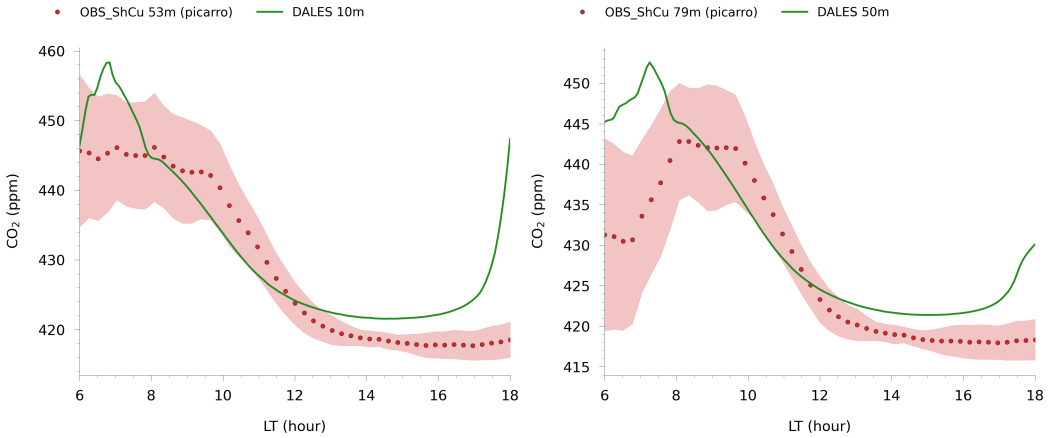

FIGURE 5.8: Diurnal cycle of the CO_2 mixing ratio for observational data taken during shallow cumulus days and DALES at a height of 53 m (left panel) and 79 m (right panel). The ∼40 m difference in height between DALES and the observations is due to the bulk canopy scheme, meaning the canopy top (which is ∼40 m) of the observations corresponds to the 0m-level in DALES. The observational data are 30-minute intervals, while DALES consists of 5-minute intervals. The observational data is aggregated over six shallow cumulus days (OBS_ShCu), where the red shading represents one standard deviation of the aggregate.

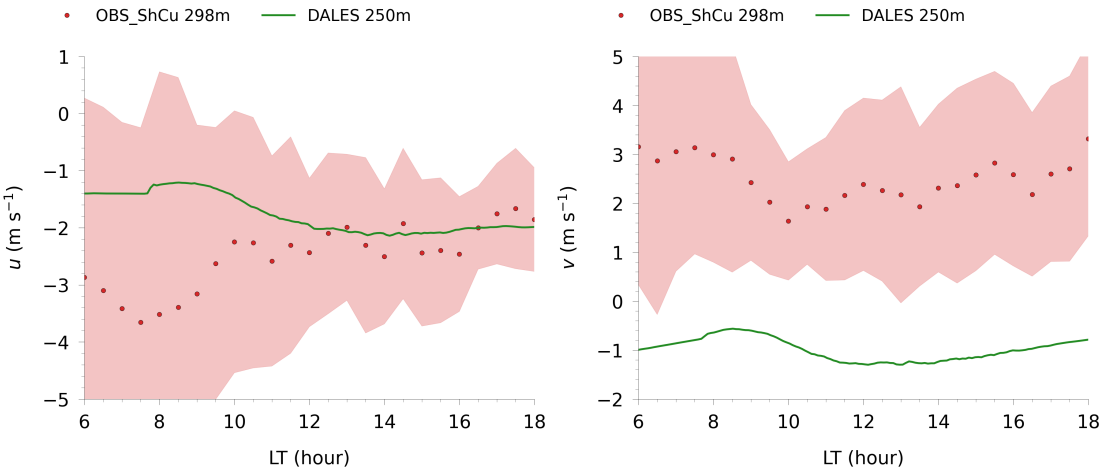

FIGURE 5.9: Diurnal cycle of the u (zonal) and v (meridional) wind components for observational data taken during shallow cumulus days and DALES at a height of 298 m. The ∼40 m difference in height between DALES and the observations is due to the bulk canopy scheme, meaning the canopy top (which is ∼40 m) of the observations corresponds to the 0m-level in DALES. The observational data are 30-minute intervals, while DALES consists of 5-minute intervals. The observational data is aggregated over six shallow cumulus days (OBS_ShCu), where the red shading represents one standard deviation of the aggregate. The text provides an explanation for the large discrepancy between DALES and the observations for v.

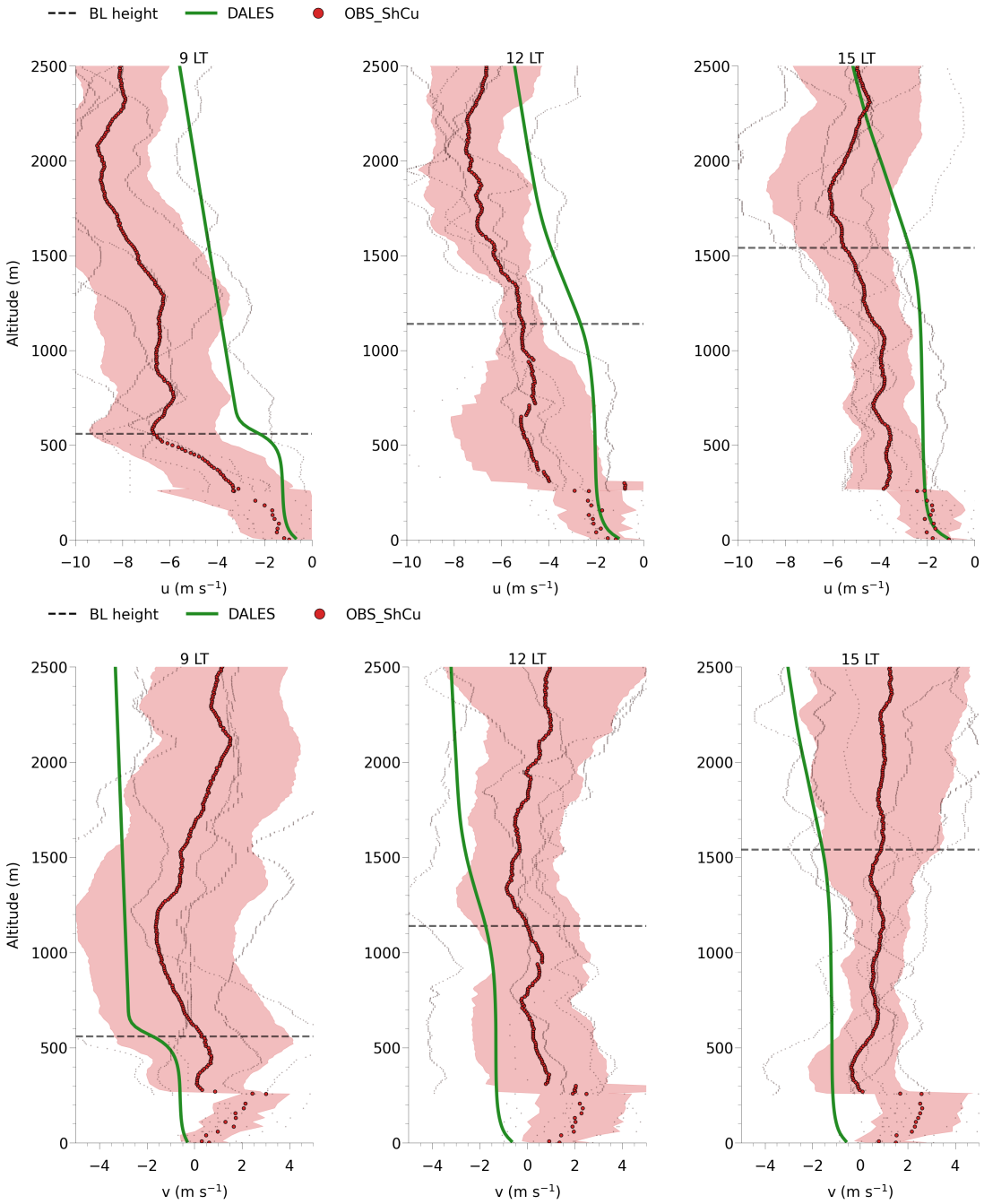

FIGURE 5.10: Vertical profiles up to 2500 m of the u (zonal) and v (meridional) wind components for the observations and DALES at 9:00, 12:00 and 15:00 local time (LT). The lower 260 m of the observational profiles consists of tower data (30-minute intervals), while above 260 m the data is from radiosondes. The observational data shown in red is an aggregate of six-day shallow cumulus days (OBS_ShCu), where the red shading represents one standard deviation of the aggregate. The grey points are individual profiles from the aggregate. The BL heights are inferred from a ceilometer. The DALES profiles are 30-minute averaged (over six 5-minute intervals).

Wind u- and v-components

In Fig. 5.9, the diurnal cycle of wind components u (left panel) and v (right panel) are shown for DALES and the observations at 298 m. For u, DALES and the observations agree well during the afternoon, although not in the morning. For v, DALES is not able to reproduce the observations, even showing an opposing sign with respect to the observations. This is however caused by the large discrepancy between the DALES initial v profile and the observations in the lower ∼260 m of the BL (see Fig. 5.13). This discrepancy for the lower ∼260 m could be related to measurements errors in tower v data, as the vertical profile of v for the observations shows an unrealistically large discontinuity between the tower and radiosonde data. In Fig. 5.10, the vertical profiles of the wind u and v components are shown at 9:00, 12:00 and 15:00 LT for DALES and the observations. Note that for the u observational profiles at 12:00 and 15:00 LT, a large discontinuity between the tower and radiosonde data is also present, indicating that the u tower data may also suffer from measurement errors. Additionally, the difference between the DALES wind profiles and the observations is substantial throughout the figures. For future research, the initial wind profiles in DALES should be closer to the observations to check if this improves the reproduction. Furthermore, as we currently do not know what causes the large discontinuities between tower and radiosonde data for u and v, this should also be further looked in to.

5.4 DALES cloud evolution

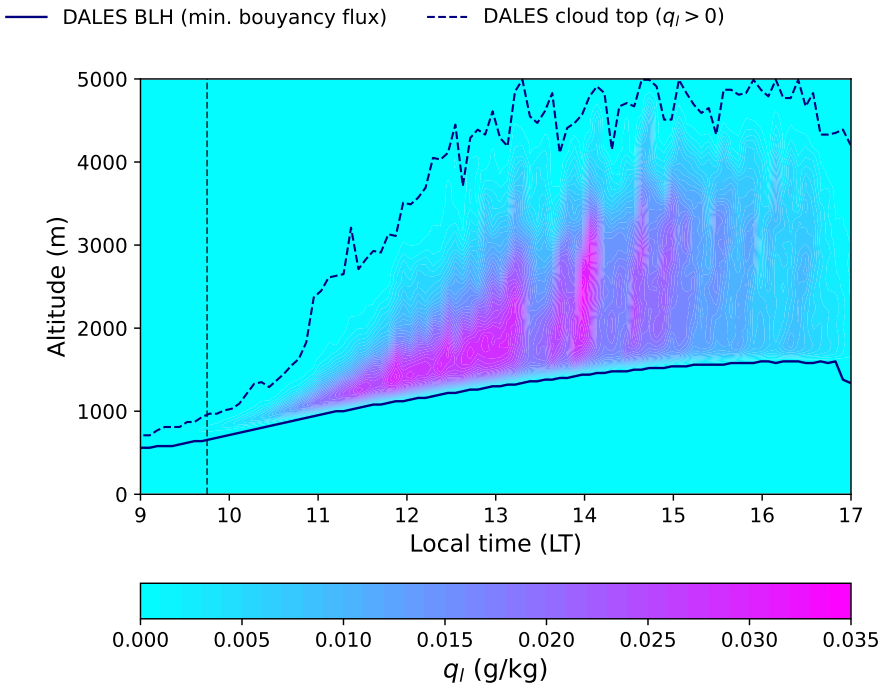

FIGURE 5.11: DALES contour plot of the liquid water specific humidity q_l between 9 and 17 local time (LT). Additionally, the BL height (from the minimum buoyancy flux) and cloud top (highest altitude at which $q_l > 0$) are shown.

In Fig. 5.11, a contour plot of the liquid water specific humidity for DALES is shown, including the BL height and cloud top. As mentioned, the DALES output is slab-averaged (see section 2.2), meaning the contour plot shows horizontally-averaged q_l. The BL height is again determined using the minimum buoyancy flux while the cloud top is the highest altitude at which $q_l > 0$. The contour plots shows that the cloud base rises along with the BL height as the day progresses (similar to Fig. 3.6, left panel). The figure highlights the evolution of the cloud top and horizontally-averaged q_l, and thereby the growth of the vertical extent of the cumulus clouds during the diurnal cycle, in addition to their dissipation. The cloud top increases rapidly during the late morning and early afternoon, even reaching the top of the domain (5 km). The clouds dissipate towards the end of the afternoon, as shown by the decreasing values of q_l (which is also reflected in the cloud cover, see right panel of Fig. 3.6).

5.5 Logbook of DALES runs

This is a logbook that lists the modifications made in DALES between different runs, including the argumentation behind the changes. This logbook represents the process of designing the DALES case, with the purpose of improving the reproduction of the measurements. Run 003 is the DALES base case adopted in Vilà-Guerau de Arellano et al. (2020).

Run 004: Copied the experiment set-up of run 003, changed julian day from 253 to 226, changed vertical profiles of θ, q, u and v (based on idealized profiles from a conceptual model, which in turn are based on observations).

Run 005: Copied the experiment set-up of run 004, changed the albedo from 0.15 to 0.1 and changed soil temperatures from 294 to 297.37 K (layer 1), 294.5 to 297.87 K (layer 2), 295 to 298.37 K (layer 3), 296 to 299.37 K (layer 4) and 297 to 300.37 K (bottom layer). The albedo was changed as the shortwave up radiation was too large compared to the observations. The 3.37 K was added to every soil layer to make sure the gradient between the upper soil layer and lower atmosphere (and between the different soil layers) are consistent between runs.

Run 006: Copied the experiment set-up of run 005, changed albedo to 0.115. The albedo was changed again as the shortwave up radiation was now too small compared to the observations.

Run 007: Copied the experiment set-up of run 005, changed albedo to 0.115 and changed vertical profiles of q and θ. As mentioned, the albedo was changed (compared to run 005) as the shortwave up radiation was now too small compared to the observations. The vertical profiles of q and θ were modified to closely match the observations (which were initially based on idealized profiles from a conceptual model as mentioned, therefore not resembling the observations).

Run 008: Copied the experiment set-up of run 007, changed the vertical profile of q to reduced values closer to the q of the conceptual model profile adopted in runs 004 – 006 for the lower ∼1800m of the troposphere, above ∼1800m same as run 007, changed soil layer temperatures by 0.37 K. The soil layer temperatures were again changed to be consistent with the soil-atmosphere gradient of 3 K, which should have been done in run 007 when changing the θ profile. The vertical profile of q was reduced as the simulation did not run well, showing a cloud cover of 100% from ∼12 LT onwards, and also a liquid water path that was too large, in addition to an unrealistic boundary layer height and cloud base and top.

Run 009: Copied the experiment set-up of run 008, changed the vertical profile of q to further reduced values closer to the q of the conceptual model profile adopted in run 004 – 006 for the lower ∼1800m of the troposphere, above ∼1800m same as run 007 and run 008. The vertical profile of q was further reduced as the simulation again did not run well, showing the same issues as run 008 such as the cloud cover of 100% from ∼ 12 LT onwards and unrealistically large liquid water path.

Run 010: Copied the experiment set-up of run 009, changed the vertical profile of θ back to the one from the conceptual model (same as run 004 – 006). This was done to investigate the possibility of θ causing the issue with the cloud cover, BL height, liquid water path and cloud base and top. These issues were avoided, which was therefore caused by the θ vertical profile close to the observations.

Run 011: Copied the experiment set-up of run 010, changed the vertical profile of q back to the one close to the observations for the full vertical profile. This was done to exclude the possibility that the q profile may cause similar or other issues when using the one close to the observations. The DALES run performed well in simulating cloud characteristics and BL height, so using a q profile that resembles the radiosonde data does not cause similar or other issues.

Run 012: Copied the experiment set-up of run 011, changed upper (>300m) vertical profile for DALES of θ due to coding error (missing incident) that caused a slight change in theta profile of observations at 6 LT. Also slightly changed q profile for DALES in the lower ∼1700m to make it more consistent with observational data. Most importantly, the lower 300m of θ and q profiles were modified to match the new observational data better, where I used an exponential fit for θ and linear fit for q. Due to the gap between tower and radiosonde data, the lower ∼300 m profile for this run for both θ and q were horizontally shifted so it aligned with the radiosonde data. The the soil temperatures were also modified to be consistent with the soil-atmosphere θ gradient of 3 K. The same issues compared to earlier runs are reported, as the DALES run again showed a 100% CC and unrealistically large liquid water path and cloud base and top.

Run 013: Copied the experiment set-up of run 012, again shifting the lower ∼300m profile of θ and q to match the radiosondes, instead connecting the lower ∼300m to a radiosonde data point at a higher altitude for both θ and q. The the soil temperatures were also

modified to be consistent with the soil-atmosphere temperature gradient of 3 K. The same issues compared to earlier runs and run 0012 are reported, again showing a 100% CC and unrealistically large liquid water path and cloud base and top.

Run 014: Copied the experiment set-up of run 013, changed the θ profile to a new profile from the conceptual model, generated based on new thermohygrometer observational data. The soil temperatures were again modified to be consistent with the soil-atmosphere temperature gradient of 3 K. The same issues compared to earlier runs and run 0013 are reported, again showing a 100% CC and unrealistically large liquid water path and cloud base and top.

Run 015: Copied the experiment set-up of run 013, kept the θ profile of run 013 that consists of an exponential fit through the new thermohygrometer tower data, however changing the profile from ∼300m onwards to a single sloped line with a certain lapse rate that is very close to the earlier θ conceptual model profile of run 004 – 006 and run 011, since these simulations did not show the earlier reported issues. The lower ∼1000m of q was changed slightly to be consistent with tower and radiosonde data (the linear fit was moved away from the tower data by 0.35 g/kg, from there it resembles the radiosonde data). The soil temperatures were already consistent with the soil-atmosphere temperature gradient of 3 K since the lower ∼300m profile of θ from run 013 was used.

Run 016: Copied the experiment set-up of run 015, but now extended the vertical domain from 4 to 5 km. The q profile was extended to continue resembling the radiosonde data. The θ profile was extended based on its lapse rate, which has also been done for the u and v initial profiles. The profile was extended to 5 km since the (maximum) cloud top (based on $q_l > 0$) reached the top of the vertical domain (4 km) in the previous simulations.

Run 017: Copied the experiment set-up of run 016, changed the initial CO_2 profile to one that start at 430 ppm from the surface until 250m, subsequently containing an inversion jump to 420 ppm, which is the value used for the rest of the profile (so up to 5 km). This was done in an attempt to more closely match the evolution of observational CO_2 data at 50 and 81m.

Run 018: Copied the experiment set-up of run 017, changed the lower ∼250 m of the initial θ profile to start at a surface value of 2 K lower compared to the previous runs (again using an exponential fit), after which it connects to the diagonal θ profile (>250m). This was done in an attempt to delay the onset of entrainment from the free troposphere to match the timing of the observations, but this did not work (it was only delayed by roughly 10 minutes).

Run 019: Copied the experiment set-up of run 018, started the simulation at 7:00 LT rather than 6:00 LT as before. This was again done in an attempt to delay the onset of entrainment to match the timing of the observations, but again did not work.

Run 020: Copied the experiment set-up of run 019, changed the lower ∼250m of the initial u and v profiles to an exponential profile starting from the surface and connecting to the diagonal (lapse rate) profiles. This was done in an attempt to more closely match the observations of u and v in the lower ∼300 m of the BL.

Run 021: Copied the experiment set-up of run 020, changed the lower ∼250 m of the initial profile to a constant $\theta = 300.3$ K profile with a ~ 1.5 K inversion jump. The soil temperatures were again modified to be consistent with the 3 K soil-atmosphere gradient. The lower θ profile was changed to an unstable profile in an attempt to delay the onset of entrainment. This worked well, but the onset of entrainment still occurred too early.

Run 022: Copied the experiment set-up of run 021, extended the constant $\theta = 300.3$ K initial profile from 250 m up to 390 m, and changed the exponential profiles of CO_2, u and v to constant profiles of $CO_2 = 445$ ppm, $u = 1.2$ m/s and $v = 0.7$ m/s. The CO_2, u and v were changed for consistency and in an attempt to improve the modelling of CO_2, the constant θ profile was extended in an attempt to further delay the onset of entrainment, which again did work but still occurred too early.

Run 023: Copied the experiment set-up of run 022, further extended the constant $\theta = 300.3$ K initial profile from 390 m up to 490 m, in addition to extending the constant profiles of CO_2, u and v profiles to 490 m while changing $u = 1.2$ to $u = 1.4$ m/s, $v = 0.7$ to $v = 1.0$ m/s and changing the CO_2 profiles above 490 m from 420 to 415 ppm. The θ profile was again extended to further delay the onset of entrainment, which now roughly matches the one of the observations. The constant CO_2, u and v profiles were extended up to 490 m for consistency with the θ profile, and the CO_2 profile above 490 m was reduced in an attempt to better simulate the evolution of CO_2. This is the run that has been used throughout this thesis.

Finally, we note that DALES has been run on the Dutch National Supercomputer Snellius (https://www.surf.nl/en/services/snellius-the-national-supercomputer), located in Amsterdam. A single run in DALES took $\sim 7 - 9$ hours on the Snellius supercomputer, therein simulating one diurnal cycle of 12 hours in DALES.

5.6 DALES initial & boundary conditions

In this section, the adopted model schemes and initial and boundary conditions of DALES are provided. In the left panel of Fig. 5.12, the initial vertical profile of potential temperature prescribed in DALES is given, including an observational profile based on radiosonde data. In the right panel of Fig. 5.12, the initial specific humidity profile implemented in DALES is shown, again including a comparison with observations. Contrary to the initial vertical profile of q, the DALES θ profile does not resemble the observations, which is discussed in section 3.6. As indicated in table 5.3, the CO_2 profile is constant at 445 ppm between 0 and 500m, with an inversion jump of 30 ppm to a constant 415 ppm profile above 500m. This constant CO_2 profile above 500m is a rough approximation based on aircraft data.

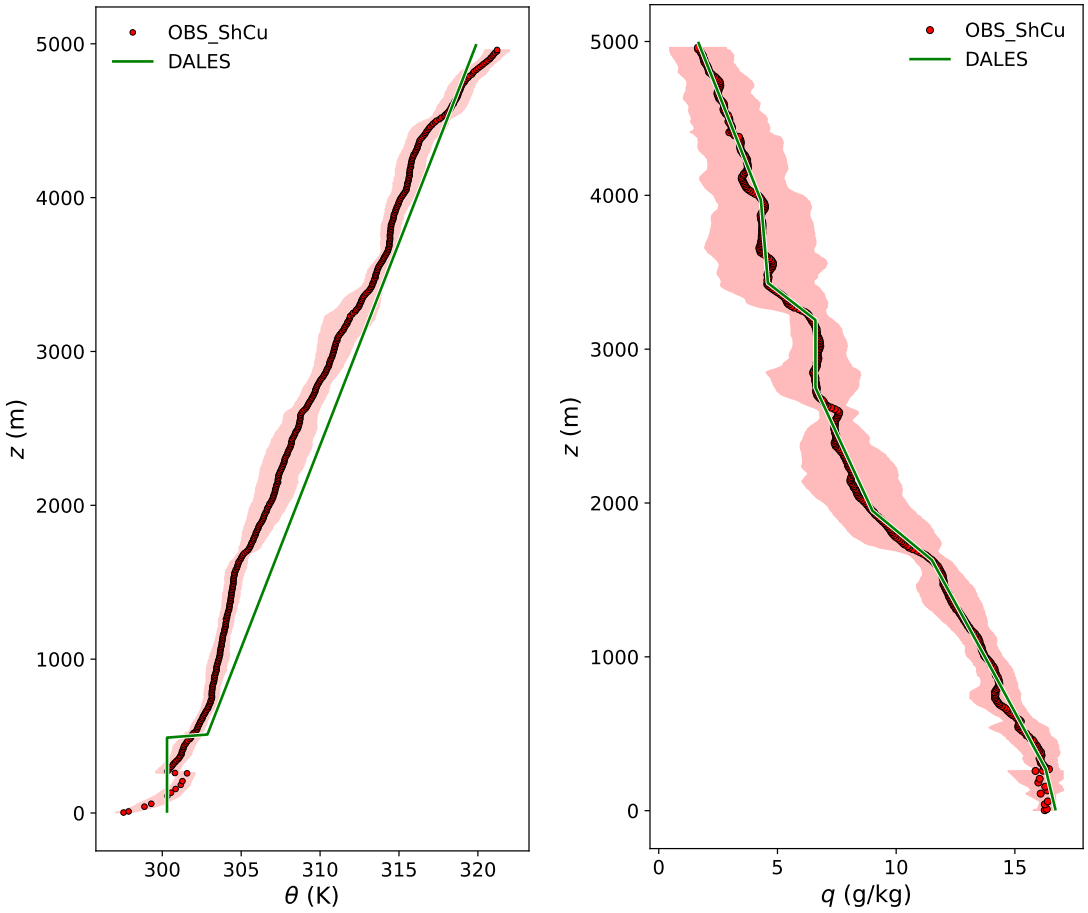

FIGURE 5.12: The initial vertical profiles implemented in DALES of potential temperature (left panel, θ) and specific humidity (right panel, q). The lower 260 m of the observational profiles consists of tower data (30-minute intervals), while above 260 m the data is from radiosondes. The observational data shown in red is an aggregate of six-day shallow cumulus days (OBS_ShCu), where the red shading represents one standard deviation of the aggregate.

In the below tables, further initial and boundary conditions are provided, in addition to the model schemes used in DALES. The tables are copied from or based on Vilà-Guerau de Arellano et al. (2020), as their DALES case is taken as a basis for the case of this research, as mentioned. In table 5.2, the model schemes that have been adopted in DALES are given. The DALES domain, numerical settings and initial vertical profiles of CO_2 and TKE are provided in table 5.3. Further initial and boundary conditions are listed in table 5.4, primarily related to surface properties. Finally, plant physiological and soil CO_2 efflux parameter values are provided in table 5.5.

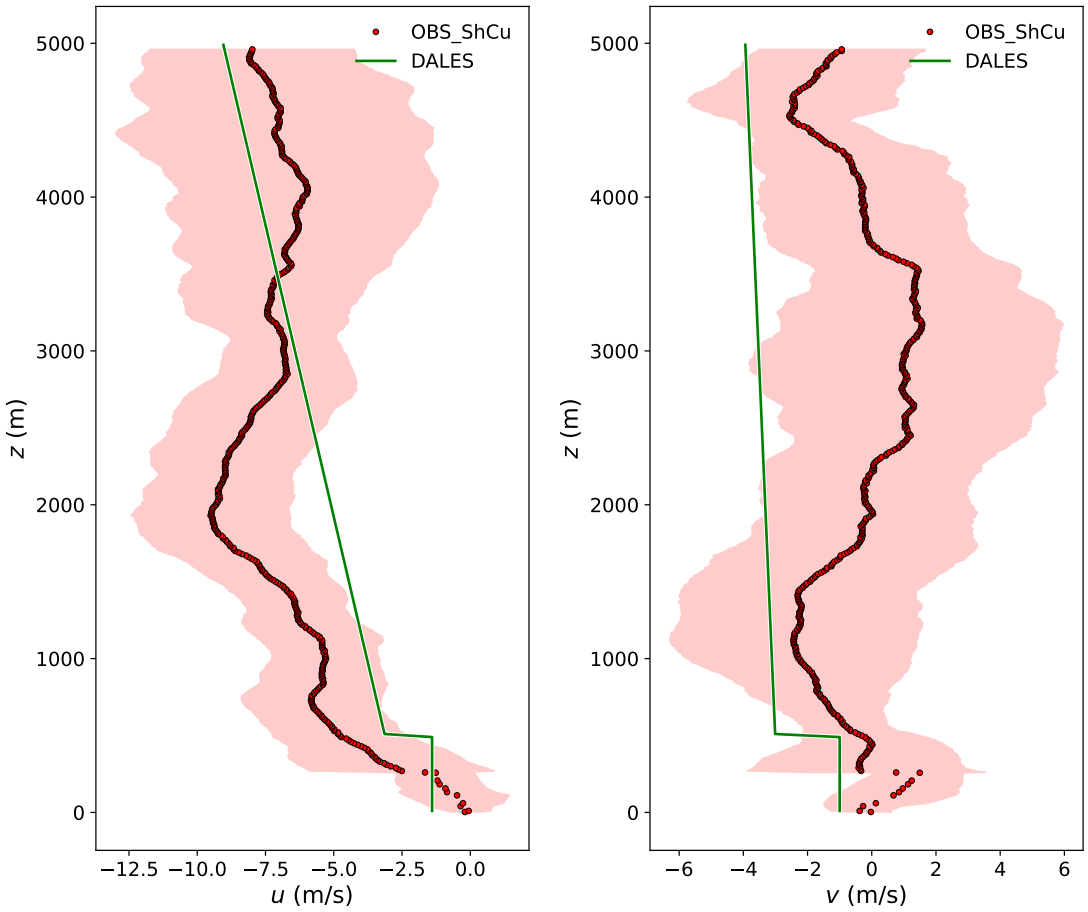

FIGURE 5.13: The initial vertical profiles implemented in DALES of the zonal wind component (left panel, u) and meridional wind component (right panel, v). The lower 260 m of the observational profiles consists of tower data (30-minute intervals), while above 260 m the data is from radiosondes. The observational data shown in red is an aggregate of six-day shallow cumulus days (OBS_ShCu), where the red shading represents one standard deviation of the aggregate.

TABLE 5.2: List of model schemes that have been adopted in DALES. RRTMG is short for Rapid Radiative Transfer Model. This table has been copied from Vilà-Guerau de Arellano et al. (2020).

Variables/process	*Model scheme*
Radiation	One-dimensional RRTMG (Pedruzo-Bagazgoitia et al., 2017)
Thermodynamics	Large-eddy simulation (Heus et al., 2010)
Microphysics	All-or-nothing (Heus et al., 2010)
CO_2 assimilation/Stomatal aperture	A-gs (Ronda, De Bruin, and Holtslag, 2001), (Pedruzo-Bagazgoitia et al., 2017)
Radiation/Stomatal aperture	Two big leafs: sunlit/shaded (Pedruzo-Bagazgoitia et al., 2017, Ronda, De Bruin, and Holtslag, 2001)
Surface fluxes	Pennman-Montheith constrained by surface energy balance (Heus et al., 2010)
Soil flux	Four-layer restore (Heus et al., 2010)

TABLE 5.3: DALES domain, numerical settings and prescribed initial vertical profiles of CO_2 and turbulent kinetic energy (TKE). CFL is short for Courant-Friedrichs-Lewy. A large part of the table has been copied from Vilà-Guerau de Arellano et al. (2020).

Domain	
Horizontal, (L_x, L_y)	(19080, 19080) m
Vertical, L_z	4990 m
Grid points, (N_x, N_y)	(360, 360)
Grid points, N_z	250
Resolution (Δx, Δy, Δz)	(53,53,20) m
Horizontal numerical scheme state variables	5^{th}-order
Vertical numerical scheme state variables	2^{nd}-order
Numerical scheme atmospheric constituent	$\kappa-$scheme
Boundary Conditions	
Horizontal	Periodic
Upper	Sponge layer
Lower	Coupled land surface model (see tables 5.4 and 5.5)
Time integration	
Time step, dt	variable, set by CFL criterion with maximum value 1 s
Integration time	43200 s
Averaging time statistics	300 s
Initial Conditions	
Subgrid turbulent kinetic energy profile	
0 m $< z <$ 90 m	TKE = 1.0 m^2 s^{-2}
90 m $< z <$ 110 m	TKE = 0.5 m^2 s^{-2}
110 m $< z <$ 4990 m	TKE = 0.0 m^2 s^{-2}
Carbon dioxide profile	
0 m $< z <$ 490 m	CO_2 = 445 ppm
490 m $< z <$ 4990 m	CO_2 = 415 ppm

TABLE 5.4: DALES initial and boundary conditions. A large part of the table has been copied from Vilà-Guerau de Arellano et al. (2020).

Geographic coordinates for radiation	
Latitude (degree)	-2.6091
Longitude (degree)	-60.2093
Julian day	226
Starting hour simulation	10 UTC (06 LT)
Roughness length momentum (z_{om})	0.5 m
Roughness length heat (z_{oh})	0.1 m
Surface properties	A-gs model (see table 5.5)
	Tropical forest
	Sandy Loam
Initial Surface Temperature (T_s)	290 K
Albedo	0.115
Skin conductivity	0.
Vegetation cover	0.9
Leaf Area Index	5
Soil properties	
Temperature soil layer 1 (T_{soil1})	297.30 K
Temperature soil layer 2 (T_{soil2})	297.80 K
Temperature soil layer 3 (T_{soil3})	298.30 K
Temperature soil layer 4 (T_{soil4})	299.30 K
Temperature soil deep layer (T_{deep})	300.30 K
Volumetric soil moisture content soil layer 1 (W_{soil1})	0.3 m^3 m^{-3}
Volumetric soil moisture content soil layer 2 (W_{soil2})	0.3 m^3 m^{-3}
Volumetric soil moisture content soil layer 3 (W_{soil3})	0.3 m^3 m^{-3}
Volumetric soil moisture content soil layer 4 (W_{soil4})	0.3 m^3 m^{-3}
Volumetric content wilting point (W_{fc})	0.171 m^3 m^{-3}
Volumetric content field capacity (W_{fc})	0.4 m^3 m^{-3}
Saturated water volumetric water (W_{sat})	0.5 m^3 m^{-3}
Skin conductivity	40 W m^{-2} K^{-1}

TABLE 5.5: Plant physiological and soil CO_2 efflux parameters from Ronda, De Bruin, and Holtslag (2001), Lloyd and Taylor (1994) and Vilà-Guerau de Arellano et al. (2015). This table has been copied from Vilà-Guerau de Arellano et al. (2020).

Plant type	Parameter	(T=298 K)	Q_{10}	T_1(K)	T_2(K)
C3	a_d (kPa^{-1})	0.07			
	α_o (mg J^{-1})	0.017			
	f_o (-)	0.89			
	K_x (m_{ground} m^{-1}_{leaf})	0.7			
	Γ (mg m^{-3}))	$68.5\rho_a$	1.5		
	g_{m298} (mm s^{-1})	7.0	2.0	278	301
	$A_{m,max}$(mg m^{-2} s^{-1})	2.2	2.0	281	311
	$g_{min,c}$(m s^{-1})	$2.5{\cdot}10^{-4}$			
Soil	Parameter				
	R10	0.15			
	Eact0	$53.3{\cdot}10^{3}$			

Bibliography

[1] Meinrat O Andreae et al. "The Amazon Tall Tower Observatory (ATTO): overview of pilot measurements on ecosystem ecology, meteorology, trace gases, and aerosols". In: *Atmospheric Chemistry and Physics* 15.18 (2015), pp. 10723–10776.

[2] Wayne M Angevine, Henk Klein Baltink, and Fred C Bosveld. "Observations of the morning transition of the convective boundary layer". In: *Boundary-Layer Meteorology* 101 (2001), pp. 209–227.

[3] Alan K Betts and Maria Assuncao F Silva Dias. "Progress in understanding land-surface-atmosphere coupling from LBA research". In: *Journal of Advances in Modeling Earth Systems* 2.2 (2010).

[4] Gordon B Bonan et al. "Moving beyond the incorrect but useful paradigm: reevaluating big-leaf and multilayer plant canopies to model biosphere-atmosphere fluxes–a review". In: *Agricultural and Forest Meteorology* 306 (2021), p. 108435.

[5] Roel JW Brienen et al. "Long-term decline of the Amazon carbon sink". In: *Nature* 519.7543 (2015), pp. 344–348.

[6] Peter M Cox et al. "Amazonian forest dieback under climate-carbon cycle projections for the 21st century". In: *Theoretical and applied climatology* 78 (2004), pp. 137–156.

[7] Vincent De Feiter. "CloudRoots: Moist Air Intrusion and Vertical Wind Shear as Limiting Factors for Deep Convective Development. M.S. thesis, Meteorology and Air Quality, Wageningen University, The Netherlands." In: (2023).

[8] VS de Feiter and SEM de Haas. "CloudRoots-Amazon22: creating a uniform dataset (in preperation)". In: (2024).

[9] Rudi J Van der Ent et al. "Origin and fate of atmospheric moisture over continents". In: *Water Resources Research* 46.9 (2010).

[10] Jeffrey M Freedman et al. "Boundary layer clouds and vegetation–atmosphere feedbacks". In: *Journal of Climate* 14.2 (2001), pp. 180–197.

[11] Virendra P Ghate and Pavlos Kollias. "On the controls of daytime precipitation in the Amazonian dry season". In: *Journal of Hydrometeorology* 17.12 (2016), pp. 3079–3097.

[12] Eliane Gomes Alves et al. "Intra-and interannual changes in isoprene emission from central Amazonia". In: *Atmospheric Chemistry and Physics* 23.14 (2023), pp. 8149–8168.

[13] Alice Henkes et al. "Morning boundary layer conditions for shallow to deep convective cloud evolution during the dry season in the central Amazon". In: *Atmospheric Chemistry and Physics* 21.17 (2021), pp. 13207–13225.

[14] Thijs Heus et al. "Formulation of and numerical studies with the Dutch Atmospheric Large-Eddy Simulation (DALES)". In: *Geoscientific Model Development Discussions* 3.1 (2010), pp. 99–180.

[15] GL Horn et al. "Cloud shading effects on characteristic boundary-layer length scales". In: *Boundary-Layer Meteorology* 157 (2015), pp. 237–263.

[16] John Lloyd and JA Taylor. "On the temperature dependence of soil respiration". In: *Functional ecology* (1994), pp. 315–323.

[17] Scot T Martin et al. "Introduction: observations and modeling of the Green Ocean Amazon (GoAmazon2014/5)". In: *Atmospheric Chemistry and Physics* 16.8 (2016), pp. 4785–4797.

[18] Roel AJ Neggers, J David Neelin, and Bjorn Stevens. "Impact mechanisms of shallow cumulus convection on tropical climate dynamics". In: *Journal of Climate* 20.11 (2007), pp. 2623–2642.

[19] HG Ouwersloot et al. "Large-eddy simulation comparison of neutral flow over a canopy: Sensitivities to physical and numerical conditions, and similarity to other representations". In: *Boundary-Layer Meteorology* 162 (2017), pp. 71–89.

[20] So-Won Park et al. "Role of cloud feedback in continental warming response to CO 2 physiological forcing". In: *Journal of Climate* 34.22 (2021), pp. 8813–8828.

[21] X Pedruzo-Bagazgoitia et al. "Direct and diffuse radiation in the shallow cumulus–vegetation system: Enhanced and decreased evapotranspiration regimes". In: *Journal of Hydrometeorology* 18.6 (2017), pp. 1731–1748.

[22] X Pedruzo-Bagazgoitia et al. "Investigating the Diurnal Radiative, Turbulent, and Biophysical Processes in the Amazonian Canopy-Atmosphere Interface by Combining LES Simulations and Observations". In: *Journal of Advances in Modeling Earth Systems* 15.2 (2023), e2022MS003210.

[23] Oliver L Phillips et al. "Drought sensitivity of the Amazon rainforest". In: *Science* 323.5919 (2009), pp. 1344–1347.

[24] David Pino et al. "Role of shear and the inversion strength during sunset turbulence over land: characteristic length scales". In: *Boundary-layer meteorology* 121 (2006), pp. 537–556.

[25] RJ Ronda, HAR De Bruin, and AAM Holtslag. "Representation of the canopy conductance in modeling the surface energy budget for low vegetation". In: *Journal of Applied Meteorology and Climatology* 40.8 (2001), pp. 1431–1444.

[26] A Pier Siebesma et al. "A large eddy simulation intercomparison study of shallow cumulus convection". In: *Journal of the Atmospheric Sciences* 60.10 (2003), pp. 1201–1219.

[27] M Sikma et al. "Interactions between vegetation, atmospheric turbulence and clouds under a wide range of background wind conditions". In: *Agricultural and Forest Meteorology* 255 (2018), pp. 31–43.

[28] DV Spracklen et al. "The effects of tropical vegetation on rainfall". In: *Annual Review of Environment and Resources* 43 (2018), pp. 193–218.

[29] Christopher J Still et al. "No evidence of canopy-scale leaf thermoregulation to cool leaves below air temperature across a range of forest ecosystems". In: *Proceedings of the National Academy of Sciences* 119.38 (2022), e2205682119.

[30] Wen-Yih Sun and Oliver M Sun. "Revisiting the parcel method and CAPE". In: *Dynamics of Atmospheres and Oceans* 86 (2019), pp. 134–152.

[31] Tanja Suni et al. "The significance of land-atmosphere interactions in the Earth system—iLEAPS achievements and perspectives". In: *Anthropocene* 12 (2015), pp. 69–84.

[32] CC Van Heerwaarden and J Vilà-Guerau de Arellano. "Relative humidity as an indicator for cloud formation over heterogeneous land surfaces". In: *Journal of the Atmospheric Sciences* 65.10 (2008), pp. 3263–3277.

[33] J Vilà-Guerau de Arellano et al. *Atmospheric boundary layer: Integrating air chemistry and land interactions*. 2015.

[34] J Vilà-Guerau de Arellano et al. "CloudRoots-Amazon22: Integrating clouds with photosynthesis by crossing scales". In: *Bulletin of the American Meteorological Society* (2024).

[35] J Vilà-Guerau de Arellano et al. "Interactions Between the Amazonian Rainforest and Cumuli Clouds: A Large-Eddy Simulation, High-Resolution ECMWF, and Observational Intercomparison Study". In: *Journal of Advances in Modeling Earth Systems* 12.7 (2020), e2019MS001828.

[36] Cort J Willmott. "Some comments on the evaluation of model performance". In: *Bulletin of the American Meteorological Society* 63.11 (1982), pp. 1309–1313.

[37] Zhaoying Zhang et al. "Large diurnal compensatory effects mitigate the response of Amazonian forests to atmospheric warming and drying". In: *Science Advances* 9.21 (2023), eabq4974.

[38] Yizhou Zhuang et al. "Seasonal variation of shallow-to-deep convection transition and its link to the environmental conditions over the Central Amazon". In: *Journal of Geophysical Research: Atmospheres* 122.5 (2017), pp. 2649–2666.